

The Science of Education

Clay M. Starlin

21st Annual International
Precision Teaching Conference
Rutgers University
November 8, 2008

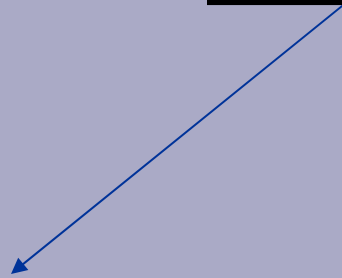


Accountability

- Should we be accountable in education?



Accountability



to count, to measure




Accountability



to count, to measure

**Educational accountability
means to measure, to be
*scientific.***



NCLB

The No Child Left Behind (NCLB) federal legislation uses the term “scientific/research based” 111 times in emphasizing the need for educational practice to become more scientific (aka: accountable, evidenced based).

However, the educational community does not agree on what it means to be scientific.

Science


Purpose of Science: to discern knowledge by distinguishing truth from falsehood.



Science

Purpose of Science: to discern knowledge by distinguishing truth from falsehood.

Effective Science Involves Two Major Components

- **Precise Measurement**
 - **Robust Research Methodology**
- 

Science

Purpose of Science: to discern knowledge by distinguishing truth from falsehood.

Effective Science Involves Two Major Components

- Precise Measurement
- Robust Research Methodology

Standardization of a Precise Measurement System is the Foundation for the Establishment of a Scientific Discipline

Science

Purpose of Science: to discern knowledge by distinguishing truth from falsehood.

Effective Science Involves Two Major Components

- Precise Measurement
- Robust Research Methodology

Standardization of a Precise Measurement System is Foundational to the Establishment of a Scientific Discipline

Components of Precise Measurement

1. Precisely Describe Phenomenon (e.g behavior) to Observe - Pinpoint
2. Accurately Record Significant Dimensions of Phenomenon - Record
3. Display Data to Facilitate Accurate Conclusions/Decisions - Chart

Pinpointing (Precise Description)



Pinpointing (Precise Description)

specify a behavior so that when ONE instance occurs, we can count it reliably.

after John Eshleman



Pinpointing (Precise Description)

specify a behavior so that when ONE instance occurs, we can count it reliably.

identifies ONE “movement cycle (MC)” (i.e. It has a start time, a do time, and a stop time.)

after John Eshleman

Pinpointing

Writing Precise Performance Objectives

1. Specify Behavior to Count

- specify count unit, attend to calibration



CALIBRATION IN ORAL READING

*Steps removed
from pure
calibration*

0	SYLLABLE
1	WORD
2	CLAUSE
3	PHRASE
4	SENTENCE
5	PARAGRAPH
6	PAGE
7	CHAPTER
8	BOOK

Pinpointing

Writing Precise Performance Objectives

1. Specify Behavior to Count

- specify count unit, attend to calibration
- use channel language (e.g. SeeSays word, HearWrites word)



MATRIX OF LEARNING CHANNELS

IN \ OUT	SAYS	WRITES	TYPES	POINTS	MARKS	CLICKS	DOES
SEE							
HEAR							
TOUCH							
SMELL							
TASTE							
THINK							
FEEL							

Pinpointing

Writing Precise Performance Objectives

1. Specify Behavior to Count

- specify count unit, attend to calibration
- use channel language (e.g. SeeSays word, HearWrites word)

2. Specify Conditions

(e.g. curriculum, setting, etc.)



Pinpointing

Writing Precise Performance Objectives

1. Specify Behavior to Count

- specify count unit, attend to calibration
- use channel language (e.g. SeeSays word, HearWrites word)

2. Specify Conditions

(e.g. curriculum, setting, etc.)

3. Specify Frequency Based Mastery Criteria

- Numeric Criteria (No Numbers → No Performance Objective)
- Wrong Numbers → Imprecise Objective (e.g. % correct)
- Criteria Result in Desired Outcomes: Retention, Endurance, Application (aka REAPS)

Writing Performance Objectives Sheet

Learner Name: _____

Channels	Singular Count Unit (e.g. word vs. words)	Conditions (e.g. setting, curriculum)	Mastery Criteria * (correct)	By

* Generally, only random error is acceptable for incorrect performance.

MATRIX OF LEARNING CHANNELS

IN \ OUT	SAYS	WRITES	TYPES	POINTS	MARKS	CLICKS	DOES
SEE							
HEAR							
TOUCH							
SMELL							
TASTE							
THINK							
FEEL							

Writing Performance Objectives Sheet

Channels Singular Count Unit Conditions Mastery Criteria Date

SeeSays

CALIBRATION IN ORAL READING

*Steps removed
from pure
calibration*

0	SYLLABLE
1	WORD
2	CLAUSE
3	PHRASE
4	SENTENCE
5	PARAGRAPH
6	PAGE
7	CHAPTER
8	BOOK

Writing Performance Objectives Sheet

Channels Singular Count Unit Conditions Mastery Criteria Date

SeeSays

Word

Writing Performance Objectives Sheet

Channels Singular Count Unit Conditions Mastery Criteria Date

SeeSays

Word

from level 5
reader in classroom

Educational Vital Signs

Skill Area	Count Unit	Mastery Criteria *
1. Speaking	ThinkSays word (correct/incorrect)	250-150 words correct/minute (with only random error)
2. Reading • orally	SeeSays word (correct/incorrect)	250-150 words correct/minute (with only random error)
• silently	SeeThinks word (correct/incorrect)	900-350 words correct/minute (with only random error)
4. Math Computation	SeeWrites digit or letter	150-100 digits or letters correct/minute
3. Writing	ThinkWrites word (correct/incorrect)	40-30 words correct/minute (with only random error)
5. Demonstrating Thinking (Comprehension) (Content Knowledge)	ThinkSay word (correct/incorrect) [ThinkSays unit of information (correct/incorrect)]	250-150 words correct/minute (with only random error) (30-20 units of correct information/minute with only random error.)

* These mastery criteria are analogous to the normal limits in medicine.

Writing Performance Objectives Sheet

<u>Channels</u>	<u>Singular Count Unit</u>	<u>Conditions</u>	<u>Mastery Criteria</u>	<u>Date</u>
-----------------	----------------------------	-------------------	-------------------------	-------------

SeeSays

Word

from level 5 @250-150 wpm
reader in class-
room

Writing Performance Objectives Sheet

Channels Singular Count Unit Conditions Mastery Criteria Date

ThinkSays

Writing Performance Objectives Sheet

Channels Singular Count Unit Conditions Mastery Criteria Date

ThinkSays nice comment

Writing Performance Objectives Sheet

Channels Singular Count Unit Conditions Mastery Criteria Date

ThinkSays nice comment in science class

Writing Performance Objectives Sheet

Channels Singular Count Unit Conditions Mastery Criteria Date

ThinkSays nice comment in science class 10/class

Precise Pinpointing Leads to Effective Assessment & Instruction (Alignment)



Precise Pinpointing Leads to Effective Assessment & Instruction (A Geography Example)

Oregon 5th Grade Geography Standard:

- Students will locate and identify on maps the continents of the world, the 50 states of the United States, and the physical features of Oregon.

Precise Pinpointing Leads to Effective Assessment & Instruction (A Geography Example)

Oregon 5th Grade Geography Standard:

Students will locate and identify on maps the continents of the world, the 50 states of the United States, and the physical features of Oregon.

Geography Example

Oregon 5th Grade Geography Standard:

Students will locate and identify on maps the continents of the world, the 50 states of the United States, and the physical features of Oregon.

Revised standard:

1. SeeWrites word (seven continents), given a map of the world, @100-120 correct letters/min. by 15 May 2009.

Geography Example

Oregon 5th Grade Geography Standard:

Students will locate and identify on maps the continents of the world, the 50 states of the United States, and the physical features of Oregon.

Revised standard:

1. SeeWrites word (seven continents), given a Euro-centric mercator map of the world, @100-120 correct letters/min. by 15 May. 2009.
2. SeeWrites letter (postal abbreviations), given a U.S. map, @100-120 correct letters/min. by 15 May 2009.

THINK - WRITE STATE ABBREVIATIONS

Name: _____

Date: _____

Directions: Write 2 letter abbreviations for 50 state.

Fluency Standard: 100 letters/minute

Correction: Count letters correct and incorrect.

Correct: _____ Incorrect: _____



Geography Example

Oregon 5th Grade Geography Standard:

Students will locate and identify on maps the continents of the world, the 50 states of the United States, and the physical features of Oregon.

Revised standard:

1. See/Writes word (seven continents), given a Euro-centric mercator map of the world, @100-120 correct letters/min. by 15 May 2009.
2. See/Writes letter (postal abbreviations), given a U.S. map, @100-120 correct letters/min. by 15 May 2009.
3. See/Writes word (the Cascade Mts., the Blue Mts., the Coast Range, Crater Lake, Mt. Hood, Klamath Lake, the Willamette River, the Snake River, the Columbia River, and the cities of Bend, Eugene, Portland, Medford, and Salem), given a physical map of Oregon, @100-120 correct letters/min. by 15 May 2009.

Components of Precise Measurement

1. Precisely Describe Phenomenon (e.g behavior) to Observe - Pinpoint
2. Accurately Record Significant Dimensions of Phenomenon - Record
 - Frequency
 - Change
 - Bounce

FREQUENCY

1. Frequency (count/time) is the most sensitive measure of human performance.



FREQUENCY

1. Frequency (count/time) is the most sensitive measure of human performance.
2. Non-frequency measures are less precise because they are Incomplete and/or Indirect

A Variety of Common But Less Precise Educational Performance Measures

Incomplete Measures: (measures that are missing critical information)	
Count Only Measures	Time Only Measures
• # done (e.g. 25 problems completed)	• elapsed time
• # correct/incorrect (e.g. 20 correct/3 errors)	• time on task
• % correct/incorrect (e.g. 90% correct)	• engaged time
• # attempted (e.g. 20 problems attempted)	• length of time of response (duration)
• # done per opportunity (e.g. 7 of 10 steps)	• time to make first response (latency)
Indirect Measures: (compilations of count and/or time information but the direct performance information has been lost.)	
• Stanine score	
• Grade equivalents	
• Score on paper & pencil tests (e.g. criterion, achievement, diagnostic)	
• Letter grades (A,B,C,D,F)	
• Letter descriptions (S - Satisfactory; US - Unsatisfactory)	
• Anecdotal records	
• Rankings	
• Rubric/rating scale (Likert scale)	

Note: All measures listed are derived from count and/or time information.

From: Starlin, C. (2002) *A science of education*. Unpublished manuscript. University of Oregon, Eugene, OR.

Examples of Incomplete Measure

	Thursday	Friday
Count	150 words correct	75 words correct
Time	3 minutes	1 minute
Count/Time	50 words/minute	75 words/minute

	Monday	Tuesday
Count	10 refusals	8 refusals
Time	100 minutes	20 minutes
Count/Time	.1 refusals/minute	.4 refusals/minute

Example of Indirect Measure

Frequency vs. % Correct

Oral Reading Scores of 131, 7th Grade Students

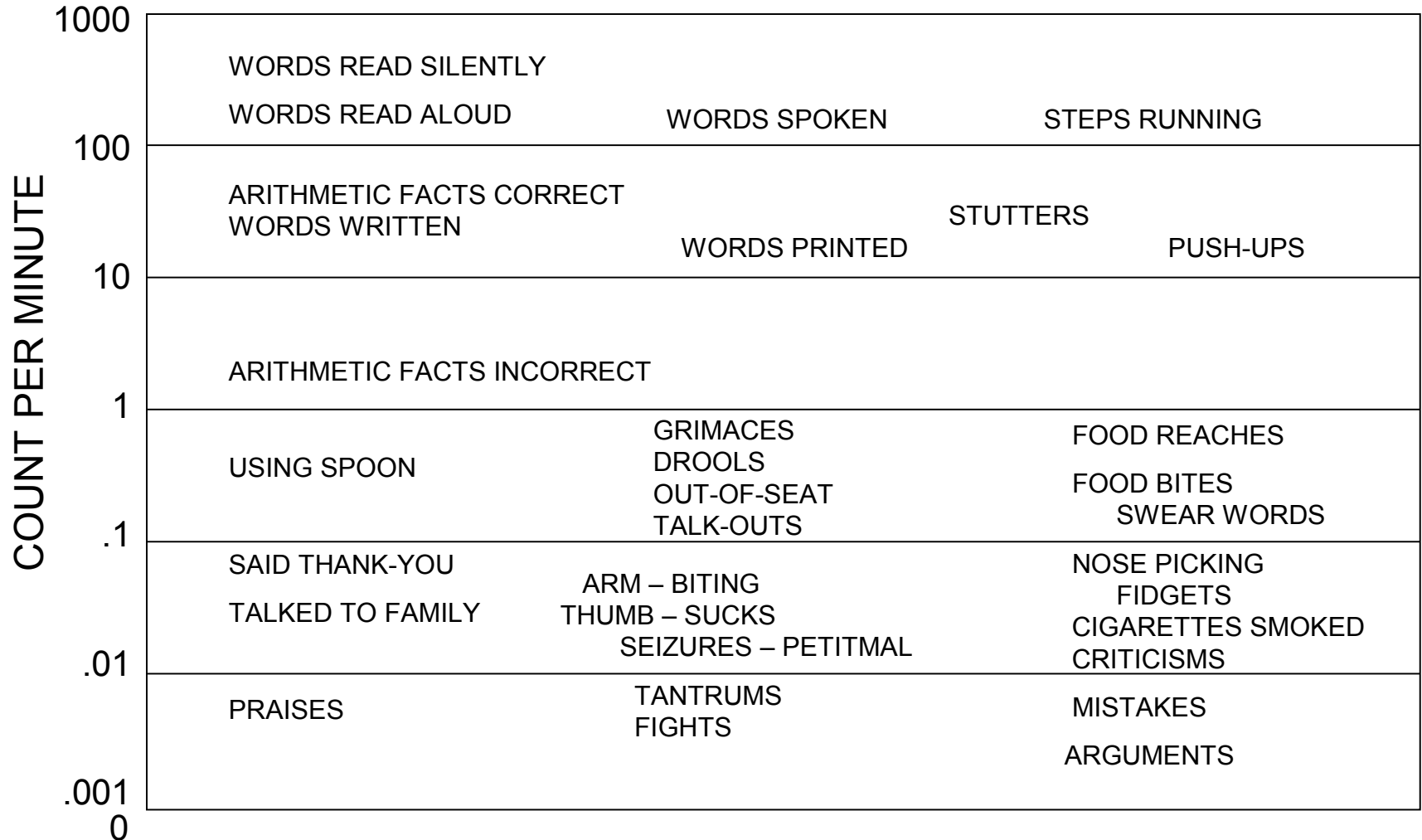
Median: Correct-100; Error-4)

TOP 10 STUDENTS	CORRECT/MIN	INCORRECT/MIN	% CORRECT
131. Doug	223	0	100%
130. Brian	187	1	99%
129. Pat	177	1	99%
128. Julie	176	0	100%
127. Lauri	175	1	99%
126. Gwen	173	2	99%
125. Linda	171	0	100%
124. Pam	170	0	100%
123. Jeff	166	0	100%
122. Mark	160	4	98%

FREQUENCY

1. Frequency (count/time) is the most sensitive measure of human performance.
2. Non-frequency measures are less precise because they are Incomplete and/or Indirect
3. Frequency is a dimension of behavior (i.e. Different behaviors have different frequencies.)


GRAPHIC REPRESENTATION OF DIFFERENT FREQUENCIES



A FREQUENCY RANGE OF HUMAN BEHAVIORS

FREQUENCY IS PIVOTAL

The decision regarding our base performance measure determines how we measure the change (celeration) and bounce of performance.



Components of Precise Measurement

1. Precisely Describe Phenomenon (e.g behavior) to Observe - Pinpoint
2. Accurately Record Significant Dimensions of Phenomenon - Record
 - Frequency
 - Change
 - Bounce
3. Display Data to Facilitate Accurate Conclusions/Decisions - Chart

Decisions Regarding Displaying Data

Narrative



NARRATIVE RECORD

Final Report: Spring Term

June, 1970

Name: Mike

Birthdate: July 5, 1959

C.A.: 10-11

Parents: Mr. And Mrs.

Address:

Phone:

Clinician: Sally

Previous Clinical Attendance: Summer, 1969, Winter, 1970

Present Grade Placement: 4th grade

School: Meadow Lark, Eugene

This report covers remedial aid given at DeBusk Clinic at the University of Oregon during the period April 20, 1970 through June 5, 1970. Included in this report are a summary of past and present achievement status, changes noted during the clinic period, basic instructional method and materials used, and recommendations for further instruction.

Past and Present Achievement Status:

Mike has studied and learned a total of 112 Dolch sight words. These are listed on a separate paper. He read the entire Merrill Linguistic Reader #3 and eight booklets of the Miami Linguistic Reader series during this clinic period.

Mike is saying the Dolch sight words well above his first time sample in the "Before" period. He said 35 words per minute the first day he was timed. He has moved consistently from 35 to 116 words per minute which was recorded on his last day in the clinic. During this rise in number of words correct, his error rate was dropping. The first day he had five errors. But out of the last 11 days, he made only two errors.

Decisions Regarding Displaying Data

Narrative, Table



TABULAR RECORD
(ISOLATED WORDS READ)

-	Correct	Incorrect	Day	Correct	Incorrect
M	34	5	M		
Tu			Tu		
W	48	5	W	80	1
Th			Th	88	0
F	53	7	F	90	0
M	59	4	M	93	0
Tu			Tu		
W	65	4	W		
Th	70	3	Th	95	0
F	77	0	F		
M	75	2	M	105	1
Tu			Tu		
W	80	1	W	119	1
Th	91	0	Th		
F			F		
M	89	1	M		
Tu			Tu		
W			W		
Th			Th		
F	90	1	F		

Decisions Regarding Displaying Data

Narrative, Table, Graph



Dpmin-11EC

DAILY per minute CHART™

© 1984 AMERICAN PROGRESS 410 - 44 CROFTS 212 352 007
119 BAY STAINBURY RD. WILMINGTON, DE 19812

10 Jan 70
DIV 40 20 10

15 Feb 70
DIV 40 20 10

15 Mar 70
DIV 40 20 10

12 Apr 70
DIV 40 20 10

10 May 70
DIV 40 20 10

7 Jun 70
DIV 40 20 10

4 SUCCESSIVE 8 CALENDAR 12 WEEKS 16 20

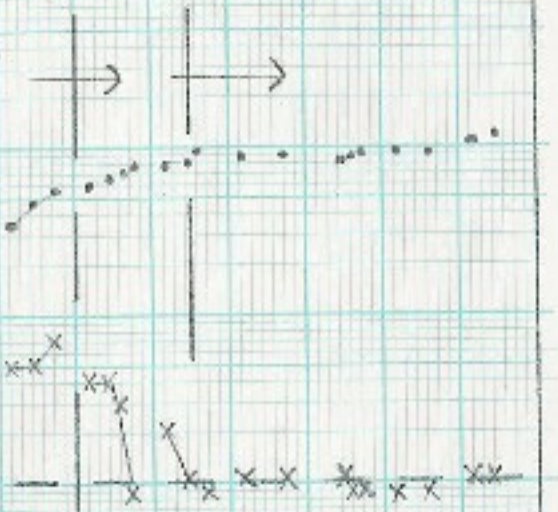
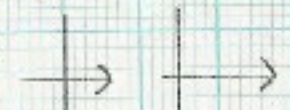
COUNT PER MINUTE

1000
500
100
50
10
5
1
.5
.1
.05
.01
.005
.001

COUNTING TIMES
16"
15"
20"
30"
1"
2"
5"
10"
20"
50"
100"



DOLCH LIST



Show Chart Pte. for library

Do own charting

CLAY STARLIN
SUPERVISOR

ADVISER

SALLY HANSON
MANAGER

SUCCESSIVE CALENDAR DAYS

Mike
PERFORMER

4th isolated
COUNTED read

Decisions Regarding Displaying Data

Narrative, Table, Graph

(a picture is worth 10,000 words)

Smith, et al (2002) Constructing Knowledge. *American Psychologist*, 57(10), 749-761.



Decisions Regarding Displaying Data

Narrative, Table, Graph

(a picture is worth 10,000 words)

Smith, et al (2002) Constructing Knowledge. *American Psychologist*, 57(10), 749-761.



Time Series Graph

(75 % of graphs are time series, Tufte, E. (2004). *The visual display of quantitative information*)

Decisions Regarding Displaying Data

Narrative, Table, Graph

(a picture is worth 10,000 words)



Time Series Graph

(75 % of graphs are time series)



Rate-of-Change Graph (logarithmic scale)

Decisions Regarding Displaying Data

Narrative, Table, Graph

(a picture is worth 10,000 words)



Time Series Graph

(75 % of graphs are time series)

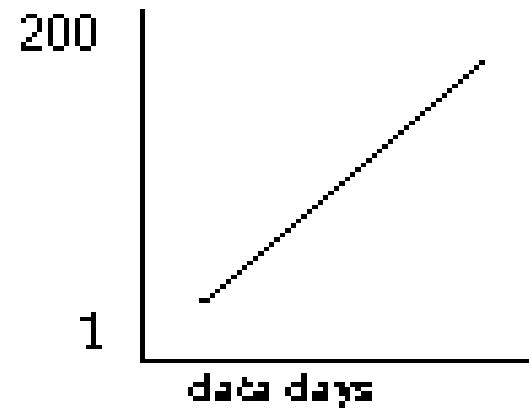
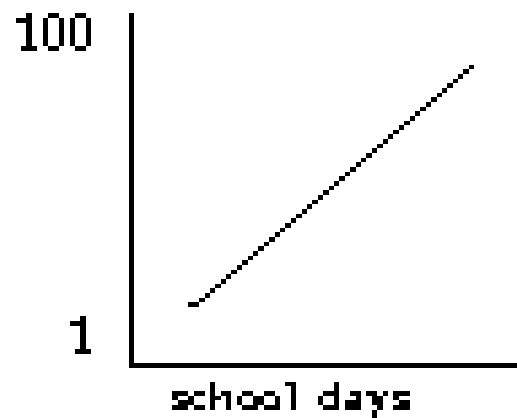
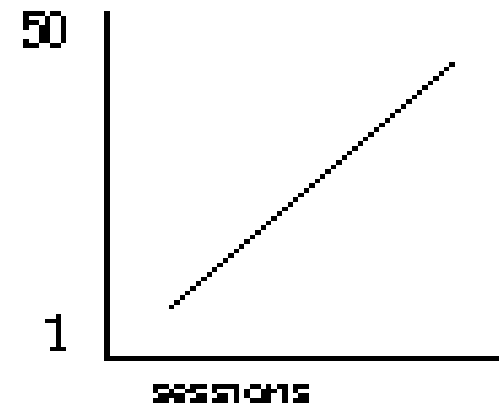
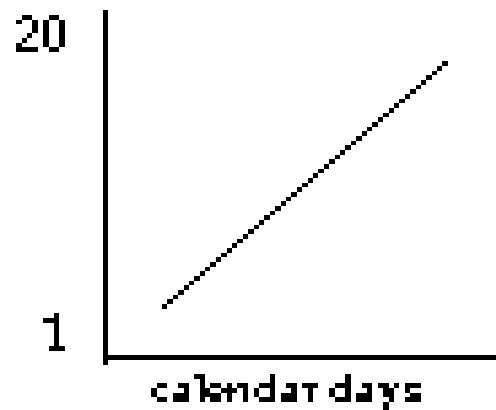


Rate-of-Change Graph (logarithmic scale)



Standard

Changing Dimensions of Graphic Axes to Create Trend Lines Which Diagonally Bisect a Chart (Fill the Frame Charts)



Decisions Regarding Displaying Data

The SCC is a standard time series, rate of change chart.

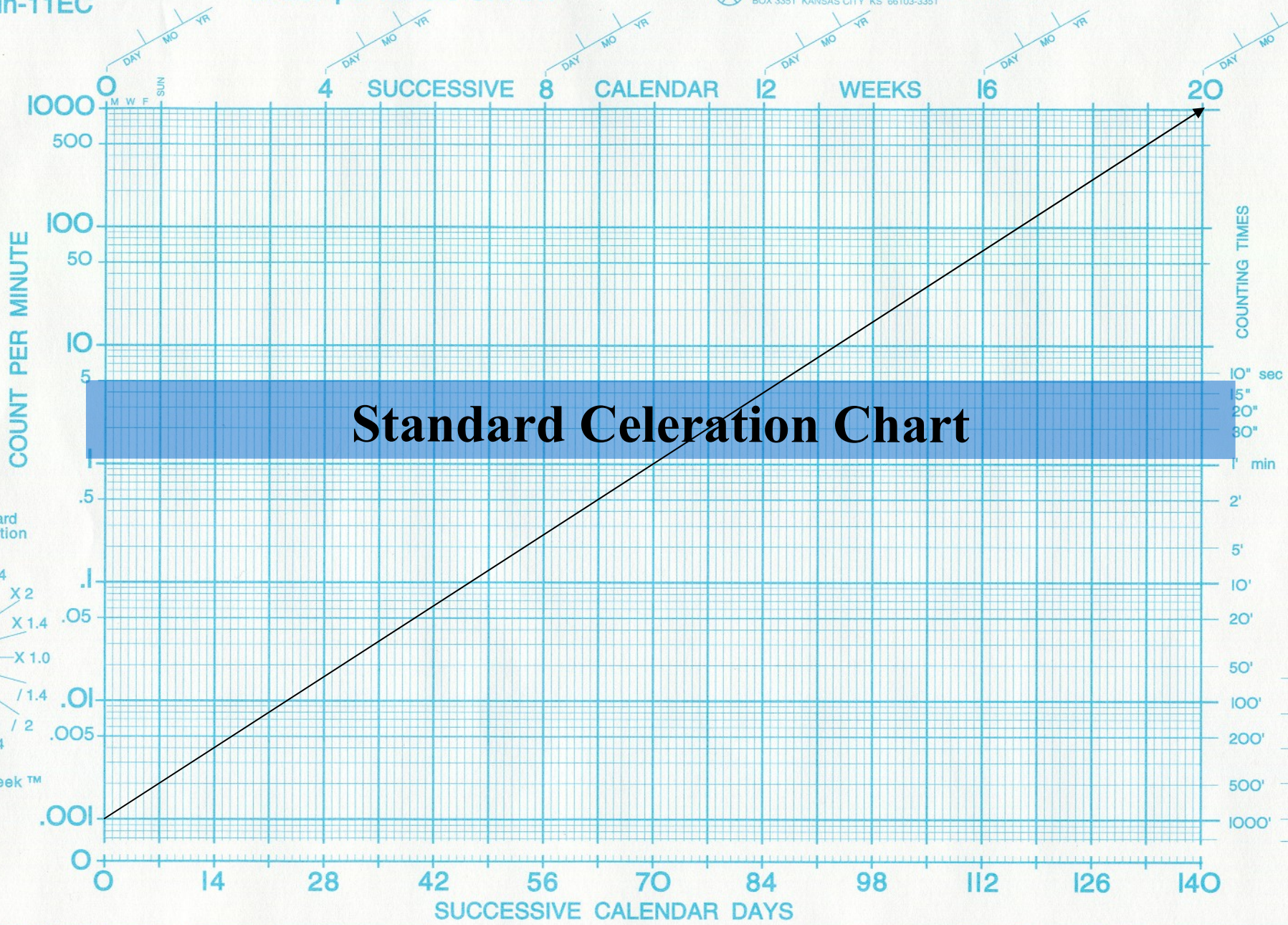


Decisions Regarding Displaying Data

The SCC is a time series, rate of change chart.

The SCC's most powerful feature:

Visually significant change =
Educational impressive change



- Standard celeration
- X16
 - X 4
 - X 2
 - X 1.4
 - X 1.0
 - / 1.4
 - / 2
 - / 4
 - / 16
- per week™

SUPERVISOR ADVISER MANAGER PERFORMER COUNTED

ORGANIZATION DIVISION ROOM TIMER COUNTER CHARTER

Sensing Differences

- See
 - Hear
 - Taste
 - Smell
 - Touch
- 

Seeing Differences

- The Greek astronomer Hipparchus (fl 146-127 BC) cataloged the brightness of 850 stars and classified them from “stars of the first magnitude” (brightest) to “stars of the sixth magnitude” (dimmiest). These orders of magnitude were based on Hipparchus’ ability to perceive differences in brightness. Subsequent calculations showed these magnitudes to vary by multiples (logarithmically).

From: Lord, J. (1994). *Sizes: The illustrated encyclopedia*. New York: Harper Perennial. Lord, 1994)

Hearing Differences

The pure tone audiometer generates tone frequencies of:

125 cycles/second (hertz) → 250 Hz → 500 Hz →
1000 Hz → 2000 Hz → 4000 Hz → 8000 Hz

Tasting Differences



Smelling Differences



Touching Differences



EDUCATIONAL IMPRESSIVE CHANGE (EIC)

1. Change in performance is always detected through sensory observations. [e.g. $\times 2$ (doubling) is always detected]



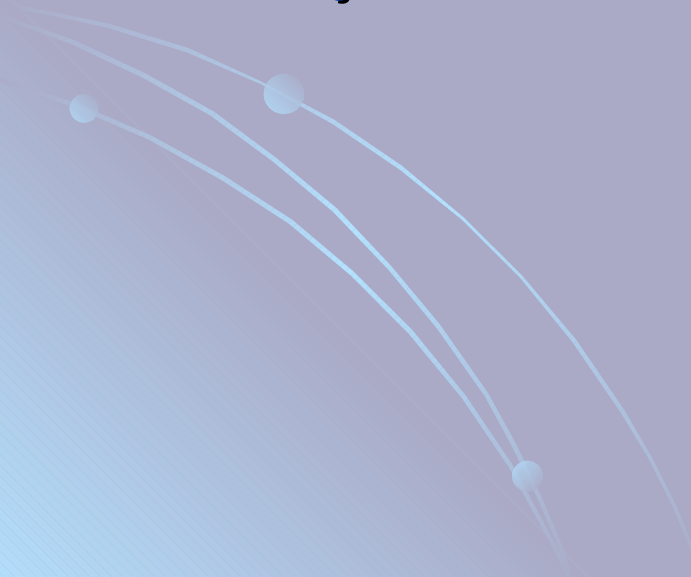
EDUCATIONAL IMPRESSIVE CHANGE (EIC)

1. Change in performance is always detected through sensory observations. [e.g. x2 (doubling) is always detected]
2. The change is rated as impressive.



EDUCATIONAL IMPRESSIVE CHANGE (EIC)

1. Change in performance is always detected through sensory observations. [e.g. x2 (doubling) is always detected]
2. The change is rated as impressive.
3. Complimentary relationship between objective and subjective information
(i.e. impressive graphic change is verified by subjective sensory reactions, #2 above)



EDUCATIONAL IMPRESSIVE CHANGE (EIC)

1. Change in performance is always detected through sensory observations. [e.g. x2 (doubling) is always detected]
2. The change is rated as impressive.
3. Complimentary relationship between objective and subjective information
(i.e. impressive graphic change is verified by subjective sensory reactions, #2 above)
4. Change data involves direct observation based learning records and summaries vs. psychometric measures (e.g. achievement, IQ tests) and group statistics.

EDUCATIONAL IMPRESSIVE CHANGE (EIC)

1. Change in performance is always detected through sensory observations. [e.g. x2 (doubling) is always detected]
2. The change is rated as impressive.
3. Complimentary relationship between objective and subjective information
(i.e. impressive graphic change is verified by subjective sensory reactions, #2 above)
4. Change data involves direct observation based learning records and summaries vs. psychometric measures (e.g. achievement, IQ tests) and group statistics.
5. EIC scales the power of the intervention effects in terms of the rapidity of the effects (e.g. x1.1 per week = 7 weeks to double performance; x2 per week = one week to double performance. - after Berquam, 1980)

Educationally Impressive Change (EIC)

Discovery of EIC Leads to Best Practice Interventions
(analogy to medicine – penicillin, polio vaccine)

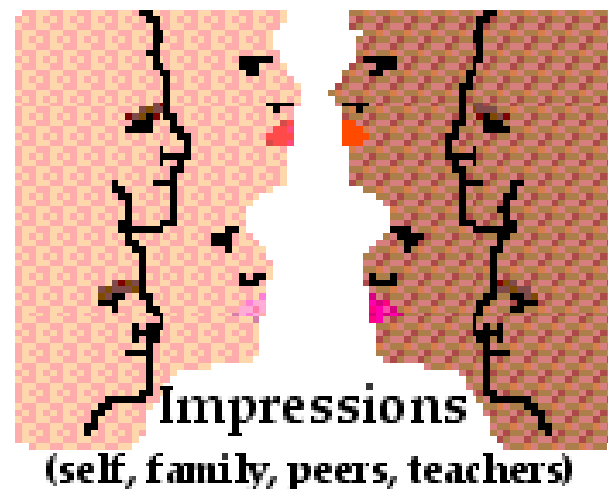
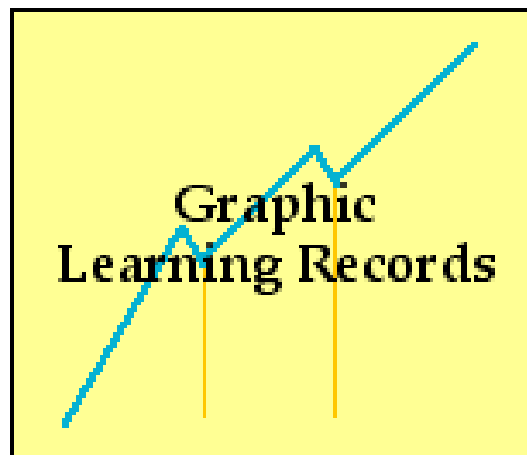
i.e. The knowledge that is the purpose of science.



“It would be possible to describe everything scientifically but it would make no sense, it would be without meaning, as if you described a Beethoven symphony as a variation of wave pressure.”

Albert Einstein

The Three Complementary Dimensions of Student Learning Evaluation



Big Idea Summary

- **Accountability = Scientific Practice**




Big Idea Summary

- **Accountability = Scientific Practice**
- **Effective Science Involves Two Major Components**
 - Precise Measurement
 - Robust Research Methodology



Big Idea Summary

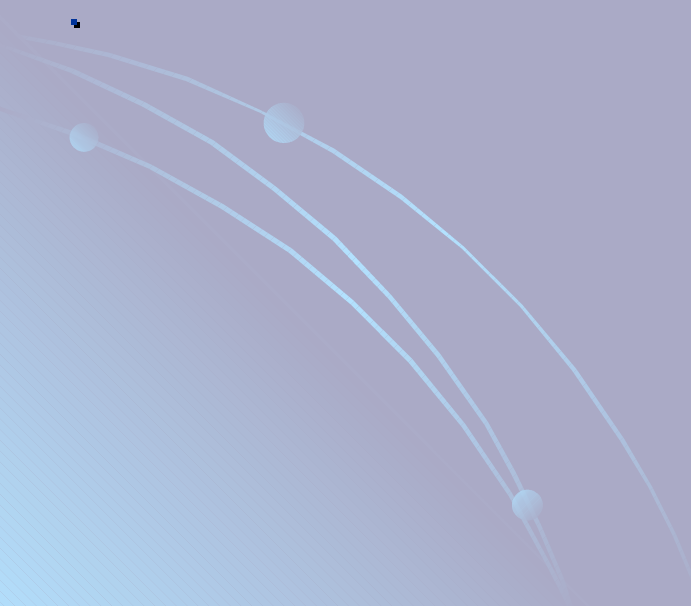
- **Accountability = Scientific Practice**
 - **Effective Science Involves Two Major Components**
 - Precise Measurement
 - Robust Research Methodology
 - **Standardization of a Precise Measurement System is Foundational to the Establishment of a Scientific Discipline**
- 

Big Idea Summary

- **Accountability = Scientific Practice**
- **Effective Science Involves Two Major Components**
 - Precise Measurement
 - Robust Research Methodology
- **Standardization of a Precise Measurement System is Foundational to the Establishment of a Scientific Discipline**
- **Components of Precise Measurement**
 1. Precisely Describe Phenomenon (e.g behavior) to Observe - Pinpoint
 2. Accurately Record Significant Dimensions of Phenomenon - Record
 3. Display Data to Facilitate Accurate Conclusions/Decisions - Chart

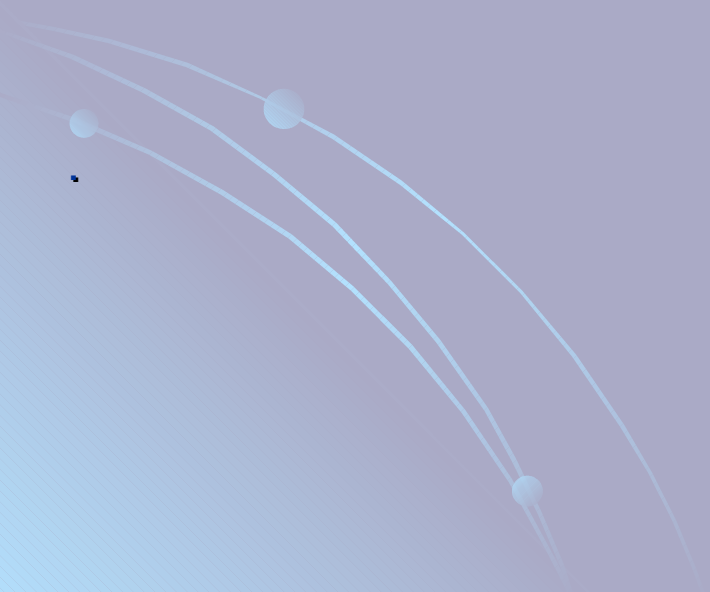
Big Idea Summary

- **Precise Pinpointing**
 - Critical to Accurate & Reliable Data
 - Leads to Effective Assessment & Instruction



Big Idea Summary

- **Precise Pinpointing**
 - Critical to Accurate & Reliable Data
 - Leads to Effective Assessment & Instruction
- **Frequency (count/time) Most Sensitive Performance Measure**
 - Learning = change in frequency
 - Bounce = variability in frequency



Big Idea Summary

- **Precise Pinpointing**
 - Critical to Accurate & Reliable Data
 - Leads to Effective Assessment & Instruction
- **Frequency (count/time) Most Sensitive Performance Measure**
 - Learning = change in frequency
 - Bounce = variability in frequency
- **A Display System Where:**
 - *“visually significant change = educationally impressive change”*
 - **Facilitates Discovery of Powerful Interventions**

Big Idea Summary

- **Precise Pinpointing**
 - Critical to Accurate & Reliable Data
 - Leads to Effective Assessment & Instruction
- **Frequency (count/time) Most Sensitive Performance Measure**
 - Learning = change in frequency
 - Bounce = variability in frequency
- **A Display System Where:**
 - *“visually significant change = educationally impressive change”*
 - **Facilitates Discovery of Powerful Interventions**
- **Educationally Impressive Change (EIC) = Our Senses Always Detect the Change**
[e.g. x2 (doubling) or greater change in performance]