Measuring Outcomes Quality Assurance

2009

Revised with permission from ECO

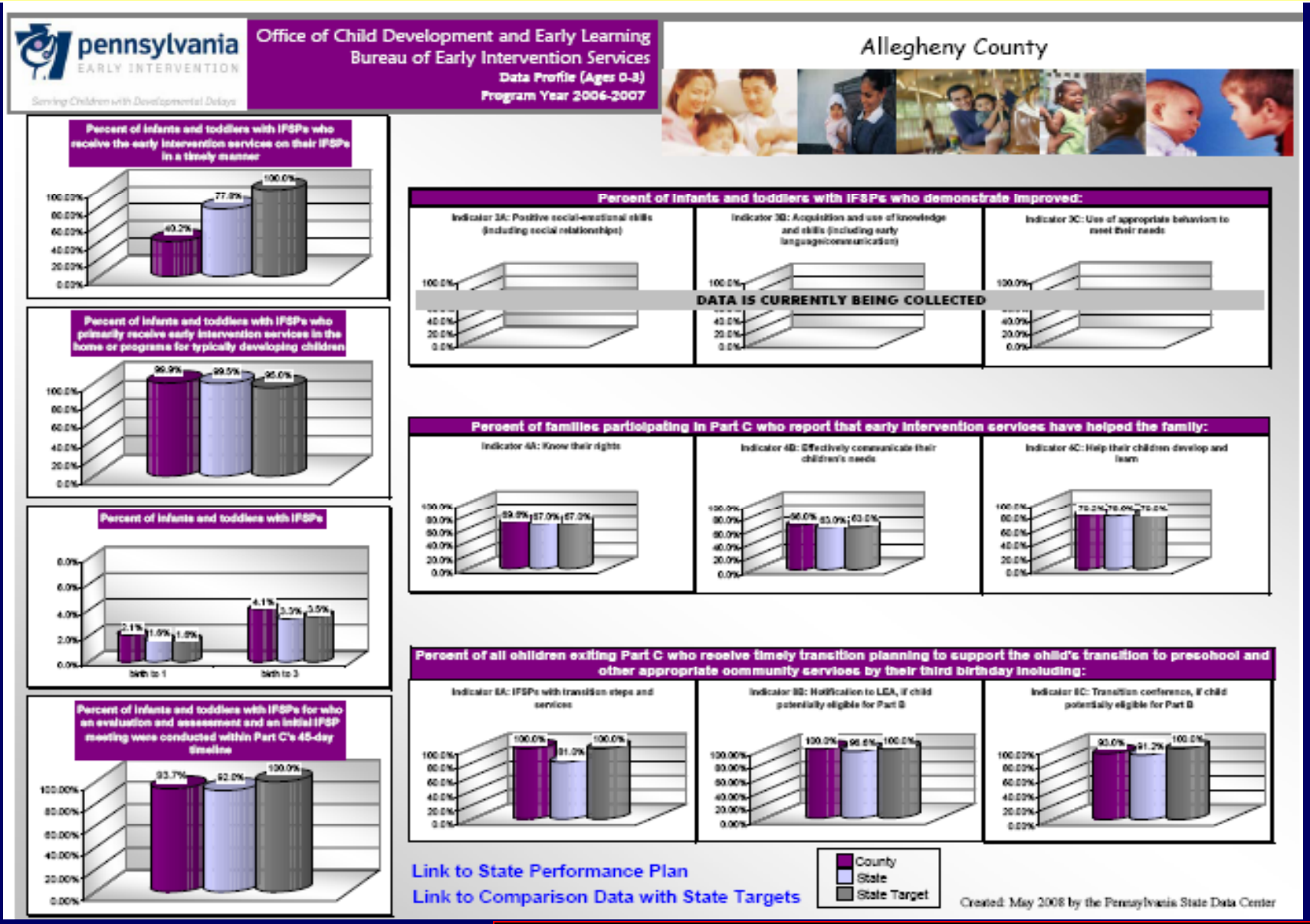
Objectives

- Understand how to evaluate completed child outcomes summary forms for accuracy.
- Understand ways to review entry outcome data.
- Understand what's needed to promote quality child outcome ratings.
- Understand the vital importance of quality local level outcome ratings and data entry.
- Understand some ways outcome data can be examined.

Purposes of Outcome Data Collection

- Demonstrate effectiveness of program for policy makers & stakeholders.
- Identify local programs in need of improvement.
- Determine whether services are equally effective for different sub-groups.
  - Race, disability groups, SES…
- Determine whether services are equally effective across different outcome areas.
- The real prize…High quality services for children and families that will lead to good outcomes.

Stakeholder Sharing



http://penndata.hbg.psu.edu/MAWA/index.aspx

National Picture

State Approaches to Measuring Child Outcomes

Approach	Part C 56 States
COSF 7-point scale	40/58 **(71%)**
One tool statewide	8/56 (14%)
Other	6/56 (11%)
Multiple Publishers' online tools	2/56 (4%)

Reporting Categories

- % of children who demonstrated improved: [OSEP]
  - Positive social-emotional skills (social relationships).
  - Acquisition and use of knowledge and skills.
  - Use of appropriate behaviors to meet their needs.

- Reporting Categories

  A. % who did not improve functioning.

  B. % who improved functioning but not sufficient to move nearer to functioning comparable to same-aged peers.

  C. % who improved functioning to a level nearer to same-aged peers but did not reach it.

  D. % who improved functioning to reach a level comparable to same-aged peers.

  E. % who maintained functioning at a level comparable to same-aged peers.

Data Reporting

- Initial progress (entry-exit) data will only represent children who have entered and exited since outcome system was put in place
  - Typically, initial data may represent children who participated in the program 6 to 12 months.
- DOD will look at entry data this year and add progress data thereafter.
- States
  - Feb 2007 – states reported entry data
  - Feb 2008 – states reported very early progress data
  - Focus for next few years will be on the quality of the data

Steps to Ensuring Quality Data

Before	**Good data collection/Training** **Good data system and data entry procedures**
During	**Ongoing supervision of implementation** **Feedback to implementers** **Refresher training**
After	**Review of COSF records** **Data analyses for validity checks**

Reviewing COSF Quality

Review Terry's COSF

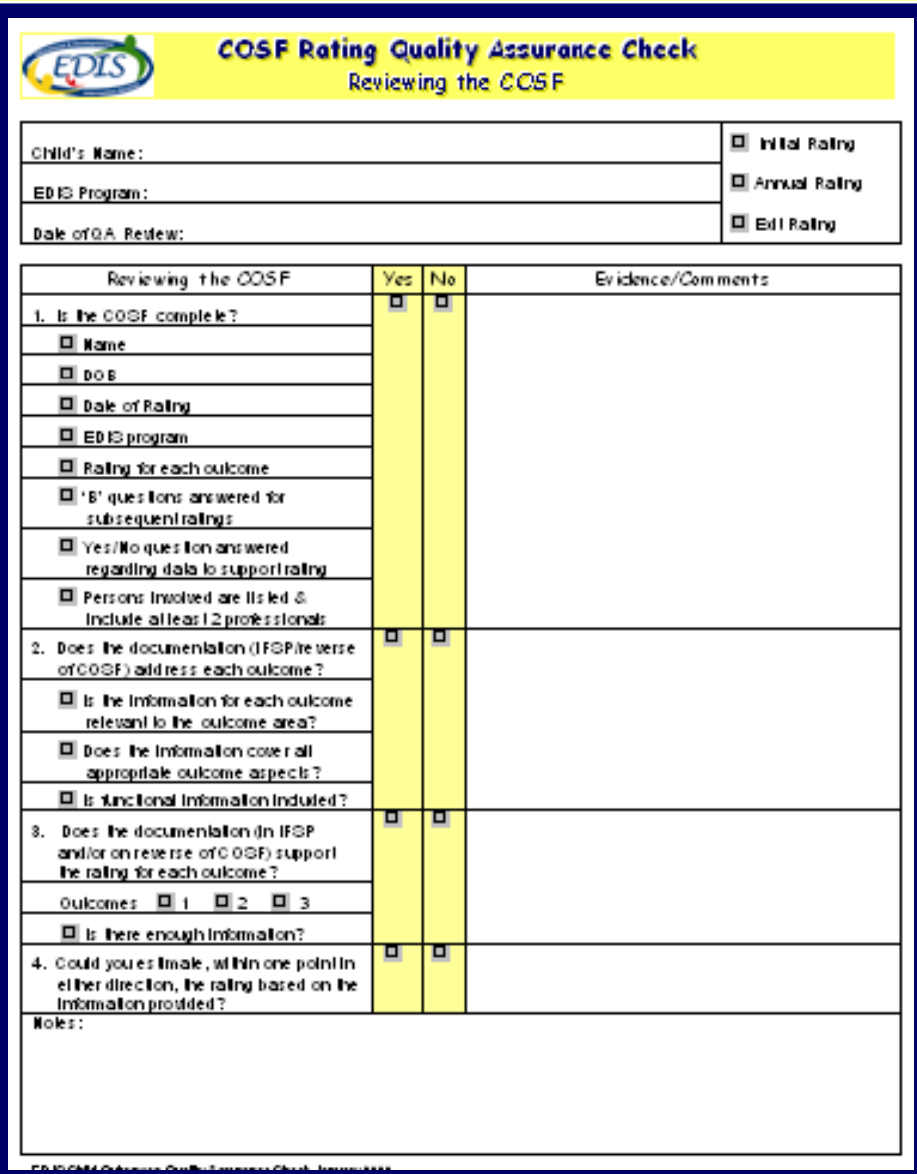
1. Is the COSF complete?

2. Does the evidence correspond with the appropriate outcome area?
   - If not what are some examples where the evidence is listed with the incorrect outcome area?

3. Does the evidence provided support the rating?
   - Where is more evidence needed?
   - Where does the evidence clearly suggest the rating is incorrect?

Review EDIS COSF

✓ Evan



COSF Rating Quality Assurance Check
Reviewing the COSF

Child's Name:	☐ Initial Rating
EDIS Program:	☐ Annual Rating
Date of QA Review:	☐ Exit Rating

Reviewing the COSF	Yes	No	Evidence/Comments
1. Is the COSF complete?	☐	☐	
☐ Name			
☐ DOB			
☐ Date of Rating			
☐ EDIS program			
☐ Rating for each outcome			
☐ 'B' questions answered for subsequent ratings			
☐ Yes/No question answered regarding data to support rating			
☐ Persons involved are listed & include at least 2 professionals			
2. Does the documentation (IFSP/reverse of COSF) address each outcome?	☐	☐	
☐ Is the information for each outcome relevant to the outcome area?			
☐ Does the information cover all appropriate outcome aspects?			
☐ Is functional information included?			
3. Does the documentation (in IFSP and/or on reverse of COSF) support the rating for each outcome?	☐	☐	
Outcomes ☐ 1 ☐ 2 ☐ 3			
☐ Is there enough information?			
4. Could you estimate, within one point in either direction, the rating based on the information provided?	☐	☐	

Notes:

Observing the COSF Rating Process

- Do all team members participate in the discussion?
- Is parent input considered?
- Does the team discuss multiple assessment sources?
- Does the team describe the child's functioning beyond test scores or isolated test items?
- Does the discussion include the child's full range of functioning (A-E; IF; F)?
- Does the team sufficiently discuss the child's functioning before making a rating?

QA – Observing the COSF Rating

COSF Rating Quality Assurance Check
Observing the COSF Rating Process

Child's Name:	☐ Initial Rating
EDIS Program:	☐ Annual Rating
Date of QA Review:	☐ Exit Rating

Observing the COSF Rating Process	Yes	No	Evidence/Comments
1. Do all team members participate in the discussion?			
2. Is parent input considered in the rating?			
3. Does the team discuss multiple assessment sources?			
4. Does the team describe the child's functioning, beyond just test scores or isolated test items?			
5. Does the discussion include the child's full range of functioning? (i.e., skills & behaviors that are age appropriate, immediate foundational, & leading to immediate foundational?			
6. Does the team sufficiently discuss the child's functioning before making a rating?			

Notes:

Does documentation support the rating?

Does documentation include

7. age-appropriate skills & behaviors?
6. skills & behaviors that are age appropriate with an identified area of concern/question?
5. a mix of skills & behaviors that are age appropriate and not?
4. a few examples of skills & behaviors that are age appropriate, but mostly not?
3. immediate foundational skills, and none that are age appropriate?
2. a few examples of immediate foundational skills, but mostly skills & behaviors that are much lower than age expected?
1. skills & behaviors that are much lower than age expectations, with none that are immediate foundational?

Looking at Data

Evidence

Inference

Action

Evidence

- Evidence refers to the numbers
  - 16% had an entry COSF rating of 7 for Outcome 2
  - Program A had 25 overdue entry Outcome Ratings

- The numbers are not debatable.

Inference – the rest of the story

- How do you interpret the #s?
- What can you conclude from the #s?
- Does evidence mean good news? Bad news? News we can't interpret?
- To reach an inference, sometimes we analyze data in other ways (ask for more evidence).
- Inference is debatable -- even reasonable people can reach different conclusions from the same set of numbers.
- Stakeholder involvement can be helpful in making sense of the evidence.

Action

- Given the inference from the numbers, what should be done?
- Recommendations or action steps.
- Action can be debatable – and often is.

Key Points

- Evidence refers to the numbers and the numbers by themselves are meaningless
- Inference is attached by those who read (interpret) the numbers
- You know your local data better than anyone
- You have the opportunity and obligation to attach meaning to it – tell the rest of the story

E, F, or A?

COSF users unaware of the need to answer the yes/no progress question	90% of exit COSFs in Program B missing a response to the yes/no progress question	Revise COSF procedures to emphasize completion of yes/no progress question
Conduct staff development on using the 7-point rating scale	75% of children in Program A received entry ratings of 2	COSF users misunderstand the definition of points on the 7-point scale
Currently used tools are not accurately assessing children's social emotional skills	Invest resources in materials for assessing social-emotional skills	45% of children reported in category 'e' for statewide progress data, Outcome 1

Data Analysis

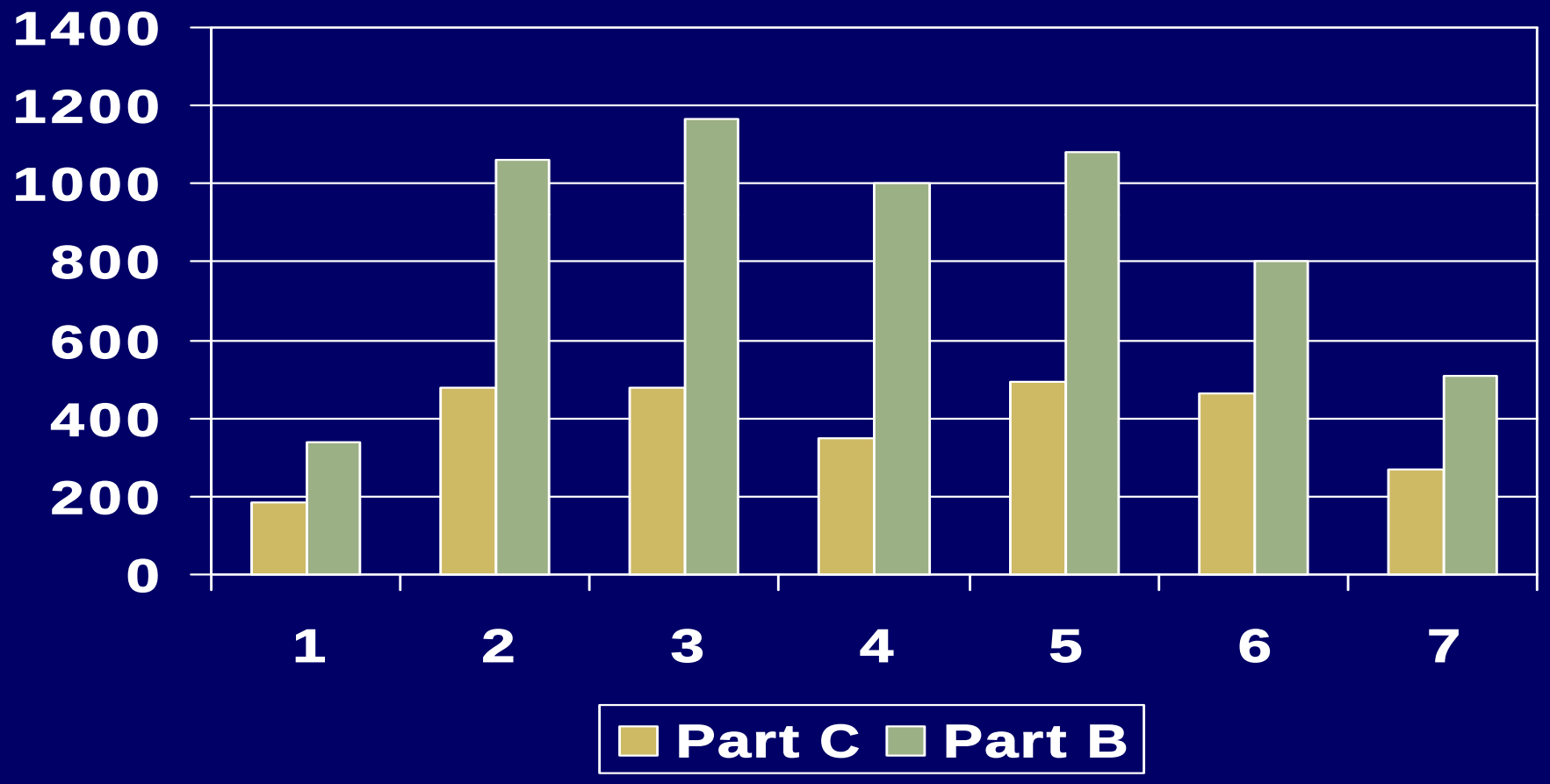
- Examine the data for inconsistencies.
- If/when you find something strange, look for other data that might help explain it.
- Is the variation caused by something other than bad data?

Predictable Patterns – Entry Data

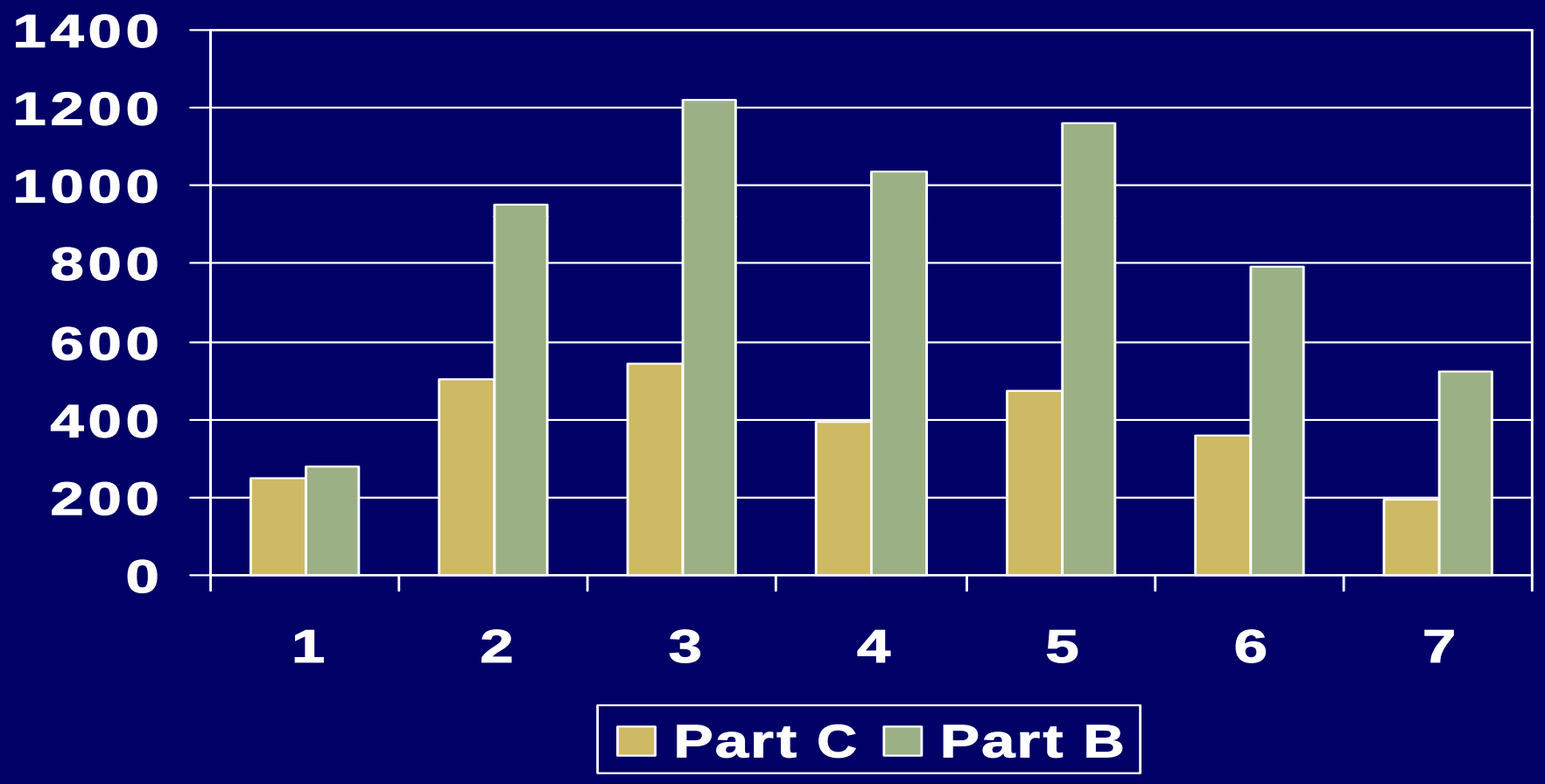
1. Children will differ from one another in their entry scores in reasonable ways (e.g., fewer scores at the high and low ends of the distribution, more scores in the middle).
   - Yet, if sensitive to differences in child functioning, should have children in every category.
   - Look at entry level distributions.

- Rationale:
  - Evidence suggests that EI serves more mildly than severely impaired children (few ratings at the lowest end – 1).
  - Few children receiving services would be expected to be considered as functioning typically (few ratings in the typical range – 6/7).

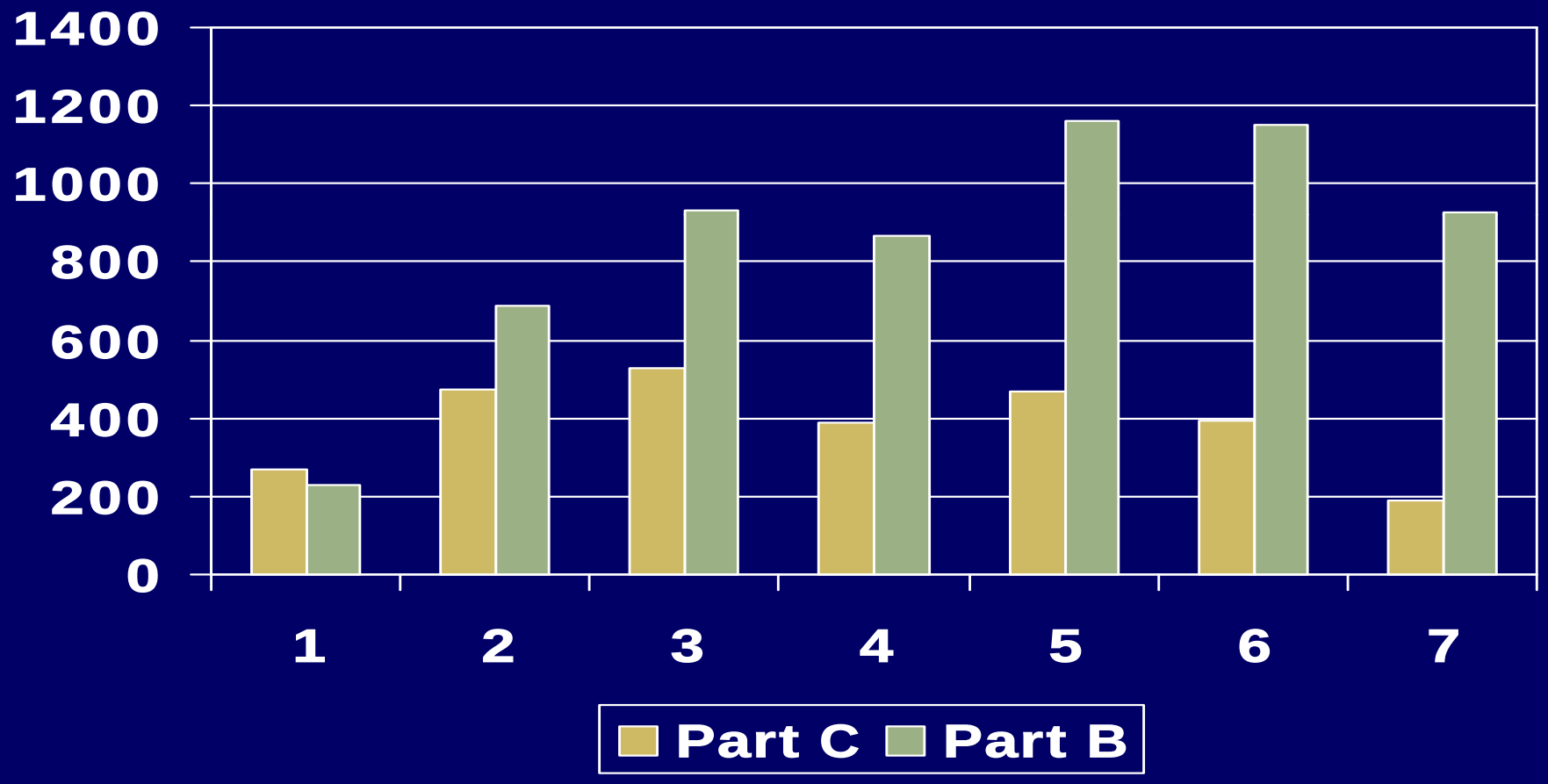
Outcome 1 Entrance: 06-07



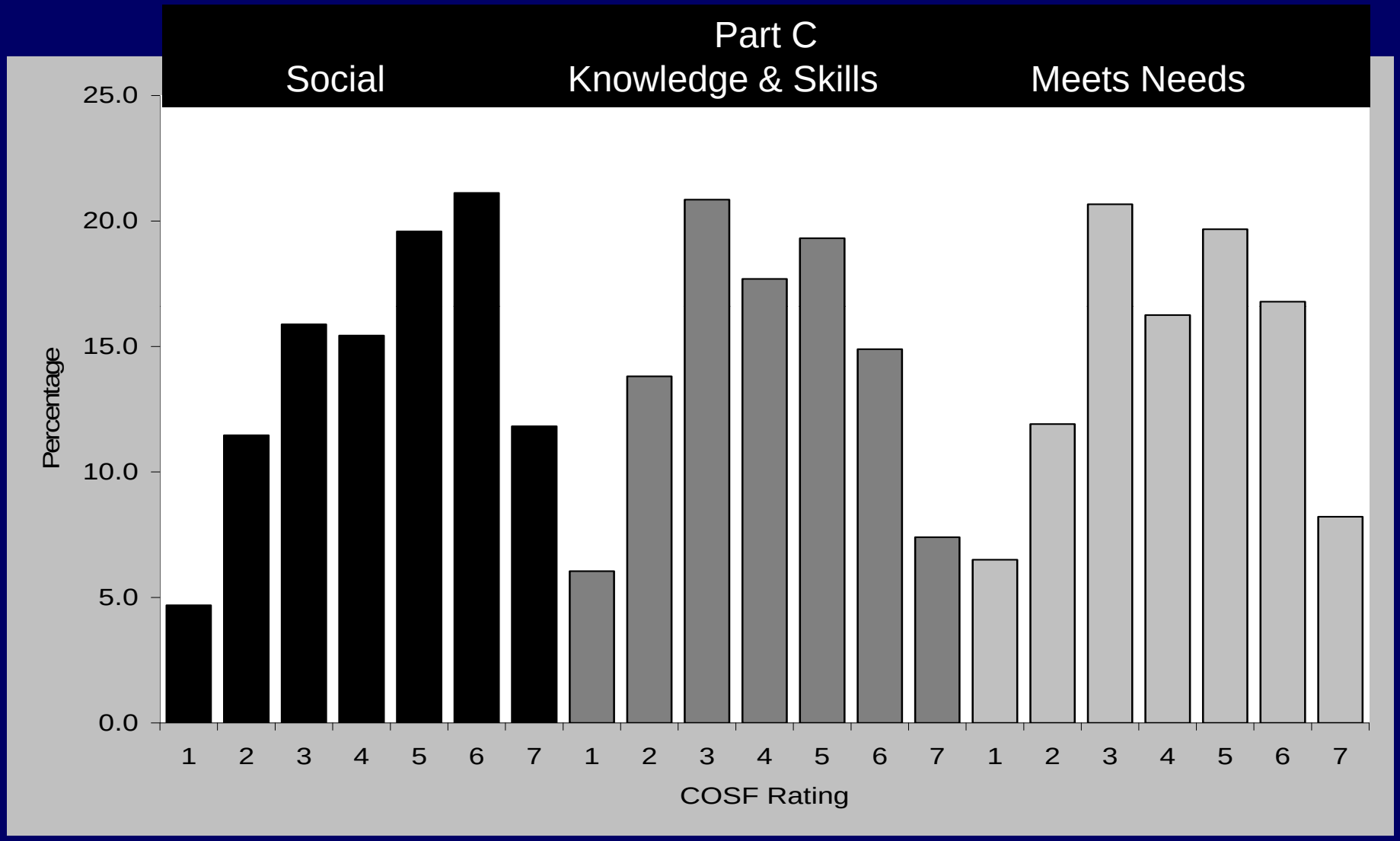
Outcome 2 Entrance: 06-07



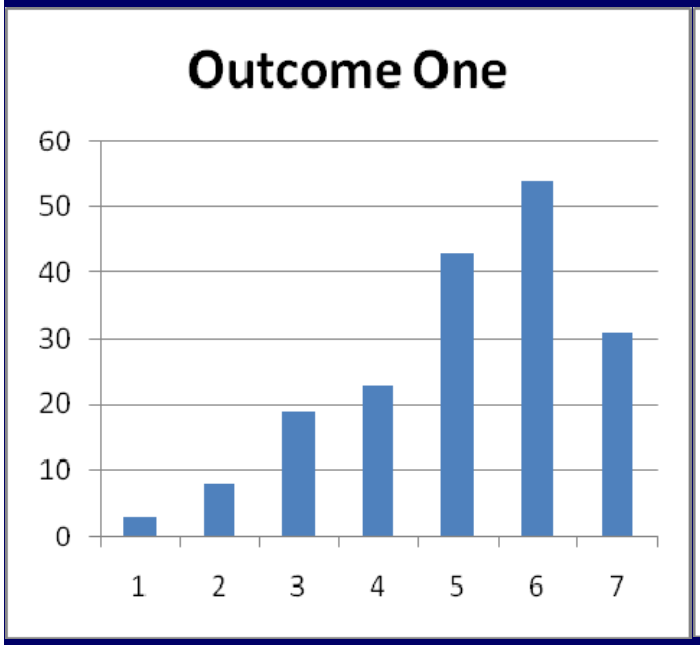
Outcome 3 Entrance: 06-07

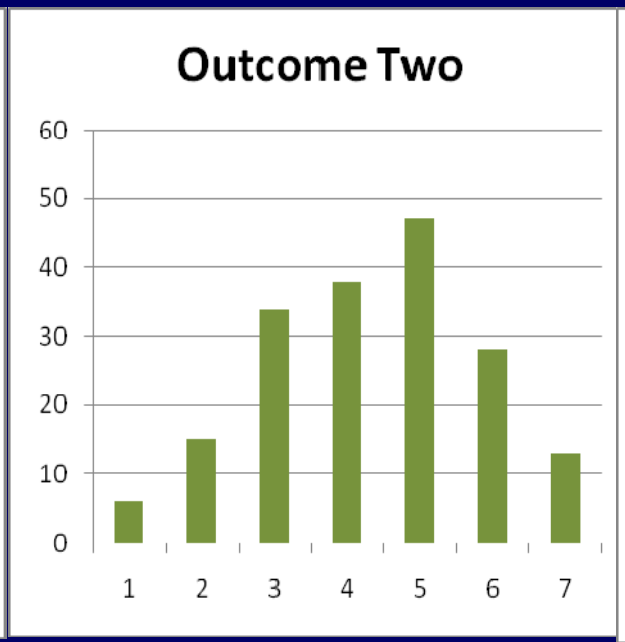


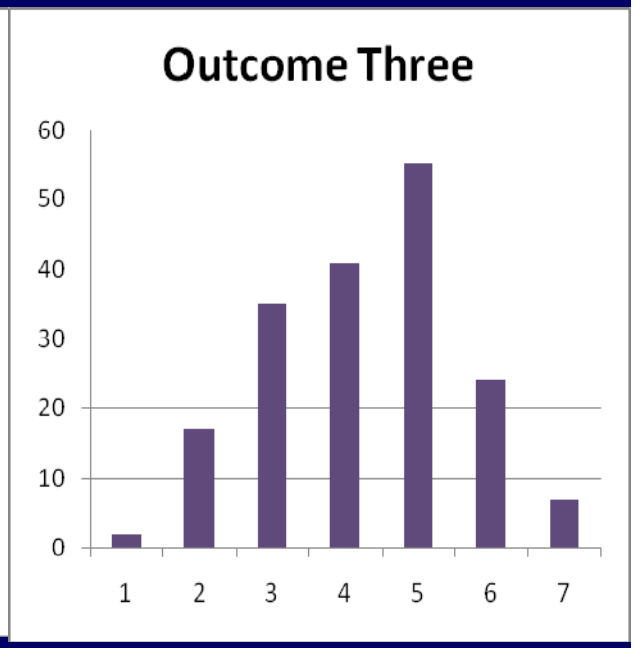
One State's – Part C Entry Data



What do you expect to see?

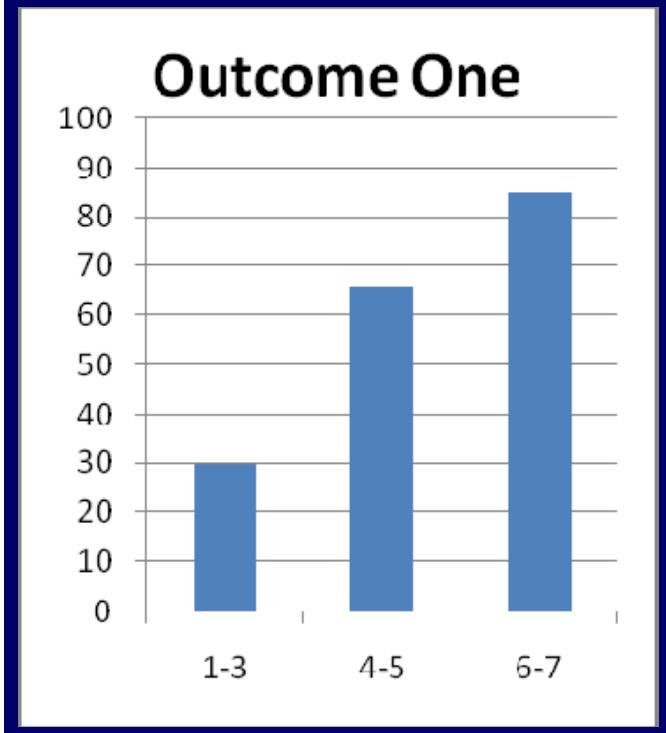


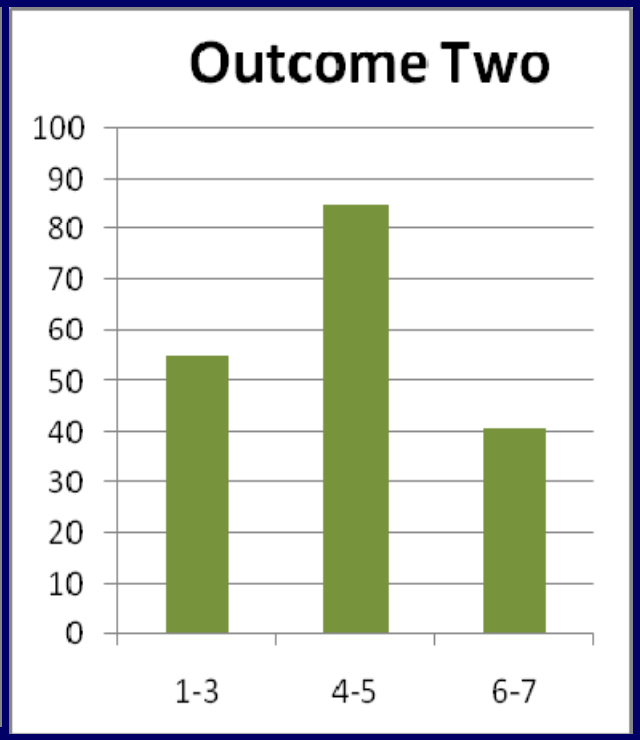


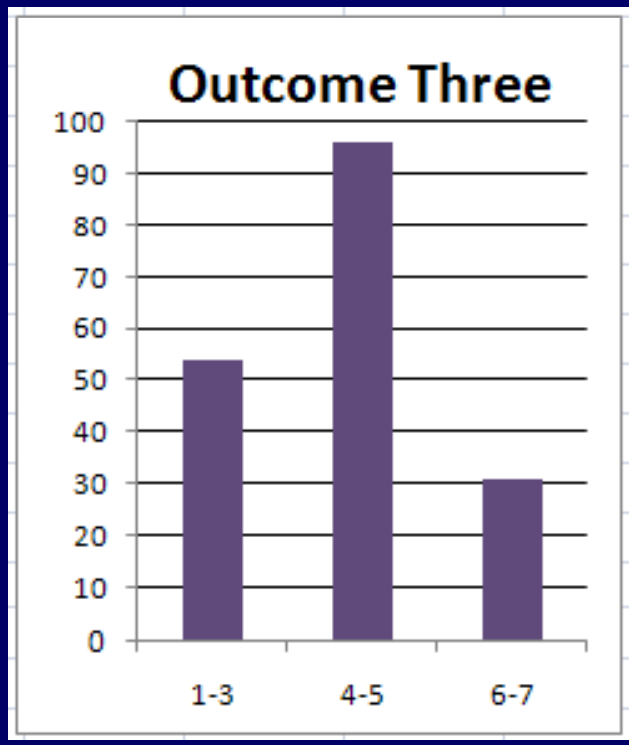


N= 181

Slice & Dice







N= 181

Dig Deeper

- What percentage of children received scores of 6 or 7 at entry on each of the three outcome areas?

  12.7% (23/181)

- What percentage of children received scores of 7 at entry on each of the three outcome areas?

  2.2% (4/181)

- The majority ***87% (153/181)*** of children show an entry rating of below 6 in at least one of the 3 outcome areas at entry.

Predictable Pattern 2

2. Functioning in one outcome area will be related to functioning in the other outcome area.

Rationale

- For many, but not all, children with disabilities, progress in functioning in the three outcomes proceeds together.

MN Crosstabulation with Progress Categories Outcome 1 & Outcome 2

1 SE → / ↓ 2 KS	a	b	c	d	e
a	4		1		
b	4	110	43	19	
c		34	187	57	7
d		13	46	113	66
d		24	23	41	193

Correlation: Outcome 1 x Outcome 2

	1	2	3	4	5	6	7	Total
1	38	7	5	1			2	53
2	24	68	18	8	3	4		125
3	8	25	40	10	8	1	1	93
4		7	30	22	9	4	3	75
5	2	5	15	33	42	15	10	122
6	1	7	14	20	34	48	9	133
7	1	2	5	3	20	21	40	92
Total	74	121	127	97	116	93	65	693

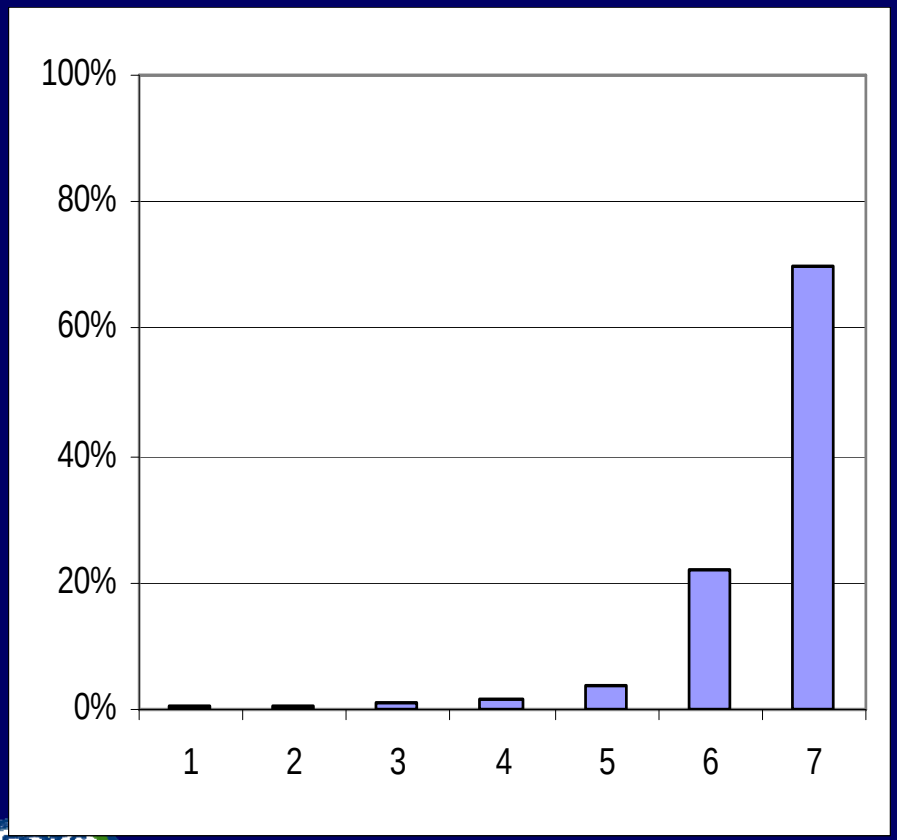
Correlation Coefficient

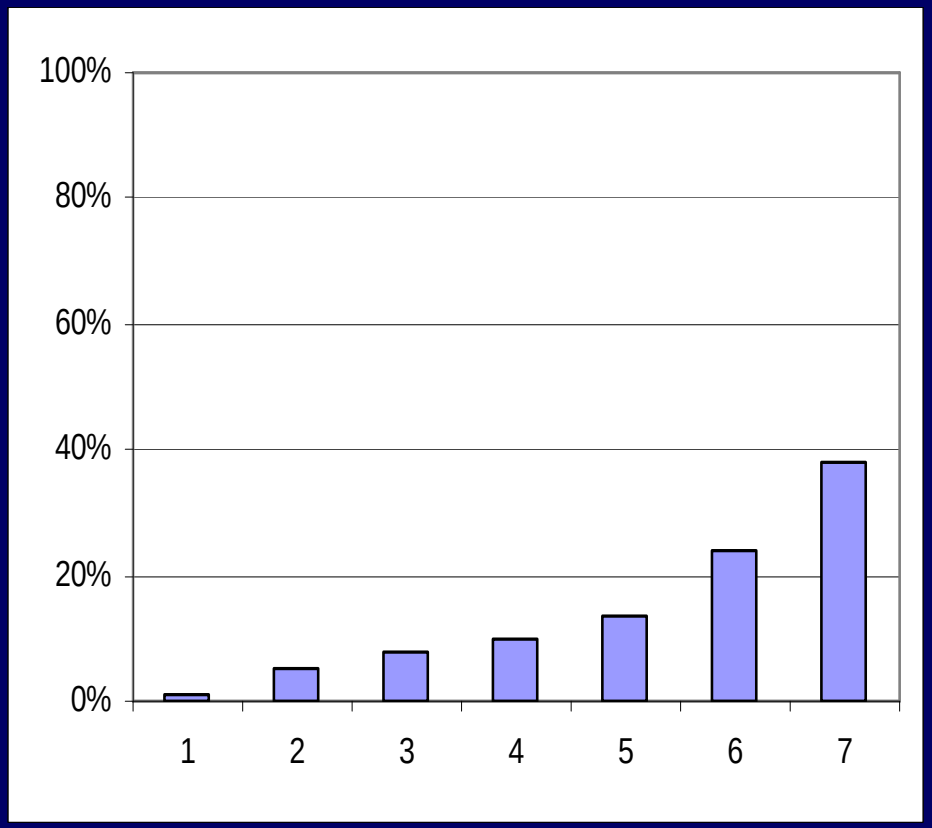
- Measure of extend of a relationship between two numbers
- Range 0 to 1 (can be negative) closer to 1 stronger the relationship
- Negative correlation means as on e set of numbers goes up, the other goes down.

1, 2, & 3 Entry Data N=181	*Outcome 1*	*Outcome 2*
Outcome 1	1	
Outcome 2	0.732611	1
Outcome 3	0.560618	0.720391

Predictable Pattern 3

3. If programs are serving similar kinds of children, entry ratings should be similar.





Comparison Data Coming

"This project is so important we can't let things that are more important interfere with it."

(Advertising/Marketing manager, United Parcel Service)

The validity of your data is questionable if …

- ….not all providers are not knowledgeable about in the COSF process
- …not all providers are careful with the COSF process

- The overall pattern in the data looks ‘strange’
  - Compared to what you expect
  - Compared to other data
  - Compared to similar states/regions/school districts

Take home message

- If you conclude the data are not (yet) valid, they cannot be used for program effectiveness, program improvement or anything else.

- What do you if the data are not as good as they should be?

- **Answer: Continue to improve data collection through ongoing quality assurance.**

Questions, Ideas?

- Please share…

Naomi.younggren2@us.army.mil