SE4831 Software Quality Assurance Assessing Review Effectiveness

inspection

- Objectives
  - Explain the problem with using bug counts as the sole measure of review effectiveness.
  - Explain the concept of capture-recapture experimental methods.
  - Explain how capture-recapture methods can be used to assess the effectiveness of formal inspections.
  - Using capture-recapture methods, estimate the remaining defects within a software artifact.

MS OE

How do we assess the effectiveness of reviews?

1. Count problems found.

2. Good feedback from stakeholders ⇒ No field problems.

3. Test failures Late

MSOE

Defects found

- Count the number of defects uncovered

Not very effective as only tool

Assessing Review Effectiveness

MSOE

**Effectiveness of the inspection**

Yield

$$Yield = \frac{TotalDefectsFound}{TotalEstimatedDefects}$$

MS
OE

Capture-Recapture

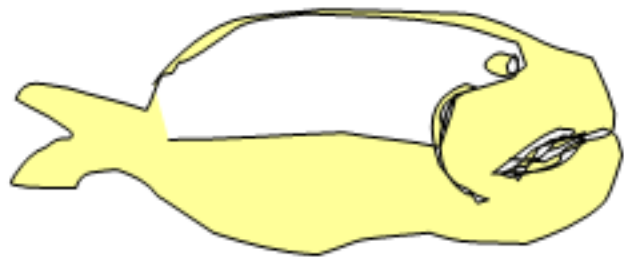- Capture-recapture methods are used for counting the total number of individuals in a population using two or more incomplete lists of those individuals

- Originially used in wildlife (birds, polar bears, wild salmon...) counting
  - Capture –> tag – > recapture – > calculate

MS OE

In class exercise

- Estimating the number of chocolates in a box before we eat them…
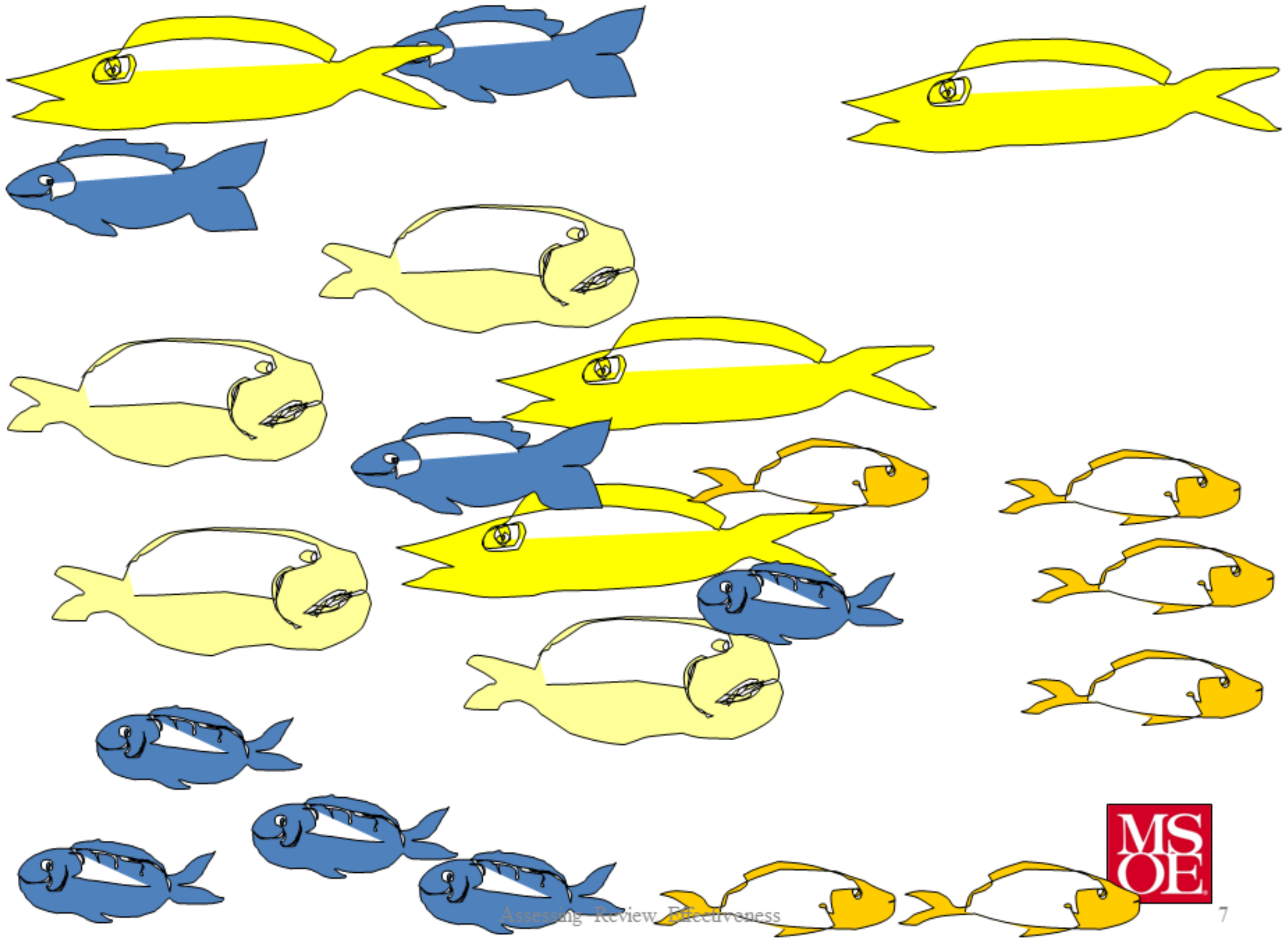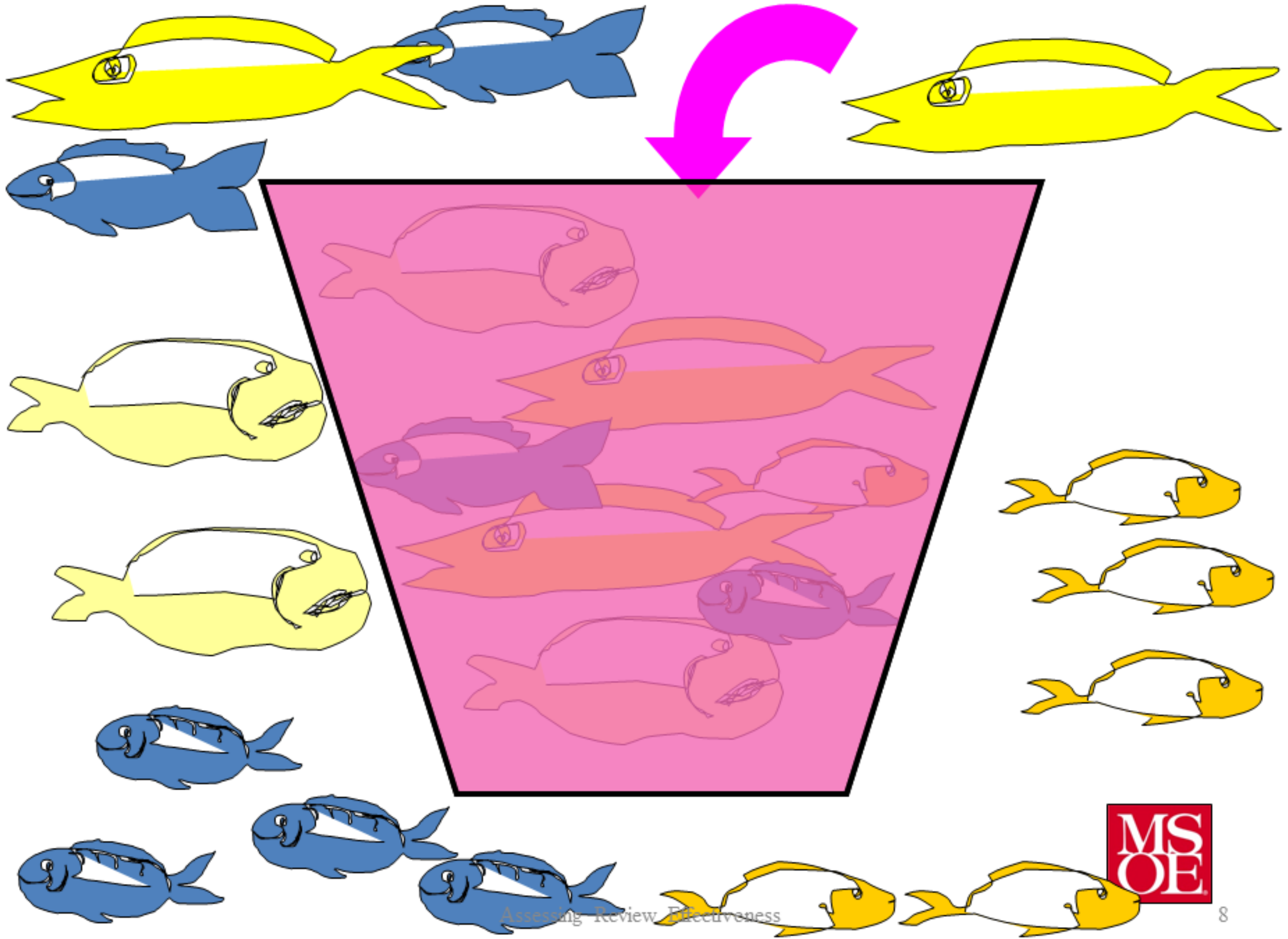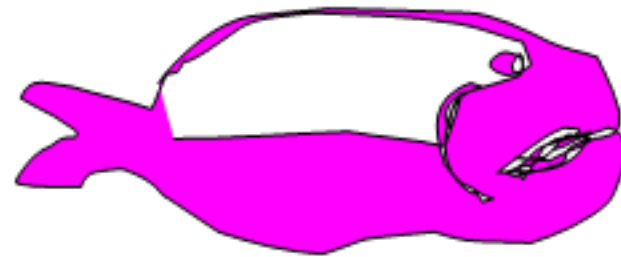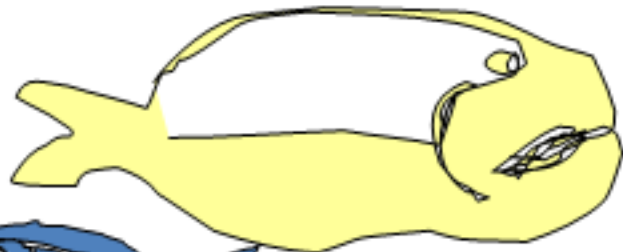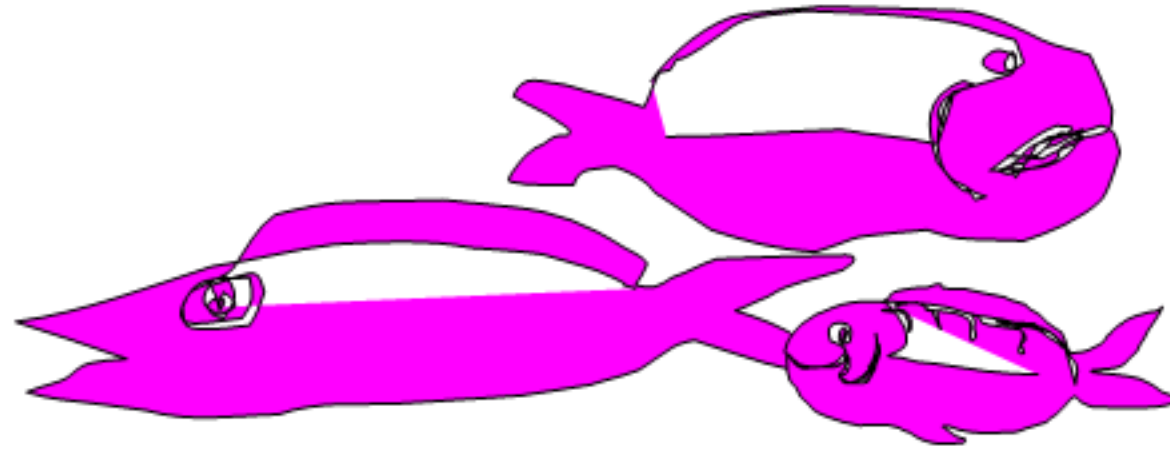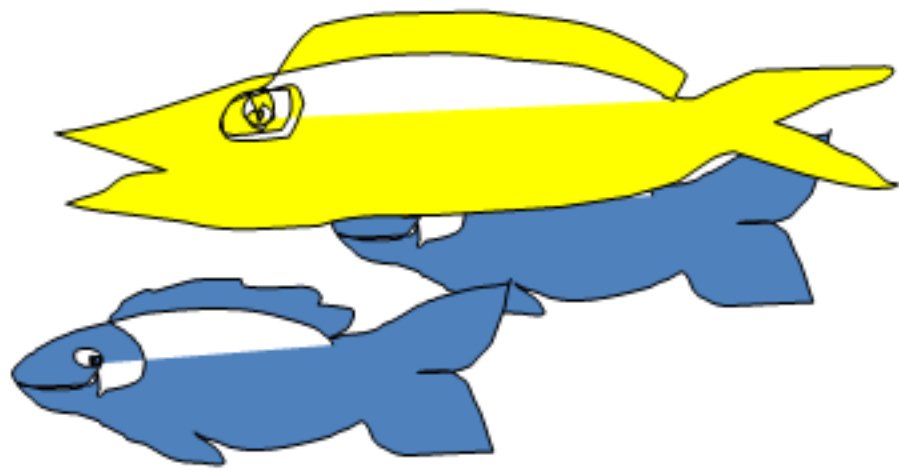
MS
OE

Two-source model

Two-source model

Two-source model

N=?

Y₁ Source Y

Z₁

Source Z

B

A

C

x=?

Assessing Review Effectiveness

15
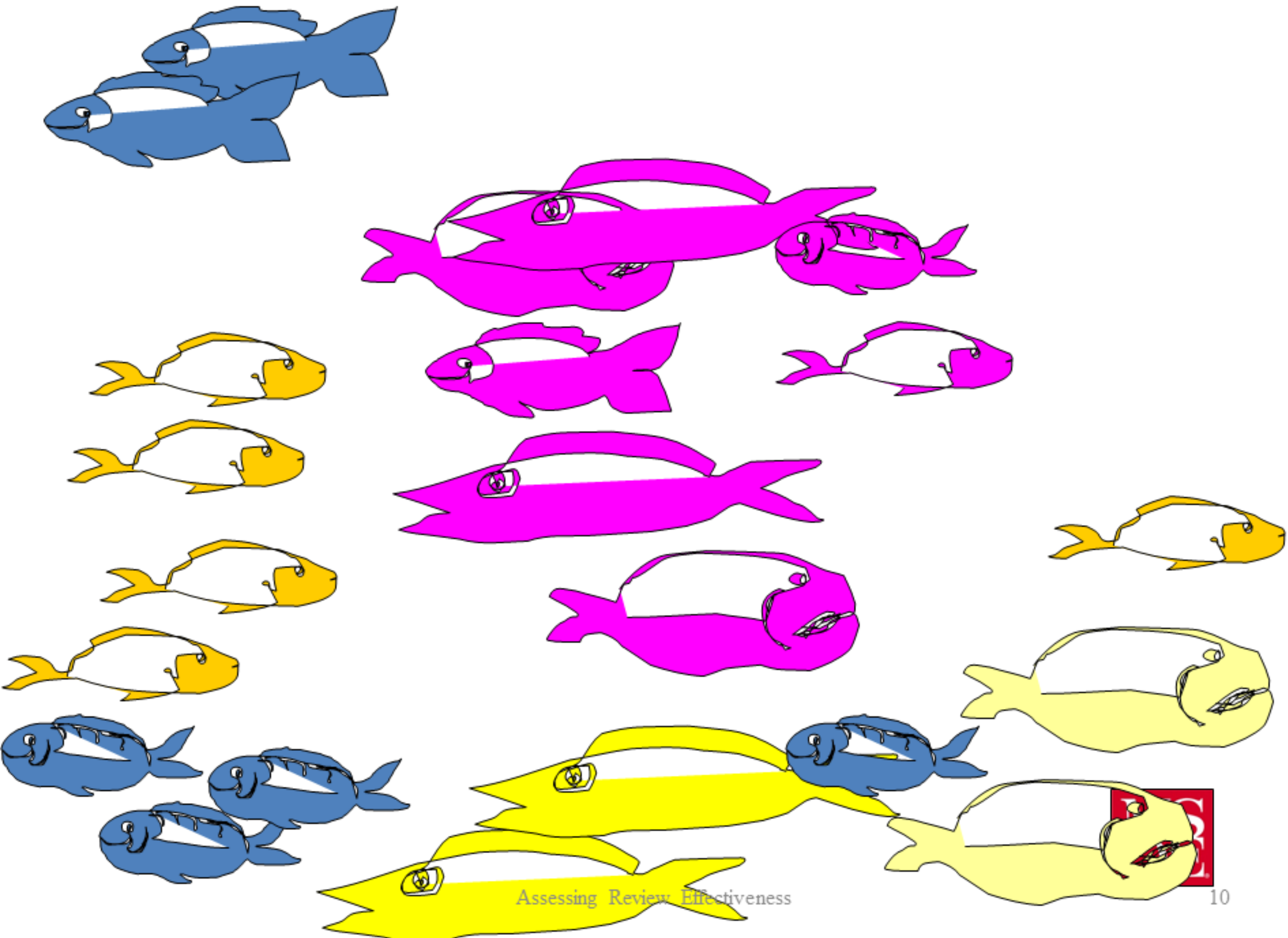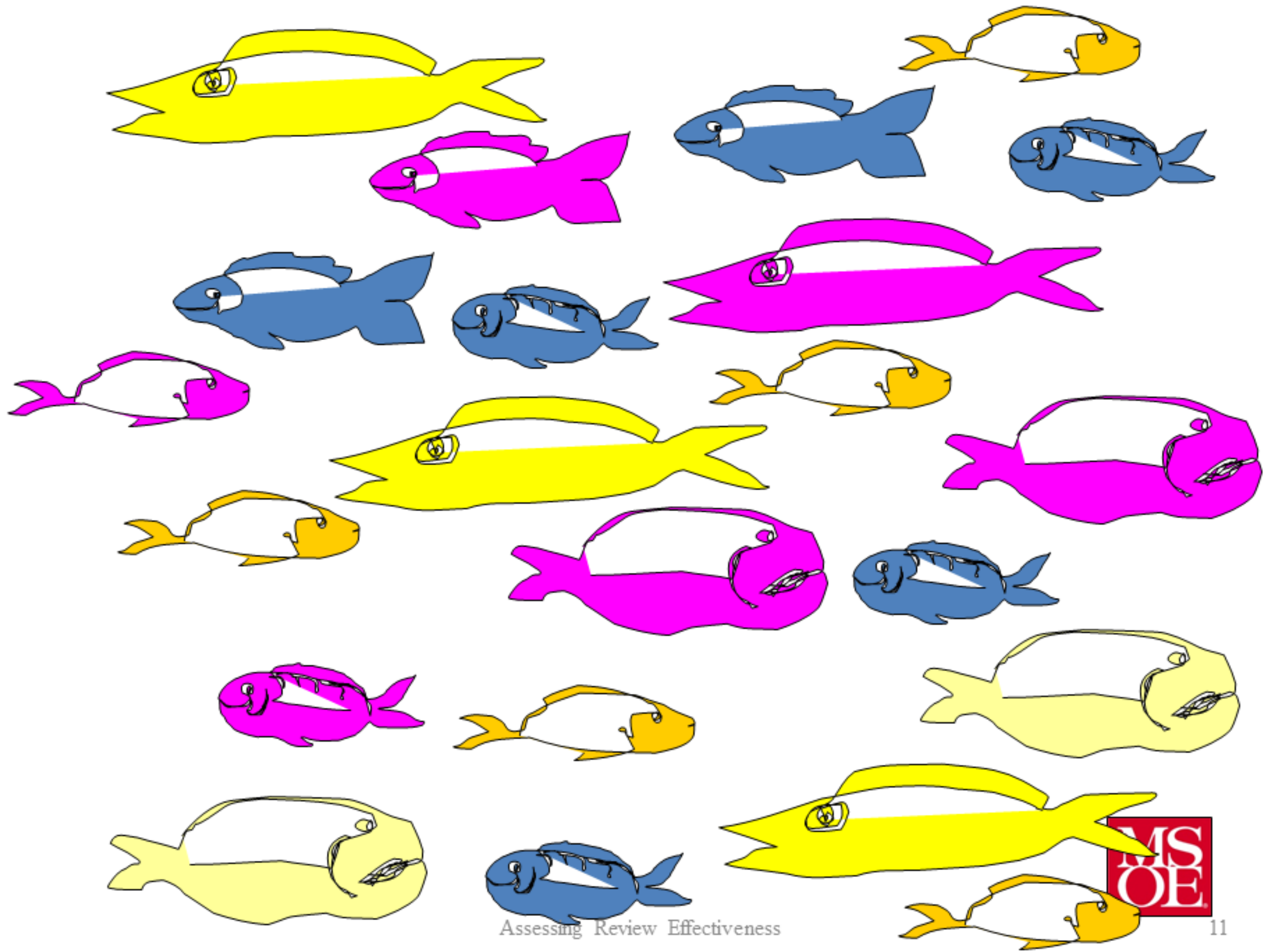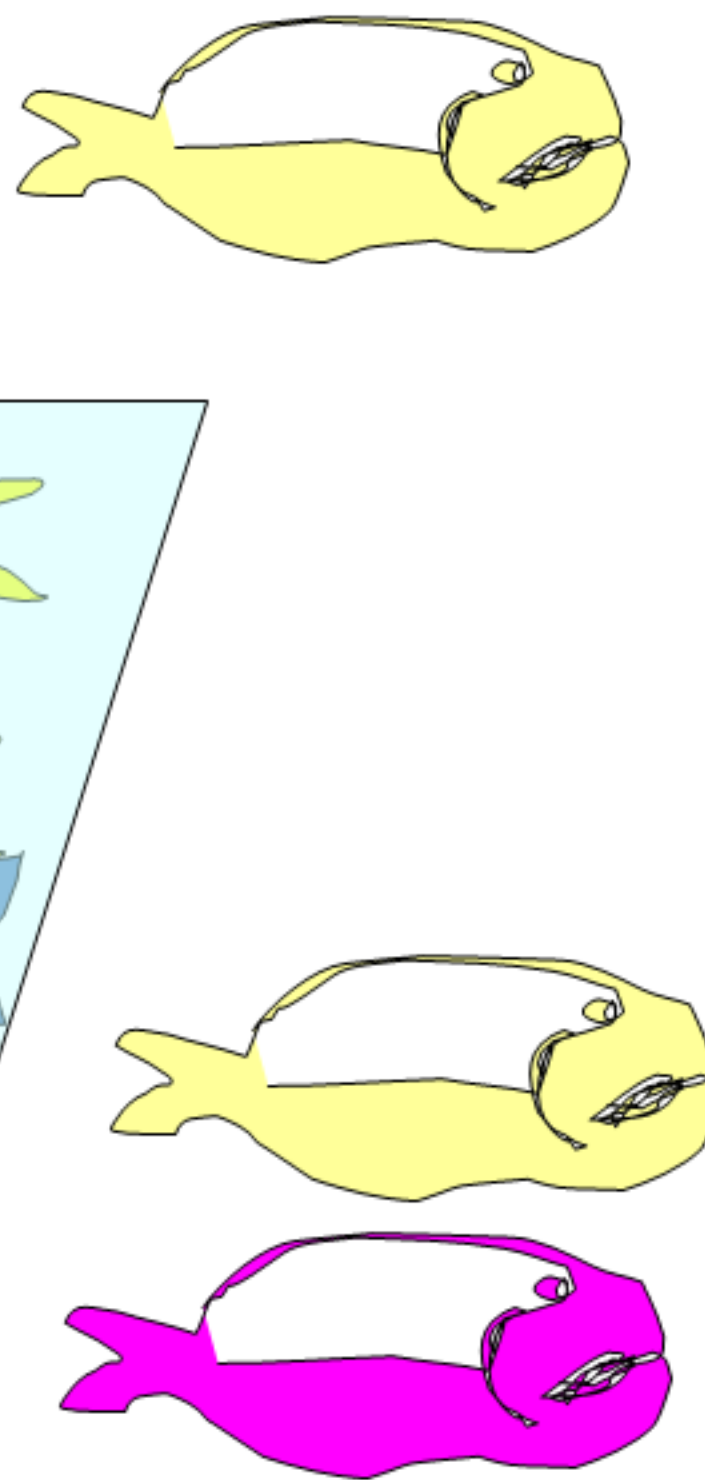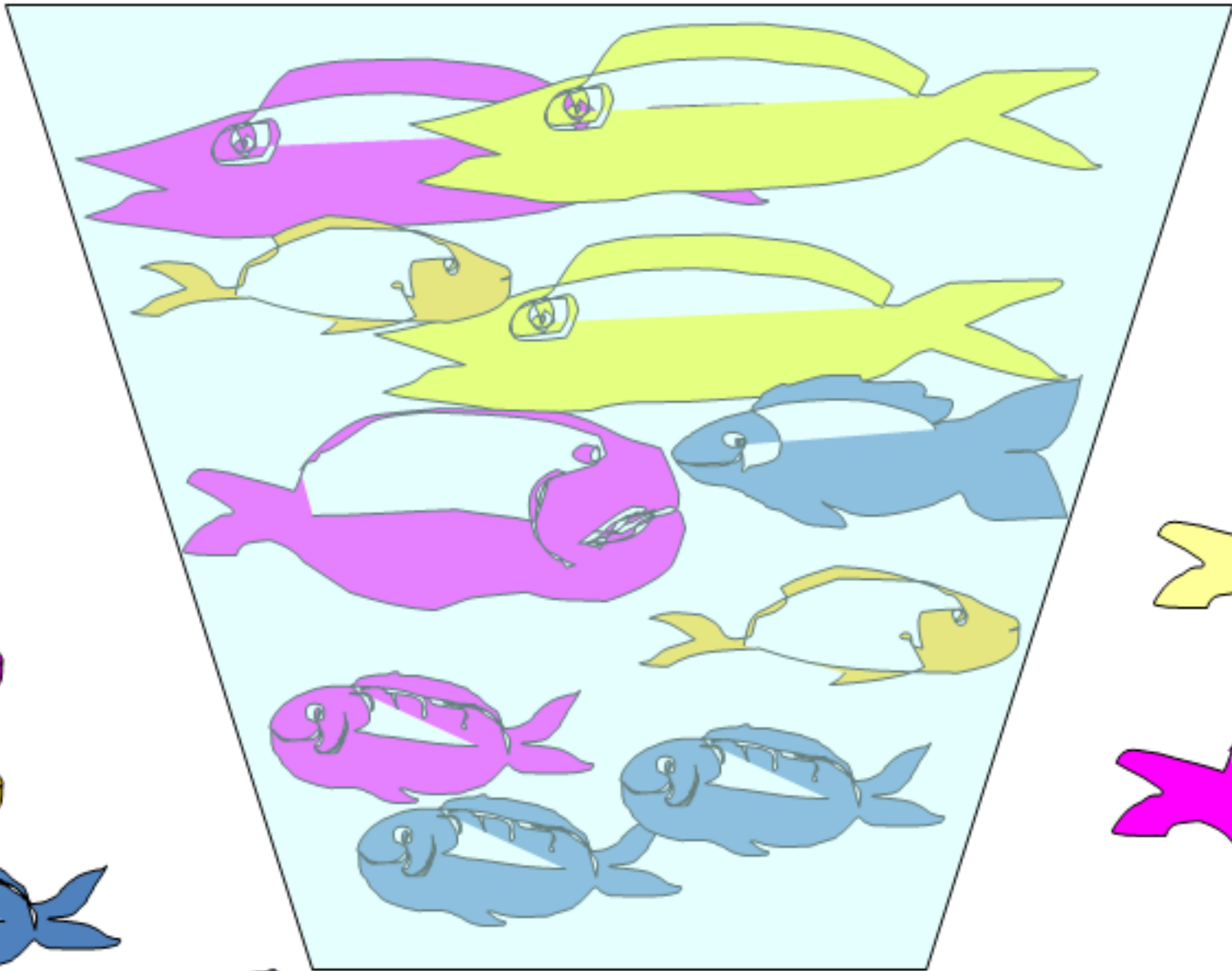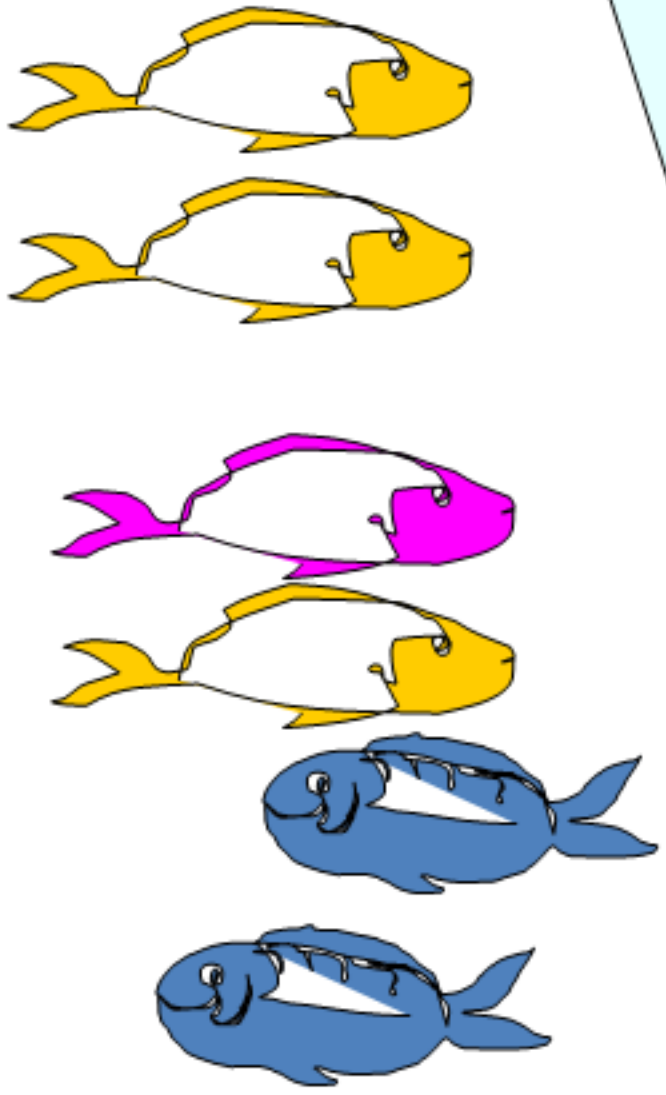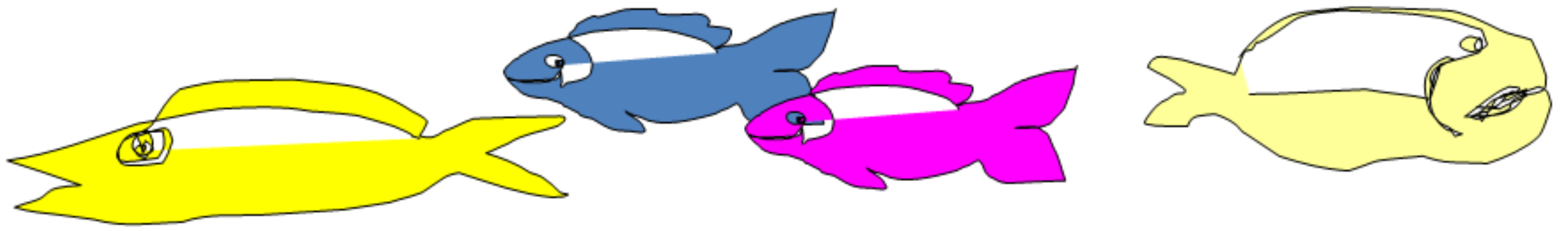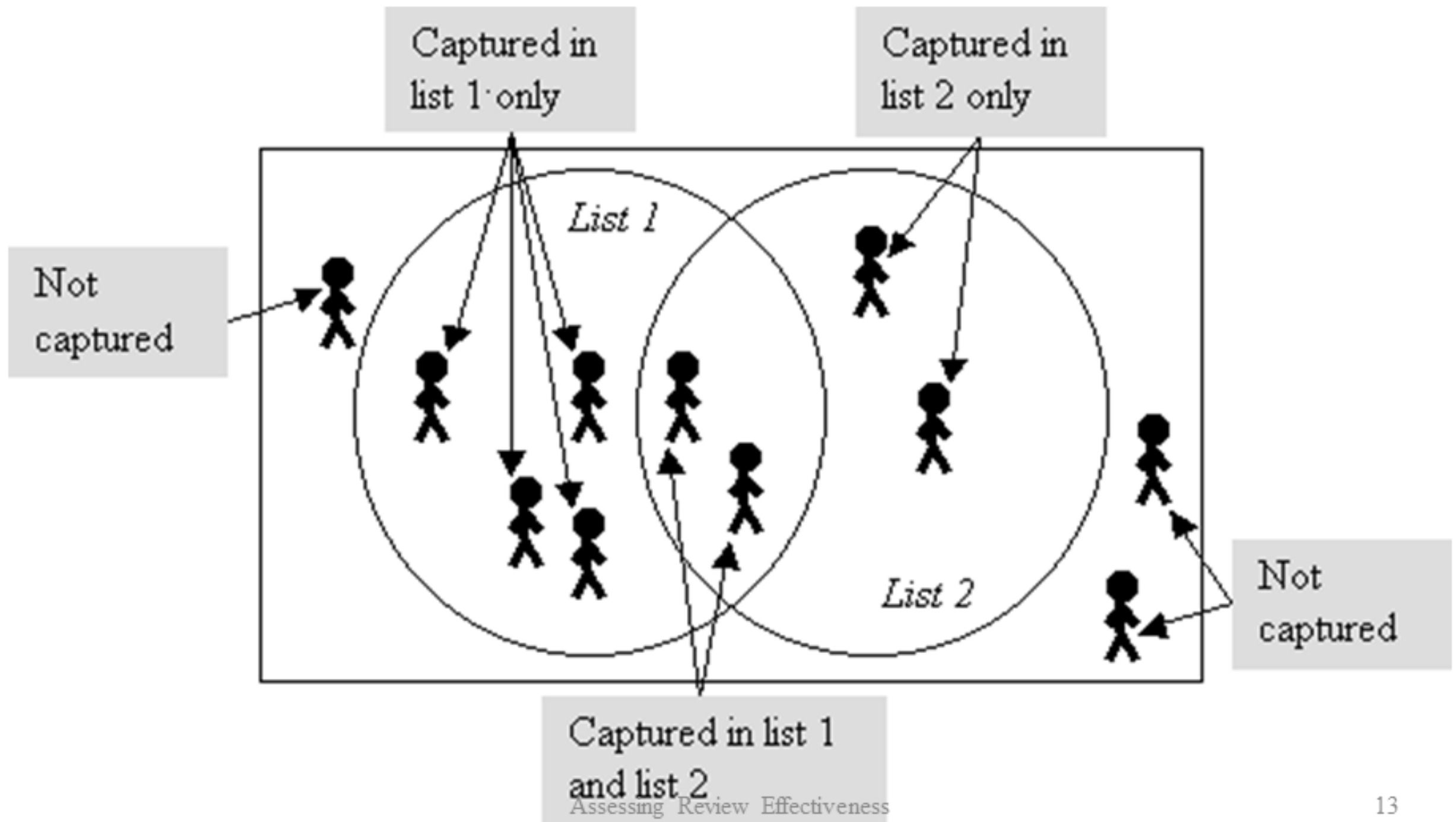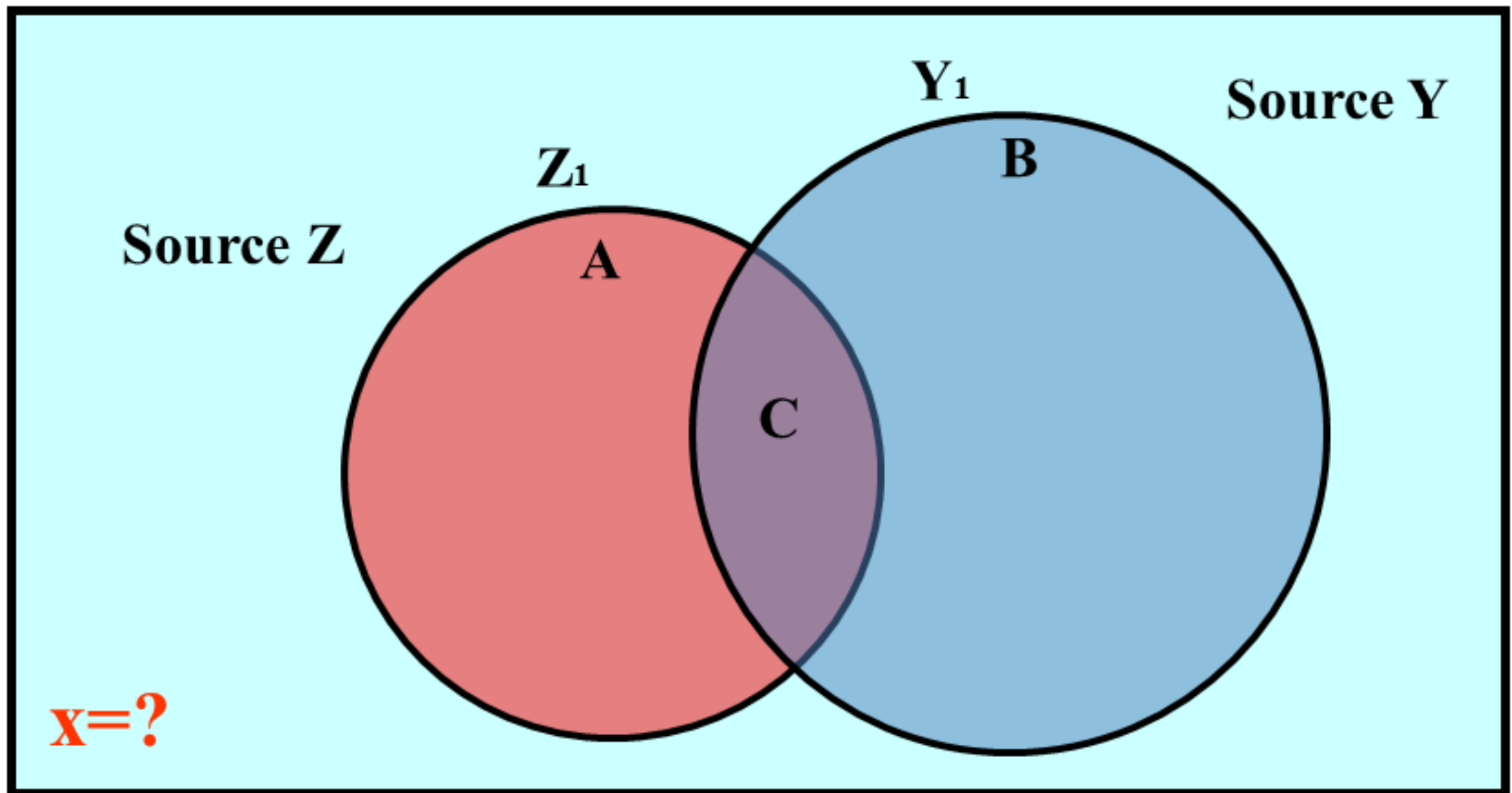
Two-source model Equations

Equations

$$N = A + B - C + x$$

$$\frac{C}{(A - C)} = \frac{B - C}{x}$$

$$Cx = (A - C) * (B - C)$$

$$Cx = AB - AC - BC + C^2$$

$$x = \frac{AB}{C} - A - B + C$$

$$N = A + B - C + \frac{AB}{C} - A - B + C$$

$$N = \frac{AB}{C}$$

$$N = \frac{A \cdot B}{C}$$

$$N = \frac{20 \times 20}{5} \Rightarrow 80$$

$12 \times 2 = 24$

$\frac{48}{2}$

R vs 00

What does this have to do with Fortran code?

Can use this to estimate defect counts.

MS OE

Capture – Recapture Inspection Steps

1. Count the number of defects found by one engineer / team. (A) — *Alice*

2. Count the number of defects found by a second engineer / team (B) — *Bill*

3. Count the number of commonly found defects and call this C. — *Common*

4. Estimate the total number of defects in the product as $T = A*B/C$

5. Calculate the total number of defects found as $F = A+B-C$ — *Found*

6. Estimate remaining defects as $R = T-F$

7. Estimate the yield $Y = F / T$

MS
OE

Team 1 Individual

- Will look at team #1 individual results

MS OE

Number	Location	Type (C, O, etc and number)	Description	Defects		Engineers Name			Total Finding
				Major	Minor	Engineer 1	Engineer 2	Engineer 3	
1	3.1	O5	Home user is not defined		1	1			1
2	3.2.1.1	O3	Not enough specifics. Are rooms controlled differently?	1					0
3	3.2.1.3	O5	More interaction details		1		1		1
4	3.2.1.4	C4	Security is out of project scope	1		1			1
5	3	C2	Not consistent using DH and DigitalHome			1			2
6	3.2.2.2	O2	Not detailed enough on "type of system"		1	1			1
7	3.2.2.4	O8	Should a technician have elevated privledges	1					0
8	3.2.2.4	O1	Does not link to a specific Use Case or document		1	1			1
9	3.3.1	O2	Specific data should be defined		1				0
10	3.3.4	O1	Link is not correct		1	1			1
11	3.3.4	C5	Should be reworded		1	1			1
12	3.3.5	O2	Define what minimizing cost is		1	1			1
13	3.3.6	C5	Spelling error (of [and] actual home)		1	1			1
14	3.3.7	C5	Should be a parentheses, not a slash		1	1			1
15	3.4	O5	Description of database		1		1		1
16	3.4.1	C5	Spelling error (require [and] internet)		1	1			3
17	3.4.1	O11	IPS is widely available isn't a requirement		1	1			1
18	3.4.1	C6	Does the web server have any other tasks?		1				0
19	3.4.2.1	O2	Not enough detail on user accounts		1	1			1
20	3.4.3.1	C5	DigitalHome is spelled wrong (DgitalHome)		1			1	3
21	3.4.3.1	C1, C6	Restricting to broadband goes against previous reqs		1	1			1
22	3.4.4.2	O11	Should specify the actual range with deviance		1			1	1
23	3.4.4.1	C6	How is the device programmed?		1			1	1
24	3.4.4.1	C6	How many sensors does the system support. How does the system work with the multiple areas?	1					0
25	3.4.4.2	C6	How is the device programmed?		1			1	1
26	3.4.4.2	C6	How many sensors does the system support. How does the system work with the multiple areas?	1					0
27	3.4.4.3	C4	Out of scope		1	1		1	2
28	3.4.4.4	C4	Out of scope		1	1		1	2
29	3.4.4.6	C4	Out of scope		1	1		1	2
30	3.4.4.5	C1	Appliance instead of small appliance		1				0
31	3	O4	Priorities not specified		1			1	1
32									0
			Totals	8	22		4	17	
			Unique Defects			12	2	4	

Assigning the defect counts (Overall)

- A set
  - The individual who found the most defects overall

- B set
  - All other engineers

- C set
  - The defects found by both A and B

$A \Rightarrow 21$

$B = 19$

$C = 9$

$$T = \frac{A * B}{C}$$

$$= \frac{21 * 19}{93} = \underline{44}$$

$$F = A + B - C$$
$$= 21 + 19 - 7 \Rightarrow 33$$

$$R \Rightarrow 21$$

$$\text{Yield} \rightarrow \frac{33}{44} \Rightarrow 2.75$$

$$Y_{BF} = \frac{21}{44} \rightarrow$$

$$Y_{BC} = \frac{5}{44} \rightarrow$$

$$Y_{13} = \frac{17}{44} \rightarrow$$

Review analysis

Review Analysis
(Major Only)

Number	Location	Type (C, O, etc and number)	Description	Defects		Engineers Name			Total Finding
				Major	Minor	Engineer 1	Engineer 2	Engineer 3	
2	3.2.1.1	O3	Not enough specifics. Are rooms controlled differently?	1					0
4	3.2.1.4	C4	Security is out of project scope	1		1			1
5	3	C2	Not consistent using DH and DigitalHome				1	1	2
7	3.2.2.4	O8	Should a technician have elevated privledges	1					0
24	3.4.4.1	C6	How many sensors does the system support. How does the system work with the multiple areas?	1					0
26	3.4.4.2	C6	How many sensors does the system support. How does the system work with the multiple areas?	1					0
27	3.4.4.3	C4	Out of scope	1		1		1	2
28	3.4.4.4	C4	Out of scope	1		1		1	2
29	3.4.4.6	C4	Out of scope	1		1		1	2
32									0
			Totals	8	0	5	0		
			Unique Defects			1	0	0	

Review analysis (Major Only)

Lets look at some data from your inspections

- Teams 3 & 6
  - Team 3 – 28 defects
    - 2 major
  - Team 6 – 46 defects
    - 1 major
  - 13 defects were found by both teams.
  - How many defects remain?

Lets take a look at Major Defects only

- Teams 3 & 6
  - Team 3
    - 2 major
  - Team 6
    - 1 major
  - 1 major was found by both teams

MS OE

Teams 2 and 5

- Team 2
  - Total of 18 found
  - 17 unique
  - 3 major

- Team 5
  - Total of 23 found
  - 9 major
  - 22 unique

- What is the value for A?
  - A. 3
  - B. 17
  - C. 23
  - D. 22
  - E. 9

Teams 2 and 5

- Team 2
  - Total of 18 found
  - 17 unique
  - 3 major

- Team 5
  - Total of 23 found
  - 9 major
  - 22 unique

- What is the value for B?
  - A. 3
  - B. 17
  - C. 23
  - D. 22
  - E. 9

MS OE

Teams 2 and 5

- Team 2
  - Total of 18 found
  - 17 unique
  - 3 major

- Team 5
  - Total of 23 found
  - 9 major
  - 22 unique

- What is the value for T?
  - A. 22
  - B. 40
  - C. 56
  - D. 72
  - E. 414

Statistically Refining Results
Unbiased Estimate

$$T = \frac{(A+1)*(B+1)}{C+1} - 1$$

Assessing Review
Effectiveness

MS
OE

Variance of Data

$$\sigma^2 = \frac{(A+1) * (B+1) * (A-C) * (B-C)}{(C+1)^2 * (C+2)}$$

MS
OE

95% Prediction Interval

$$UPI = T + 1.96\sigma$$

$$LPI = T - 1.96\sigma$$

Assessing Review
Effectiveness

MS
OE

For the team 2 and 5 review, what is the UPI and LPI?

$$\sigma^2 = \frac{(A+1)*(B+1)*(A-C)*(B-C)}{(C+1)^2*(C+2)}$$

$$UPI = T + 1.96\sigma$$

$$LPI = T - 1.96\sigma$$