

## Examining the Role of Testing in Teaching

Michelle Drouin &  
Michael Bendele  
Department of Psychology

## How do you presently use your tests?

- *Summative* assessment vs. *formative* assessment
- Larger, introductory psychology classes
  - Tell class answers after test OR
  - Don't tell answers after class

## My dilemma

- Medium-sized upper level (Area VI) course
- Many students (50%) failing exam
- How can I use my exams in a better way?
- *Formative summative* assessment (FSA; Wininger, 2005)

## Let's try it

- Mini-exam
- 1 minute to study
- 1 minute to take test
- 2 minutes to take retest

### How it really worked in the classroom

- Students had 90 minutes for exam (approx. 30 minutes MC, 60 minutes essay)
- Students completed FSA activity in next class period (30 minutes)
- Student completed Individual FSA activity after Unit 1 exam and Group FSA activity after Unit 2 exam

### How did I evaluate effectiveness of activity?

- Exam over same MC material 5 weeks after they completed retest activity in class
- Student evaluations

### 5-week posttest performance— Individual FSA

- Scores on Unit 1 exam ranged from 10 to 25 ( $M = 19.10$ ,  $SD = 3.57$ ).
- Scores on 5-week posttest ranged from 14-25 ( $M = 19.32$ ,  $SD = 3.20$ ).
- This is a 1.2% increase, which is not significant.

### 5-week posttest performance—Group FSA

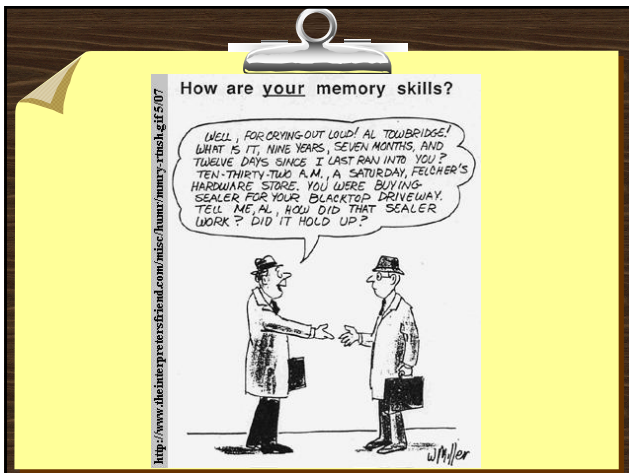
- Scores on Unit 1 exam ranged from 12-22 ( $M = 17.94$ ,  $SD = 2.78$ ).
- Scores on 5-week posttest ranged from 16-24 ( $M = 19.83$ ,  $SD = 2.30$ ).
- This is a 10% increase, which is highly significant.

Table 1  
Means, Standard Deviations, and Significant Differences for Students' Evaluations of Individual vs. Group FSA Activity

Measure	Individual FSA		Group FSA	
	M (SD)	M (SD)	t (29)	d
I enjoy working with classmates	3.90 (0.96)	4.30 (0.70)	-1.56	-.48
I prefer working independently	2.73 (0.98)	2.33 (0.99)	1.44	.41
I want a study group for the next exam	3.56 (0.90)	4.23 (0.94)	-2.88**	-.73
I found the exercise useful	4.00 (0.87)	4.40 (0.56)	-1.93	-.55
I enjoyed the structure <sup>a</sup> of the exercise	3.17 (1.1)	4.45 (0.57)	-5.26***	-1.46
This was a good review of key concepts	4.07 (0.83)	4.27 (0.73)	-0.88	-.26
I feel closer <sup>b</sup> to my classmates	2.83 (0.95)	3.93 (0.69)	-4.96***	-1.32
I want another exam review	4.53 (0.68)	4.17 (0.91)	1.65	.45
I prefer group over individual review	--	4.47 (0.82)--	--	--

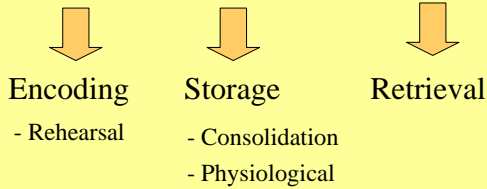
Note. N = 30.  
<sup>a</sup>Structure was either "working independently" or "working with classmates" depending on activity (individual or group). <sup>b</sup>Because of this exercise. Scale from 1 (strongly disagree) to 5 (strongly agree).  
 \*\* p < .01  
 \*\*\* p < .001

- ### What does this all mean for me?
- How are you presently using your exams?
  - Do you want to increase students' knowledge?
  - Do you want to increase classroom collegiality?
  - Group-based exam reviews might be a good way to do this



- ### After an Exam ...
- I studied for hours – why isn't my grade higher!
  - I have read the chapter 4 times – why did I get an F?
  - I studied- why am I failing your class?

### Three Basic Memory Processes:



**Testing Effect:** the phenomenon of improved performance from taking a test

1. Promotes learning by making later studying more effective.

**Testing Effect:** the phenomenon of improved performance from taking a test

1. Promotes learning by making later studying more effective.
2. The act of testing (retrieval) itself influences learning & retention of information.
  - Testing vs. additional studying

### Some Terminology

- **Metacognition:** your knowledge of and awareness about cognitive processes
- **Metamemory:** memory about your own memory
- **Judgments of learning (JOL):** an estimate of how well something has been learned.

For **each item** on the exam indicate your level of confidence in your answer next to the question.

1-----2-----3-----4-----5-----6-----7

Not at all                      Some                      Very  
Confident                      What                      Confident

\_\_\_ 1. T F A theory is more complex than a hypothesis.

\_\_\_ 2. T F The Nuremberg Code was developed based on Nazi war crime trials that took place at the end of World War II.

\_\_\_ 3. T F Multivariate experimental designs have more than one dependent variable.

### Three Basic Memory Processes:

Encoding                      Storage                      Retrieval

**Examples:** tests, quizzes, in class questions [e.g., CATs], review sessions [with participation], & homework

### References

- Glenn, D. (2007). You will be tested on this. *The Chronicle of Higher Education*, 53(40), A14.
- Roediger, III, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1, 181-210.
- Wininger, S. R. (2005). Using your tests to teach: Formative summative assessment. *Teaching of Psychology*, 32, 164-166.