

Test Preparation

Writing Questions

Grading and Feedback

Improving Exams

**Design table of specification.**  
**Write structured white-space – rational argumentation essay questions.**  
**Apply good grading practices.**  
**Improve tests using surveys and indices of difficulty and discrimination.**

Copyright Harvey J. Brightman 1

How to construct a table of specification for exams (and share it with students).

Topic\Level	Explain, Interpret, Distinguish	Cases (Application)	Critical Thinking: Analyze, Propose
Decision Support	10		20
Decision Making		12	4
Creativity			5
Root Cause Analysis		21	
Forecasting	20		8

**Percentages reflect time spent. All objectives in syllabus.**

WSJ 3 August 1988

**Comment on this article. What are the implications? What is the message? 15 points**

**Dow Chemical  
To Buy Back  
5% of its Shares**

Dow Chemical flush with cash after sharp earnings increase said it will repurchase more than 5% of its common shares outstanding.

Dow said it believes the stock is undervalued. Dow is trading at more than seven times analysts' estimates of \$12 a share.

"Given the price of the stock it's a good use of the cash." said an analyst with Drexel Burnham Lambert Inc.

Four More Paragraphs.

Copyright Harvey J. Brightman

3

WSJ article question generated wide variation in grades among finance professors.



**Comment on article. What's the message!!**



**Use unambiguous verbs @different levels  
Meaningful:**

**Analysis:**

**Evaluate:**

**Synthesis**

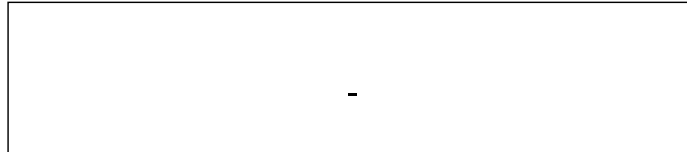
Copyright Harvey J. Brightman

4

Even with well-stated questions we still have problems.

**Describe and defend one weakness of MCDM. Suggest and defend an improvement to overcome your weakness.**

**Students ramble on, hard to grade reliably, and takes forever.**



Preparing Exams

Number of tests  
Table of specification

Writing Questions

Structured white-space  
- critical thinking format  
short answer

Grading Exams

Improving Exams

Here are some examples of **unstructured** short answer essay questions

1. **Describe and defend one weakness of MCDM. Suggest and defend an improvement to overcome your weakness. {DA 06, 09-011}**

List objectives that form basis for question.

2. **Discuss and defend two similarities and two differences between Porter's competitive strategies: (1) cost and (2) differentiation. {PCS 1,3}**

Copyright Harvey J. Brightman

7

3. **Shown above is a line graph. Do the data exhibit a meandering or seasonal pattern? Defend. {FM 40}**

4 **Kindig (1998) published a specific proposal for designing the financing of health care around rewarding improvements in population health rather than paying providers for what they do. Discuss the impact on patients and health care providers. {PHP 2, 5}**

5 **The central limit theorem is critical to constructing a confidence interval on a population mean. Discuss. {BS 7, 9}**

Copyright Harvey J. Brightman

8

A brilliant all-bluff answer by one of my undergraduate students.

1. The central limit theorem is critical to constructing a construct confidence interval on a population mean.
2. Dr. Deming noted that central limit theorem was the key to drawing conclusions about a population mean.
3. As you said in class, "the central limit theorem is a critical topic."
4. The central limit theorem says that .....
5. Finally, the CLT is also important to hypothesis testing which is similar to confidence intervals in that ...

Copyright Harvey J. Brightman

9

### Using Mason/Mitroff model of rational thinking to learn critical thinking

Element	Definition	Example
<b>Claim</b>	A belief or statement, <b>not a fact</b>	Yes the CLT is critical to constructing confidence intervals for a population mean.
<b>Data</b>	Support for your claim- why should we believe you? What are your reasons? Start with " <b>because</b> ."	<b>because</b> CLT ensures that sampling distribution is normally distributed. We can use t-distribution to compute width of CI.
<b>Warrant</b>	Support for your data; Why should others believe the data presented in the "because." Start with " <b>given</b> ."	My data are correct <b>given</b> that text states that you can only t-distribution if sampling distribution is normally distributed.
<b>Rebuttal</b>	Weakness in either data or warrant statements. Start with " <b>unless</b> ."	My claim is correct <b>unless</b> population is normally distributed. Then sampling distribution will be normal shaped and we don't need CLT.
Mason and Mitroff, <i>Challenging Strategic Planning Assumptions</i> , John Wiley, 1981. Toulmin Model: Copyright Harvey J. Brightman <a href="http://www.concentric.net/~Creyn266/COMM335/Toulmin.htm">http://www.concentric.net/~Creyn266/COMM335/Toulmin.htm</a>		

10

Another example using the Mason/Mitroff model		
Element	Definition	Example
<b>Claim</b>	A belief or statement, <b>not a fact</b>	Starbucks decline in customer satisfaction scores are due to 1. 2
<b>Data</b>	Support for your claim- why should we believe you? What are your reasons? Start with <b>"because."</b>	My first reason is correct <b>because</b> .....
<b>Warrant</b>	Support for your data; Why should others believe the data presented in the "because." Start with <b>"given."</b>	My data or reasons are correct <b>given</b> they are based on the empirical data presented in case.
<b>Rebuttal</b>	Weakness in either data or warrant statements. Start with <b>"unless."</b>	My claim is correct <b>unless</b> Starbucks data are tainted or were improperly collected, computed, or interpreted.
Mason and Mitroff, <i>Challenging Strategic Planning Assumptions</i> , John Wiley, 1981. Toulmin Model: <small>Copyright Harvey J. Brightman</small> <a href="http://www.concentric.net/~Creyn266/COMM335/Toulmin.htm">http://www.concentric.net/~Creyn266/COMM335/Toulmin.htm</a>		11

Rewriting question #1 as structured white space

**Claim: One serious weakness of MCDM is**

**Data: This is a serious weakness because**

**Claim: My improvement for overcoming this weakness is**

**Data: This would minimize the weakness because**

Copyright Harvey J. Brightman

12

The remainder of question #1 as structured white space

**Warrant: My data for my improvement are correct given**

**Rebuttal: My improvement for reducing the weakness will work unless**

Copyright Harvey J. Brightman

13

Rewriting question #2 as structured white space

**Claim: Similarity 1 is**

**Data: This is an important similarity because**

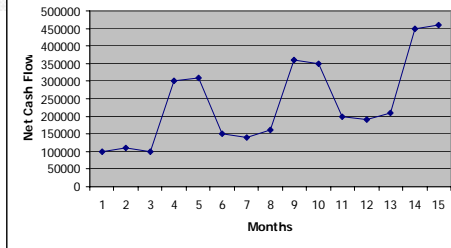
**Claim: Difference 1 is**

**Data: This is an important similarity because**

Copyright Harvey J. Brightman

14

Rewriting question #3 as structured white space



**Select one of the answers and complete.**

**1. The line graph indicates a meandering pattern because**

**2. The line graph indicates a seasonal pattern with a pattern of length \_\_\_\_\_ months because**

Rewriting question #4 as structured white space

**If Kindig's proposal were implemented, one benefit for patients would be**

:

**This benefit would occur because**

**One disadvantage for patients would be:**

**This disadvantage would occur because**

The remainder of question #4 as structured white space

One specific change in the behavior of health plan providers would be

This would happen **because**

The change in behavior would happen **unless**

Rewriting question #5 as structured white space

**Select one of the answers and complete.**

**1. It would never be possible to construct CI because**

**2. It would always be possible to construct CI because**

**3a. It would only be possible to construct CI when**

**3b. This is true because**

## Why use structured white space with Mason/Mitroff model?



**Minimizes “hunt and peck” grading answers. Reduces time and improves reliability. Give partial credit.**



**Provides model of how to think critically.**

Copyright Harvey J. Brightman

19

Preparing Exams

Writing Questions

Grading Exams

Improving Exams

**Number of tests questions per test, and table of specification**

**Structured white-space and critical thinking**

**Good practice and scoring rubrics**

Copyright Harvey J. Brightman

20

## Some recommendations on grading essay exams.

**Do “blind” grading.**

**Use detailed scoring key. Skim several answers to determine if you need to change scoring key. See next slide.**

**Grade one question at a time. When finished again grade one question to check for grading consistency.**

**Shuffle papers before grading next question.**

**Use scoring rubrics. See slides 24-26.**

Copyright Harvey J. Brightman

21

## Here is a partial scoring key.

**How are Van Gundy’s why problem reframing method and defining decision objective similar and different?  
[CM3] 10 points**

**One similarity between these two concepts is**

***4-5 points for anyone of the following; two points for any other partially correct answer; zero if left blank or totally incorrect***

- a We use both methods in the intelligence activity of solving managerial problems.**
- b. Or both are qualitative methods**
- c. Or neither method is used in disturbance problem solving.**
- d. All other correct answers that are not shown above.**

Copyright Harvey J. Brightman

22

## More on grading issues.

**Return promptly and/or provide key after quiz or place on portal.**



**for question reviews and return quickly.**

**Selected written comments highly motivating.  
But .....**

Copyright Harvey J. Brightman

23

## Beyond grading – providing detailed and informative feedback



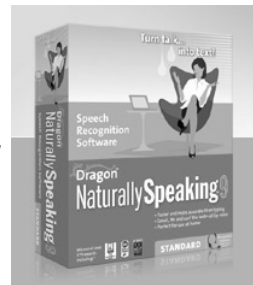
use



or



or



Copyright Harvey J. Brightman

24

Here is analytic scoring rubric for multi-page essay.

<b>Knowledge</b> 18 points	<b>Answers all issues with evidence, examples and elaboration</b> 17-18	<b>Answers most issues but with little or no evidence etc.</b> 15-16	<b>Doesn't answer all issues or has errors of fact</b> 11-14
<b>Organization</b> 9 points	<b>Very logical sequence of paragraphs.</b> 9	<b>Data-claim flow or sequence hard to follow.</b> 7-8	<b>No logical sequence to answer.</b> 5-6
<b>Grammar</b> 3 points	<b>Grammatically correct prose.</b> 3	<b>Most grammar correct.</b> 2	<b>Many spelling errors and poor grammar</b> 1

Copyright Harvey J. Brightman

25

Can also use analytic scoring rubrics for case analyses – partial scoring rubric.

Criteria	Above Average	Average	Below Average
Identifies main issues 20%	States clearly and defends	States clearly but does not defend	Incomplete or incorrect
Application of course content 20%	Frequent use of relevant frameworks	Some use of relevant frameworks	Uses incorrect frameworks
Analysis 30%	Insightful with proper defense	Good but with little or no defense	Rehashes data in case

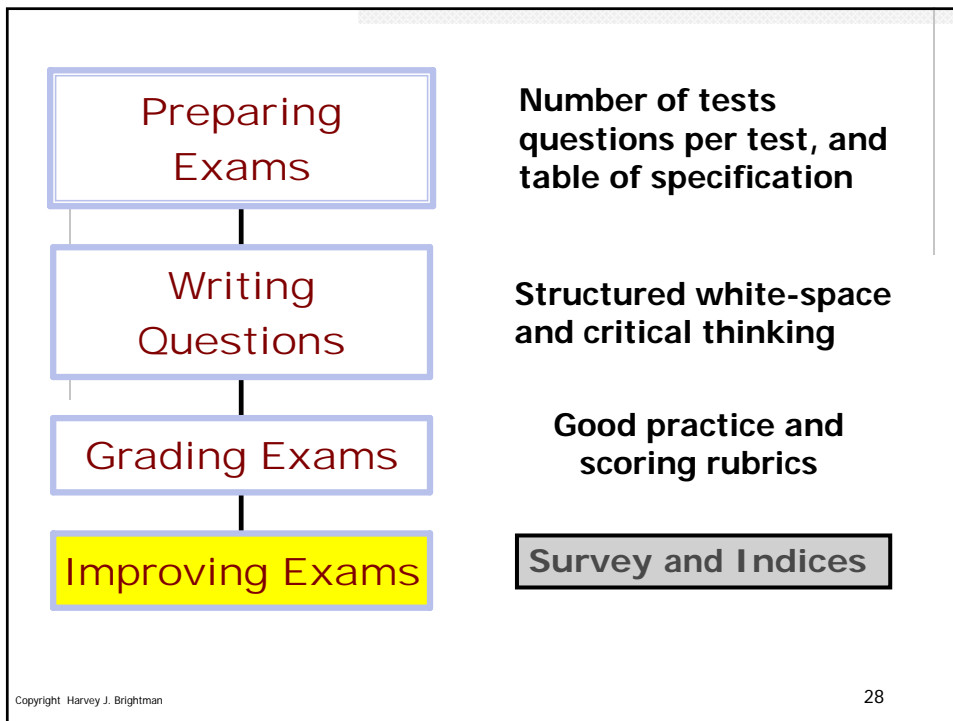
Copyright Harvey J. Brightman

26

## Can also use analytic scoring rubrics for assigning team contribution points.

CATEGORY	4 Points	3 Points	2 Points	1 Point
Contributions	Routinely provides useful ideas when participating in the group and in classroom discussion. A definite leader who contributes a lot of effort.	Usually provides useful ideas when participating in the group and in classroom discussion. A strong group member who tries hard!	Sometimes provides useful ideas when participating in the group and in classroom discussion. A satisfactory group member who does what is required.	Rarely provides useful ideas when participating in the group and in classroom discussion. May refuse to participate.
Focus on the task	Consistently stays focused on the task and what needs to be done. Very self-directed.	Focuses on the task and what needs to be done most of the time. Other group members can count on this person.	Focuses on the task and what needs to be done some of the time. Other group members must sometimes nag, prod, and remind to keep this person on-task.	Rarely focuses on the task and what needs to be done. Lets others do the work.
Working with Others	Almost always listens to, shares with, and supports the efforts of others. Tries to keep people working well together.	Usually listens to, shares, with, and supports the efforts of others. Does not cause "waves" in the group.	Often listens to, shares with, and supports the efforts of others, but sometimes is not a good team member.	Rarely listens to, shares with, and supports the efforts of others. Often is not a good team player.
Quality of Work	Provides work of the highest quality.	Provides high quality work.	Provides work that occasionally needs to be checked/redone by other group members to ensure quality.	Provides work that usually needs to be checked/redone by others to ensure quality.

Copyright Harvey J. Brightman 27



## Differences in norm and criterion based tests

### Norm-Based Tests (SAT)

Goal: ensure grade distribution

25% or more students not finished at period's end.

Exam time: 3 or 4 times longer than for faculty to complete.

Short-answer essay questions at meaningful-integrated level requiring four or five sentences should take about 6-7 minutes.

MC questions at rote level should take about .50 to .75 minutes.

### Criterion-Based Tests (CPA)

Goal: Ensure mastery

Less than 10% not finished at period's end.

Exam time: 5 or 6 times longer than for faculty to complete.

Short-answer essay questions at meaningful-integrated level requiring four or five sentences should take about 9-10 minutes.

MC questions at rote level should take about 0.75 to 1.45 minutes.

Copyright Harvey J. Brightman

29

## Brief survey for criterion-based test

1. Have exams reflected topic objectives?
2. Have exams reflected important subject content?
3. Has there been sufficient time to complete exams?
4. Has instructor reacted positively to student concerns on questions, grading, etc.?

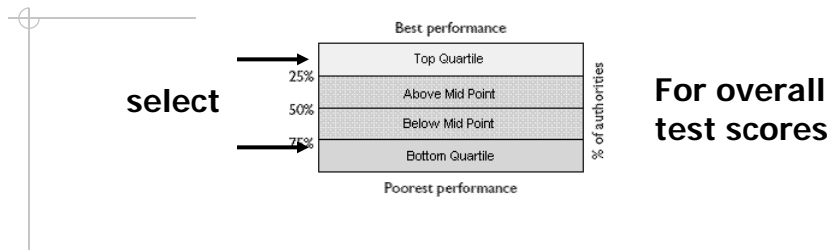
1                      2                      3                      4                      5  
Not at all                      Moderately                      Very much

5. How could I improve exams? Be specific and illustrate improvements please.

Copyright Harvey J. Brightman

30

## Computing index of difficulty and discrimination for essay exam



For each test question, sum the scores for top quartile group (Sum of Scores<sub>high</sub>) and for lowest lowest group (Sum of Scores<sub>low</sub>)

Copyright Harvey J. Brightman

31

## The formulae for index of difficulty and discrimination

### Difficulty

$$\frac{\text{Sum of Scores}_{\text{HIGH}} + \text{Sum of Scores}_{\text{LOW}}}{2} * \text{Max Score for Question}$$

### Discrimination

$$\frac{\text{Sum of Scores}_{\text{HIGH}} - \text{Sum of Scores}_{\text{LOW}}}{4} * \text{Max Score for Question}$$

Copyright Harvey J. Brightman

32

Here is an example for computing index of difficulty

<b>Class of 40. Based on <u>overall test</u> scores:</b>		
<b>10 students in top quartile</b>		
<b>10 students in bottom quartile</b>		
	<b>Q1 = 20 Pts</b>	<b>Q2 = 20 Pts</b>
<b>Sum of Scores<sub>HIGH</sub></b>	<b>180</b>	<b>160</b>
<b>Sum of Scores<sub>LOW</sub></b>	<b>130</b>	<b>90</b>
	$\frac{(180+130)}{(40/2) \times 20} = .78$	$\frac{(160+90)}{(40/2) \times 20} = .63$

Copyright Harvey J. Brightman

33

How to interpret the index of difficulty score

**Norm-Based: Revise or discard question if index less than 0.25 (too hard) or greater than 0.75 (too easy).**

**Criterion-Based: Keep if index greater than 0.75 if critical to learning objectives. If index less than 0.25 revise teaching if critical to learning objectives.**

Copyright Harvey J. Brightman

34

Here is an example for computing index of discrimination

<b>Class of 40. Based on <u>overall test</u> scores:</b>		
<b>10 students in top quartile</b>		
<b>10 students in bottom quartile</b>		
	<b>Q1 = 20 Pts</b>	<b>Q2 = 20 Pts</b>
<b>Sum of Scores<sub>HIGH</sub></b>	<b>180</b>	<b>160</b>
<b>Sum of Scores<sub>LOW</sub></b>	<b>130</b>	<b>90</b>
	$\frac{(180-130)}{(40/4) \times 20} = .25$	$\frac{(160-90)}{(40/4) \times 20} = .35$

Copyright Harvey J. Brightman

35

How to interpret the index of discrimination score

**Norm-Based: Revise or discard question if index less than 0.25 (does not discriminate).**

**Criterion-Based: NA**

Copyright Harvey J. Brightman

36

## Readings list

- ◆ Walvoord, B and Anderson, V. *Effective Grading*.  
Jossey Bass. 1998
- ◆ Stevens D. and Levi, A. *Introduction to Rubrics*.  
Stylus, 2005.
- ◆ Toulmin method:  
<http://writing.colostate.edu/guides/reading/toulmin/>