



# Let's talk about teaching – EXAMS

A structured (open) discussion

May 21, 2019: Steve Hines

## Overall Objectives

### **Break down the silos!**

Through discussion & reflection, explore how our individual philosophies and practices align with our peers – and whether they seem appropriate.

Are there practices more of us should be following?



## Part 1: Objectives

1. Explore the concept of “**desirable difficulty**” as it might apply to exams – and be reflected in exam statistics.
2. Share methods by which we adjust exam scores – and discuss whether we even should.



## Part 2: Objectives

3. Collaboratively explore how & why we share **EXAM STATISTICS** with our students – and the pros & cons of our methods.
4. Explore our shared (or not) philosophies on the **PURPOSE** of exams – e.g. assessment tools versus learning tools.

### Controversy!

5. Share our ideas on whether we should **RETURN** graded exams to our students – and why or why not.



## Discussion Questions



- ➔ 1. What is an appropriate exam mean?
- ➔ 2. Is it appropriate to adjust student scores to hit a targeted mean? If so, when and how?
3. Should we share exam statistics – with our students; with each other?
4. Should graded exams always be returned to students – why or why not? \*\*
5. What is the purpose of an exam?

## Survey: Exam Means

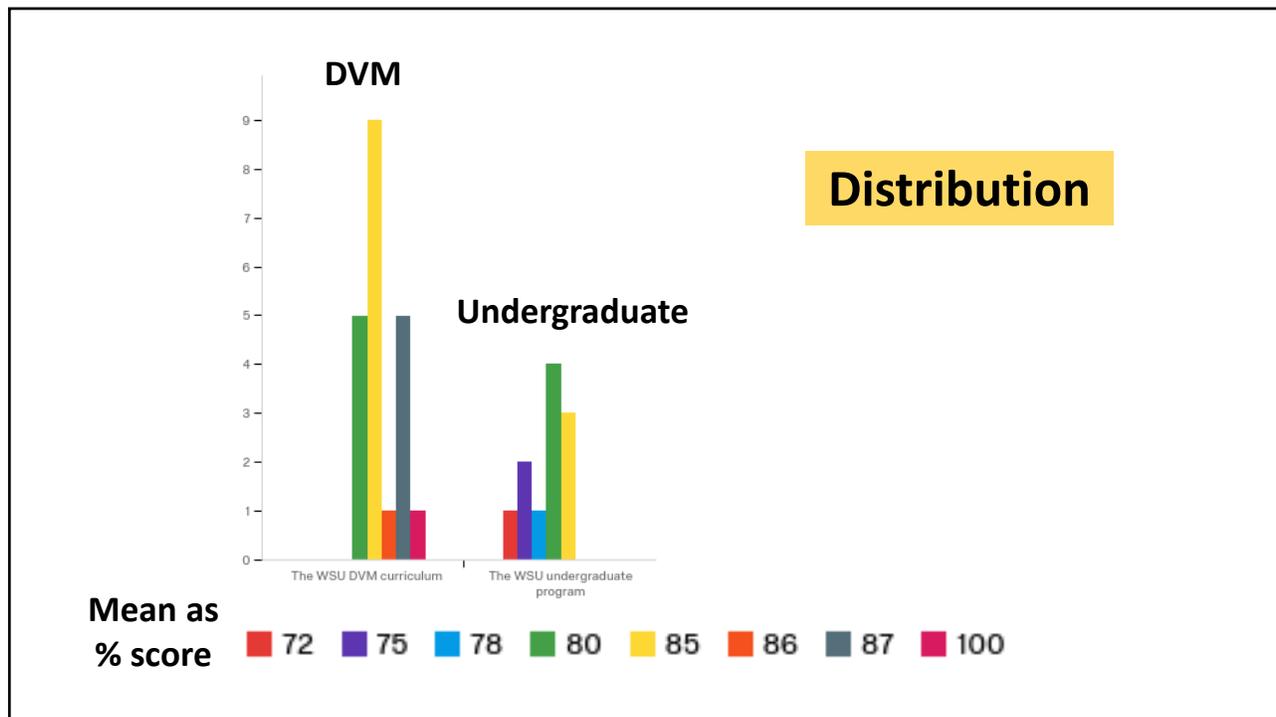
11:05 AM  
05/22/19

### DVM program

- n = 21
- **MEAN = 85.1%**
- SD = 4.2
- Range = 80 – 100 %
- Mean Min. Pass = 72%

### Undergraduate program

- n = 11
- **MEAN = 79.6%**
- SD = 4.16
- Range = 72 – 85%

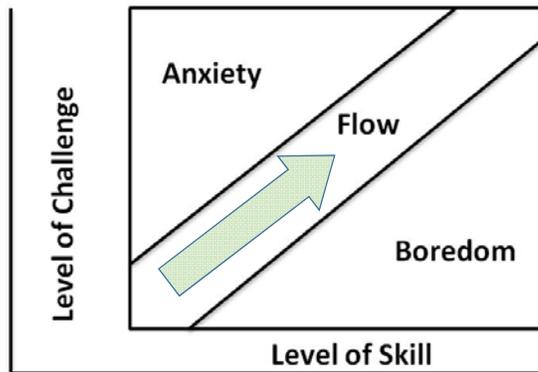


**The term **DESIRABLE DIFFICULTY** was first coined by Robert A. Bjork in 1994. [\[1\]](#)**

- A **desirable difficulty** is a learning task that requires a considerable but desirable amount of effort, thereby improving long-term performance.
- Research suggests that while difficult tasks might slow down learning initially, the long term benefits are greater than with easy tasks. [\[2\]](#)
- However, to be desirable, the tasks must also be accomplishable (*i.e. not so difficult as to be overly frustrating or stress inducing*).



## The “sweet spot”



<https://blogs.scientificamerican.com/observations/how-wrong-should-you-be>

<https://www.biorxiv.org/content/10.1101/255182v1>

### The Eighty Five Percent Rule for Optimal Learning

Robert C. Wilson<sup>a,1</sup>, Amitai Shenhav<sup>b,c</sup>, Mark Straccia<sup>d</sup>, and Jonathan D. Cohen<sup>e</sup>

<sup>a</sup>Department of Psychology and Cognitive Science Program, University of Arizona

<sup>b</sup>Cognitive, Linguistic, & Psychological Sciences, Brown University

<sup>c</sup>Brown Institute for Brain Science, Brown University

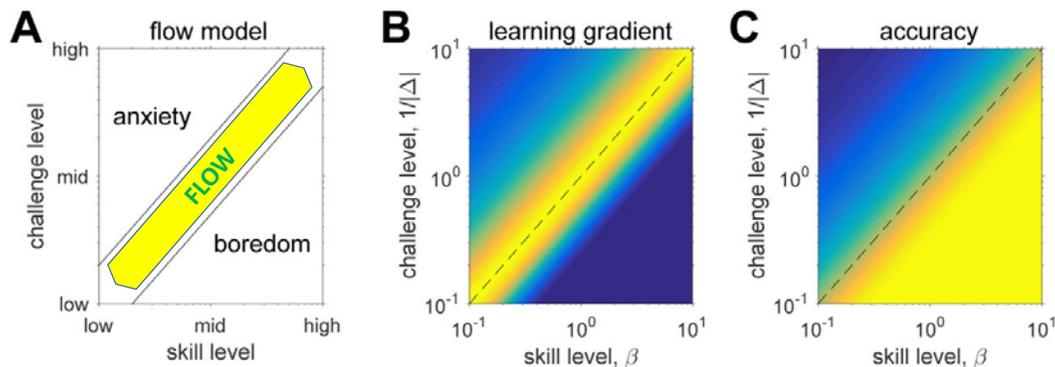
<sup>d</sup>Department of Psychology, UCLA

<sup>e</sup>Princeton Neuroscience Institute, Princeton University

<sup>1</sup>To whom correspondence should be addressed. E-mail: bob@arizona.edu

## Abstract

- Researchers and educators have long wrestled with the question of how best to teach their clients be they human, animal or machine. **Here we focus on the role of a single variable, the difficulty of training, and examine its effect on the rate of learning.**
- In many situations we find that there is **a sweet spot** in which training is neither too easy nor too hard, and where learning progresses most quickly. We derive conditions for this sweet spot for a broad class of learning algorithms in the context of binary classification tasks, in which ambiguous stimuli must be sorted into one of two classes.
- For all of these gradient-descent based learning algorithms **we find that the optimal error rate for training is around 15.87% or, conversely, that the optimal training accuracy is about 85%.**
- We demonstrate the efficacy of this ‘Eighty Five Percent Rule’ for artificial neural networks used in AI and biologically plausible neural networks thought to describe human and animal learning.



**Figure 4:** Proposed relationship between the Eighty Five Percent Rule and Flow.

(A) Original model of flow as a state that is achieved when skill and challenge are well balanced.

(B) Normalized learning rate,  $\partial ER/\partial \beta$ , (B) and accuracy (C) as a function of skill and challenge suggests that **flow** corresponds to high learning and accuracy, **boredom** corresponds to low learning and high accuracy, while **anxiety** is associated with low learning and low accuracy.

## Discussion Questions



1. What is an appropriate exam mean?
- ➔ **2. Is it appropriate to adjust student scores to hit a targeted mean? If so, when and how?**
3. Should we share exam statistics – with our students; with each other?
4. Should graded exams always be returned to students – why or why not? \*\*
5. What is the purpose of an exam?



### Let's talk about teaching – EXAMS

A structured (open) discussion

May 21, 2019: Steve Hines



## Part 2: Objectives

- ➔ **3. Collaboratively explore how & why we share EXAM STATISTICS with our students – and the pros & cons of our methods.**
4. Explore our shared (or not) philosophies on the PURPOSE of exams – e.g. assessment tools versus learning tools.

### Controversy!

5. Share our ideas on whether we should RETURN graded exams to our students – and why or why not.

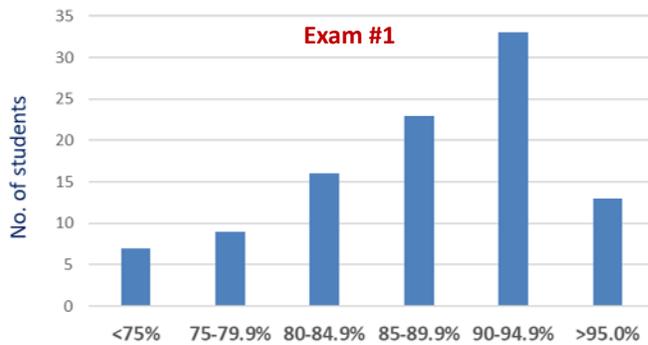


## Collaboratively explore how & why we share EXAM STATISTICS with our students – and the pros & cons of our methods.

- How many share exam statistics with students?
- Why?
- Your methods? (e.g. statistics shared)

Exam Statistics	
Exam Summary	Question Details
Number of Students	13
Maximum Score Possible	100 pts (100%)
High Score	89 (89%)
Median Score	80 (80%)
Mean Score	78.38 (78%)
Low Score	45 (45%)
Top Quartile Score (Q4)	87.5 - 89 (88- 89%)
Quartile 3 Scores (Q3)	80 - 87.5 (80- 88%)
Quartile 2 Scores (Q2)	75 - 80 (75- 80%)
Bottom Quartile Scores (Q1)	45 - 75 (45- 75%)

## Fall, Year 2 – DVM curriculum



- Scores are put into pools to discourage meaningless comparisons
- Additional statistics are NOT provided

### Concept: Students can -

- gain some sense of overall exam difficulty – as reflected in class performance
- measure their own performance relative to their peers
- NOT define themselves as “above or below” average

## Part 2: Objectives

3. Collaboratively explore how & why we share EXAM STATISTICS with our students – and the pros & cons of our methods.



4. **Explore our shared (or not) philosophies on the PURPOSE of exams – e.g. assessment tools versus learning tools.**

### Controversy!

5. Share our ideas on whether we should RETURN graded exams to our students – and why or why not.



