Scott E Page

University of Michigan

Santa Fe Institute

Leveraging Diversity

Outline

The Path of Inclusion

Identity and Cognitive Diversity

Prediction

Problems Solving

Case: The Netflix Prize

Takeaways

Framework

Tools: Diverse Perspectives, Heuristics, and Interpretations

Tasks: Problem Solving and Prediction

The Path of Inclusion

Hiring diverse people is the **right** thing to do.

Hiring diverse people is the **required by law.**

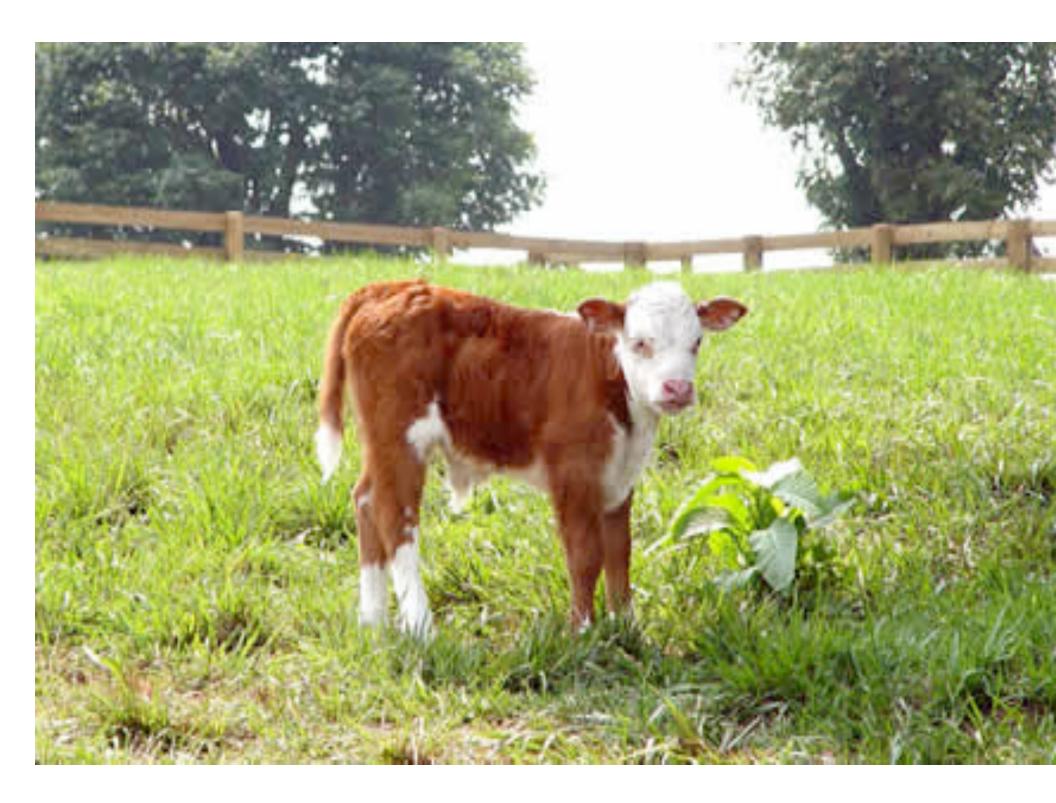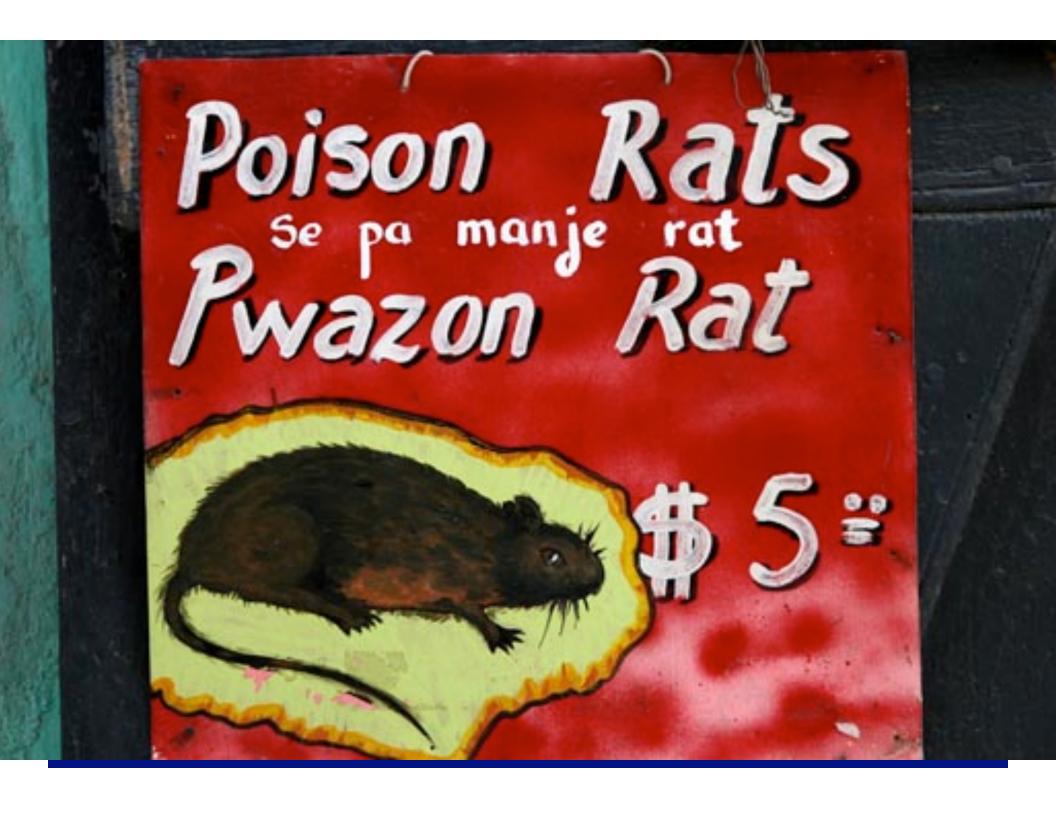Seeking diversity enlarges the pool and results in better employees.

Diversity

Ability

Diversity is a strategic advantage. It makes organizations more productive and more innovative on cognitive tasks.
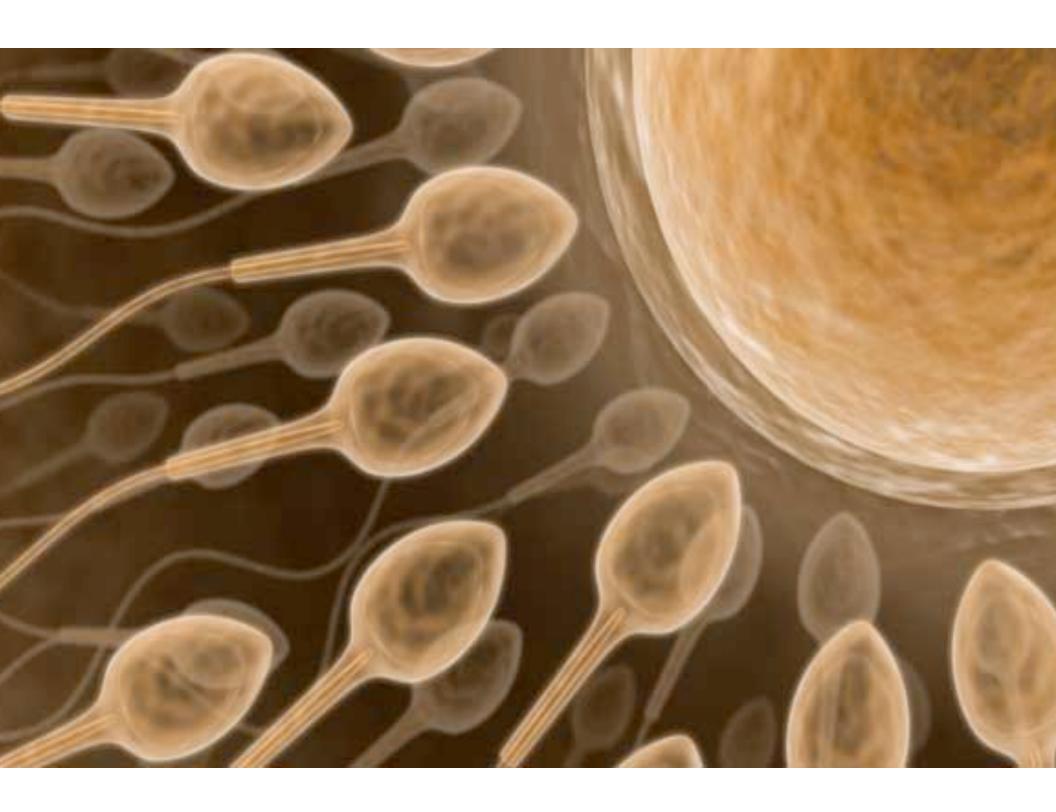
Diversity
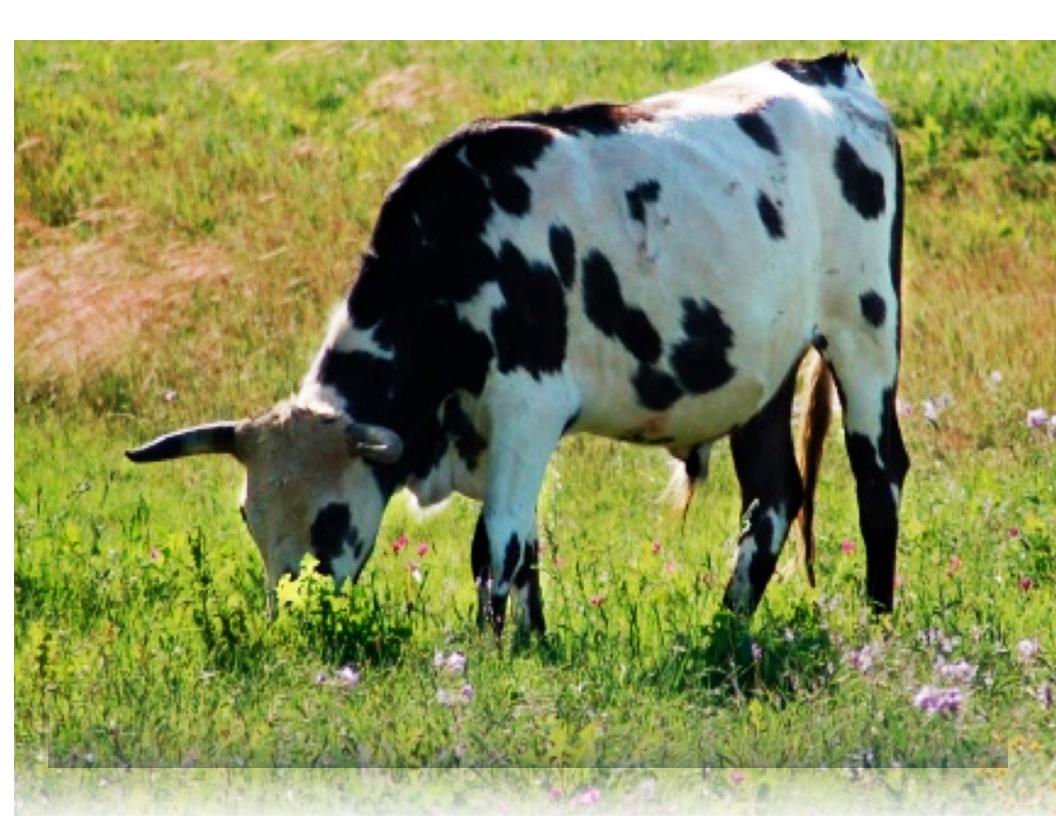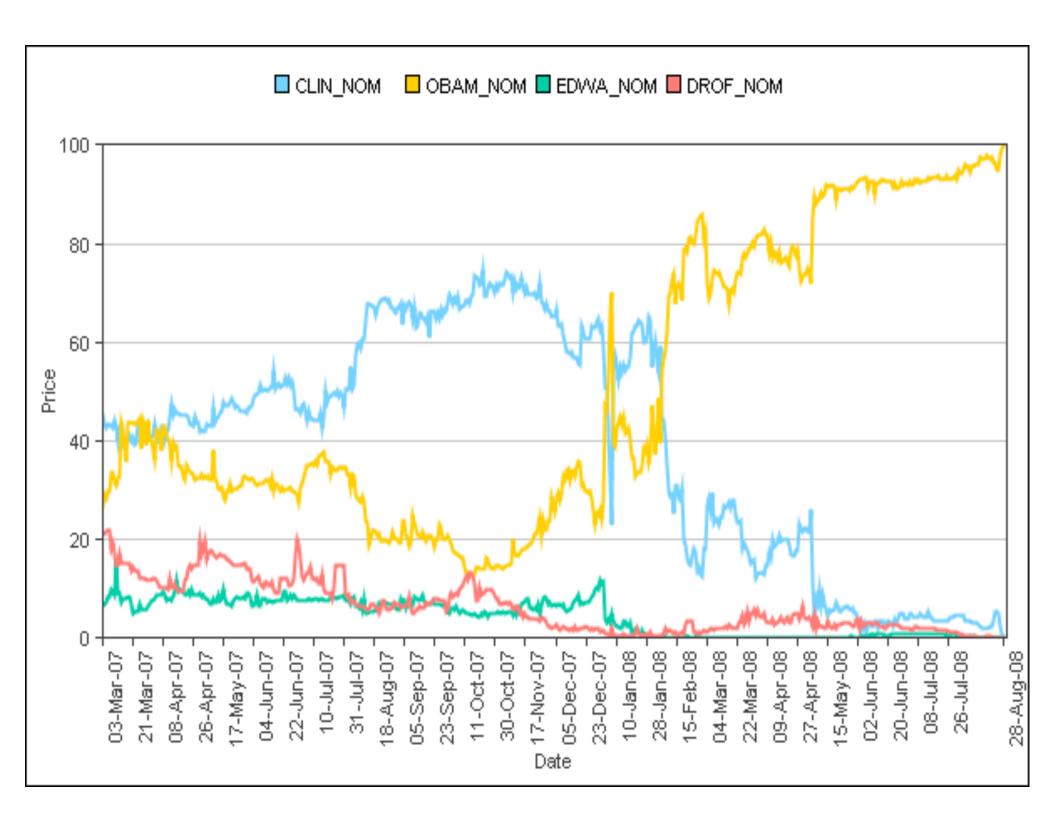# Ability

Identity and Cognitive Diversity

Prediction

Iowa Electronic Markets

IEM Prices
Obama 0.535
McCain 0.464

Final Gallup Poll
Obama 0.55
McCain 0.44

Actual Outcome
Obama 0.531
McCain 0.469

Methods of Divination

Stars and Planets (astrology)

Rolling Dice

Tarot Cards

Palm Reading

Crystal Balls

Head Shape (Phrenology)

Atmospheric Conditions

Dreams

Animal Entrails

Moles on the body

Lightning

Smoke and Fire

Flight of Birds

Neighing of Horses
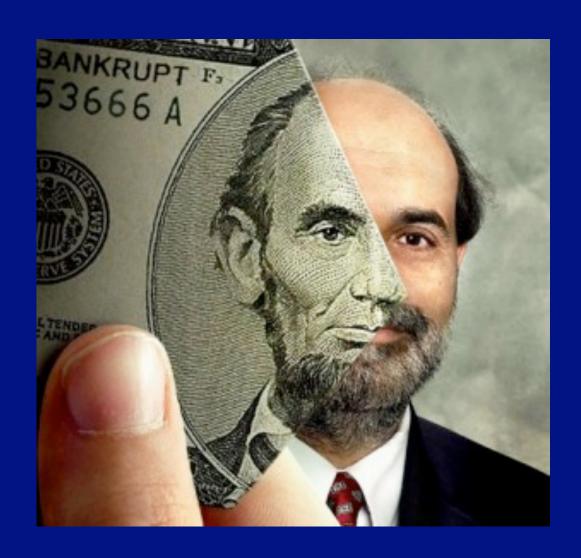
Tea Leaves and Coffee Grounds

Passages of Sacred Texts

Numbers

I Ching

Guessing

MODELS
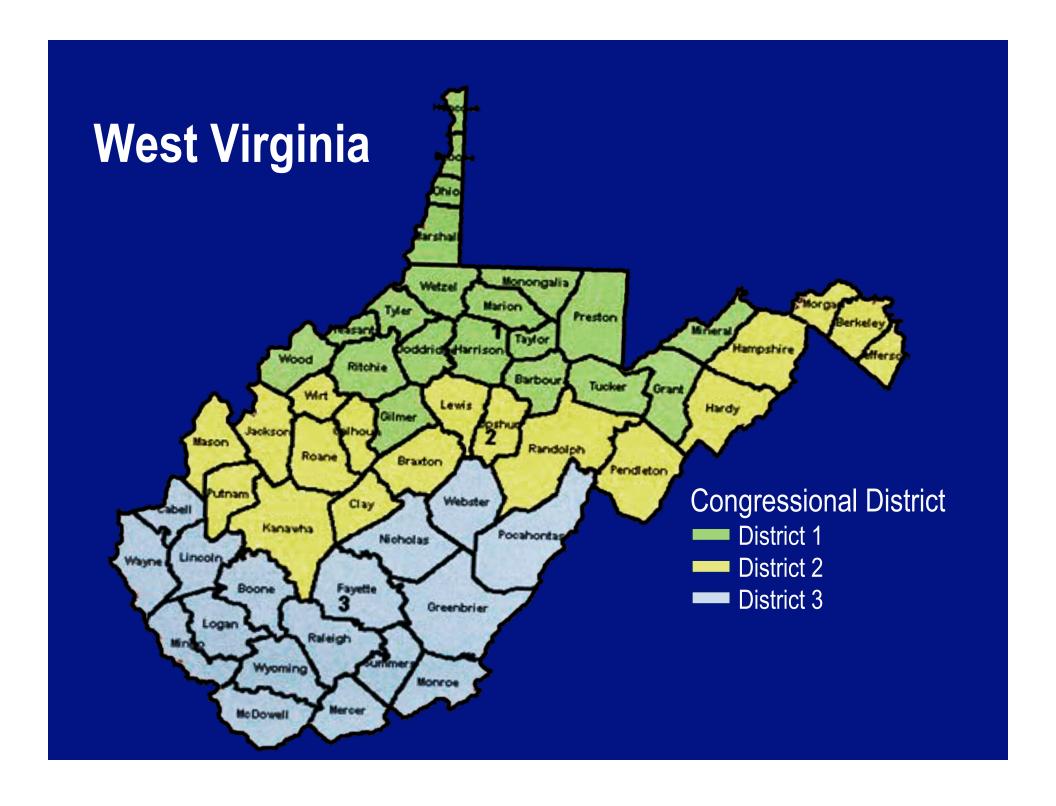
David Orrell "The Future of Everything.

West Virginia

Slaw available on request
Slaw standard on hot dogs
Slaw not available
No data available
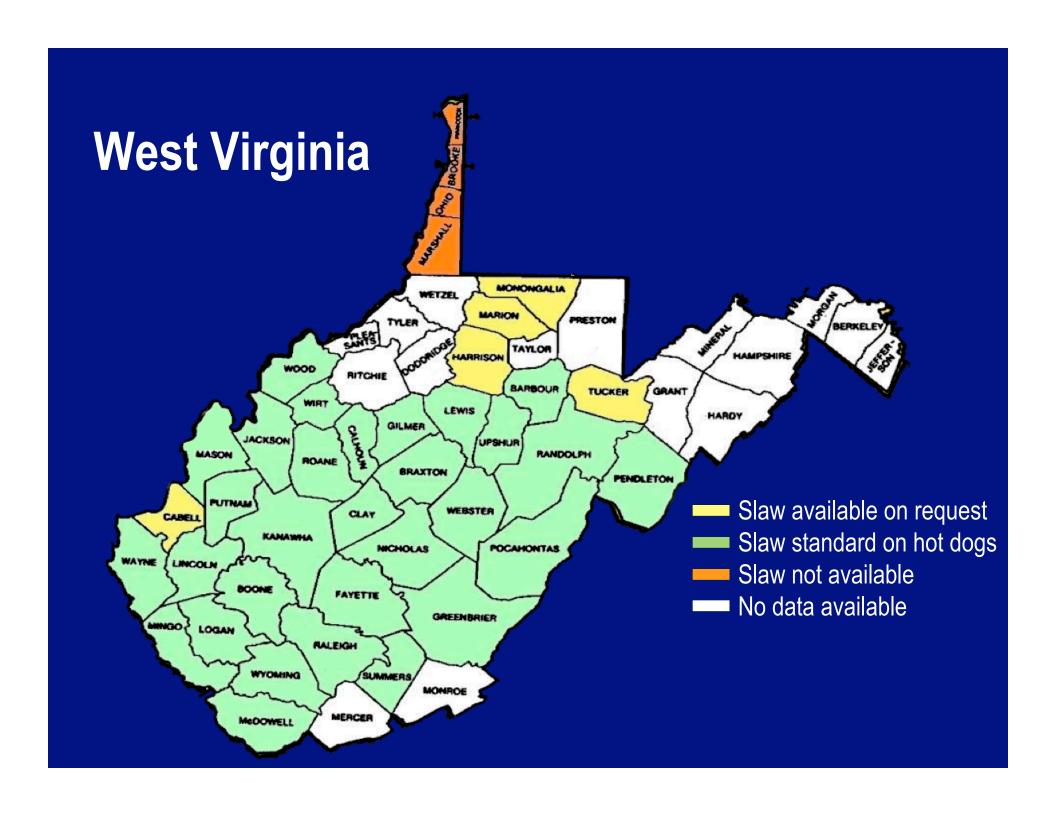
Interpretations: Pile Sort

Place the following food items in piles

Broccoli	Carrots	Canned Beets
Fresh Salmon	Arugula	Fennel
Spam	Ahi Tuna	Canned Posole
Niman Pork	Sea Bass	Canned Salmon

BOBO Sort

Veggie	Organic	Canned
Broccoli	Fresh Salmon	Canned Beets
Arugula	Sea Bass	Spam
Carrots	Niman Pork	Canned Salmon
Fennel	Ahi Tuna	Canned Posole

Airstream Sort

Veggie	**Meat/Fish**	**Weird?**
Broccoli	Fresh Salmon	Canned Posole
Fennel	Spam	Sea Bass
Carrots	Niman Pork	Arugula
Canned Beets	Canned Salmon	Ahi Tuna

Diversity Prediction Theorem

Crowd Error = Average Error - Diversity

$$(c - \theta)^2 = \frac{1}{n}\sum_{i=1}^{n}(s_i - \theta)^2 - \frac{1}{n}\sum_{i=1}^{n}(s_i - c)^2$$

Galton's Steer

Crowd Error = Average Error – Diversity

0.6 = 2,956.0 - 2955.4

2005 NFL Draft

Player	A	B	C	D	E	F	G	CROWD
Alex Smith	1	1	1	1	1	1	1	**1**
Ronnie Brown	2	2	4	2	2	5	2	**2.7**
Braylon Edwards	3	3	2	7	3	2	3	**3.3**
Cedric Benson	4	4	13	4	8	4	8	**5.9**
Carnell Williams	8	5	5	5	4	13	4	**6.4**
Adam Jones	16	9	6	8	6	6	9	**8.1**

2005 NFL Draft

Predictor	A	B	C	D	E	F	G	CROWD
Squared Error	158	89	210	235	112	82	75	**34.4**

NFL Experts

Predictor	A	B	C	D	E	F	G	CROWD
Squared Error	158	89	210	235	112	82	75	**34.4**

Average Error: 137.3
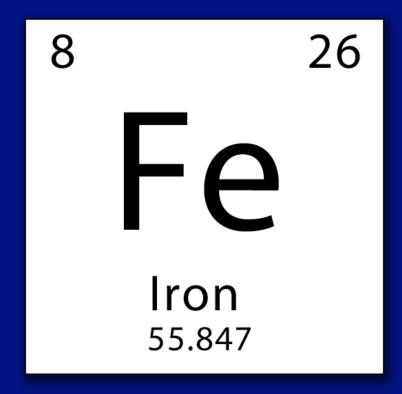Diversity: 102.9
Crowd Error: **34.4**

Problem Solving

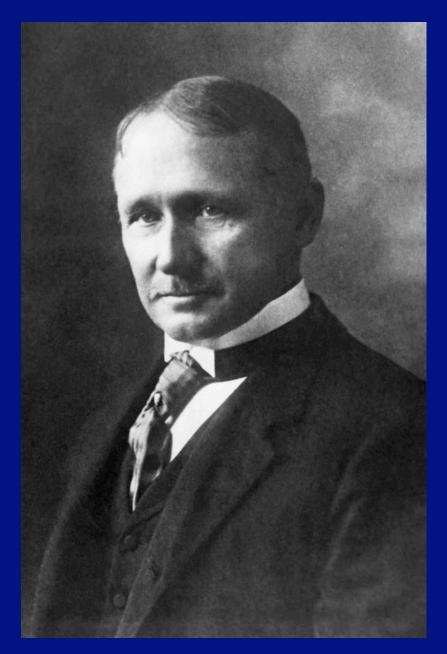Gunter Blobel: The exception
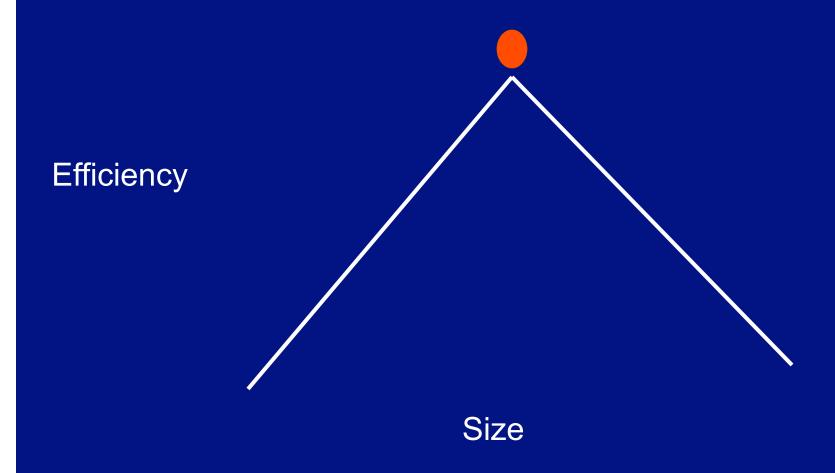
Perspectives

The Technocratic Ideal

Frederick

Winslow

Taylor

1856-1915
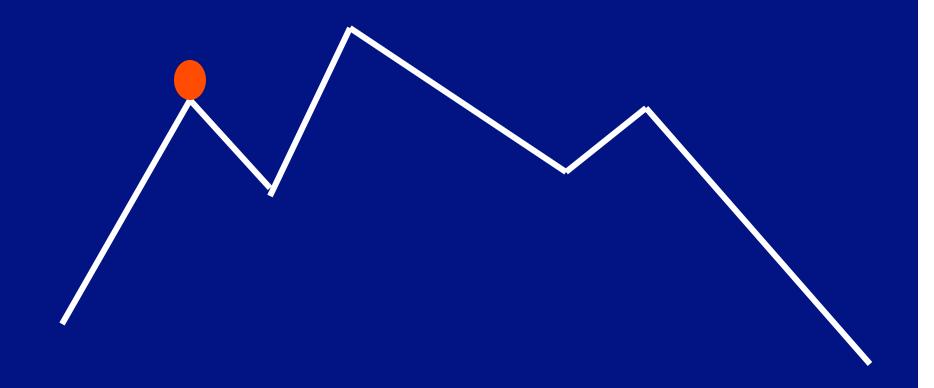
Simple: Shovel Landscape

Efficiency

Size
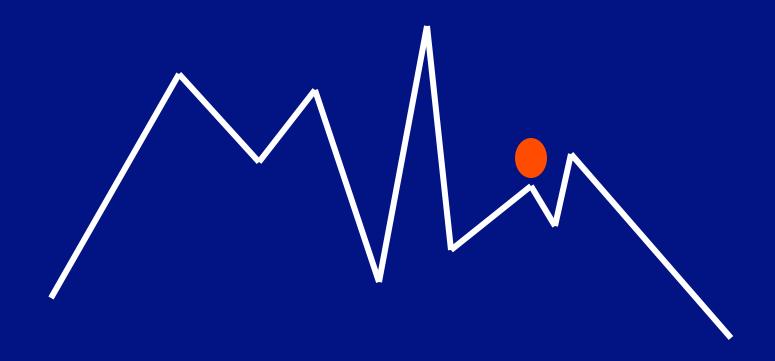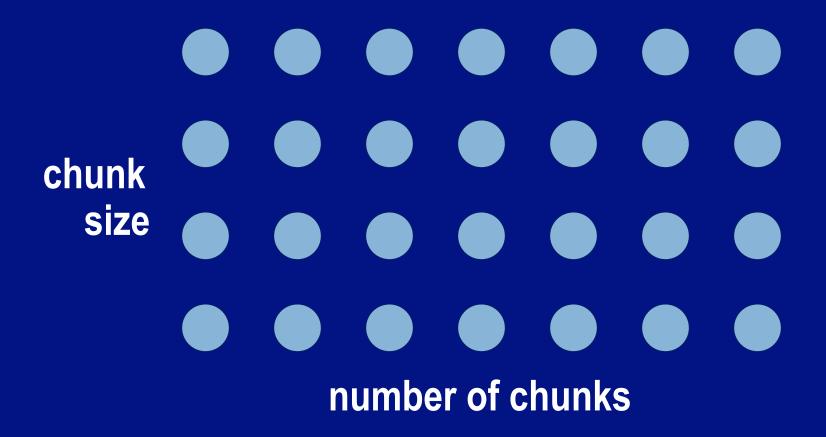
Caloric Landscape

Masticity Landscape

Consultant's Perspective

caloric rank

Ben and Jerry's Perspective

Consultant's Perspective

caloric rank
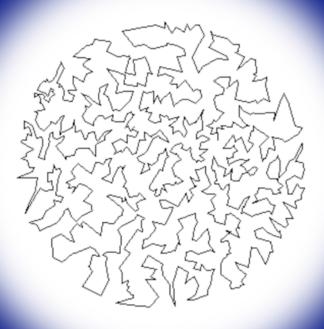
Heuristics

IQ Question:

Fill in the Blank: 1   2   3   5   _   13

$$1 \quad 2 \quad 3 \quad 5 \quad \mathbf{8} \quad 13$$

$$x_{i+2} - x_{i+1} = x$$

IQ Question:

1   4   9   16   _   36

1   4   9   16   **25**   36

$$x_i^2$$

IQ Question:

1   2   6   _   1806

$$1 \quad 2 \quad 6 \quad \mathbf{42} \quad 1806$$

$$x_{i+1} - x_i = x_i{}^2$$

$2 - 1 = 1^2$

$6 - 2 = 2^2$

$42 - 6 = 6^2$
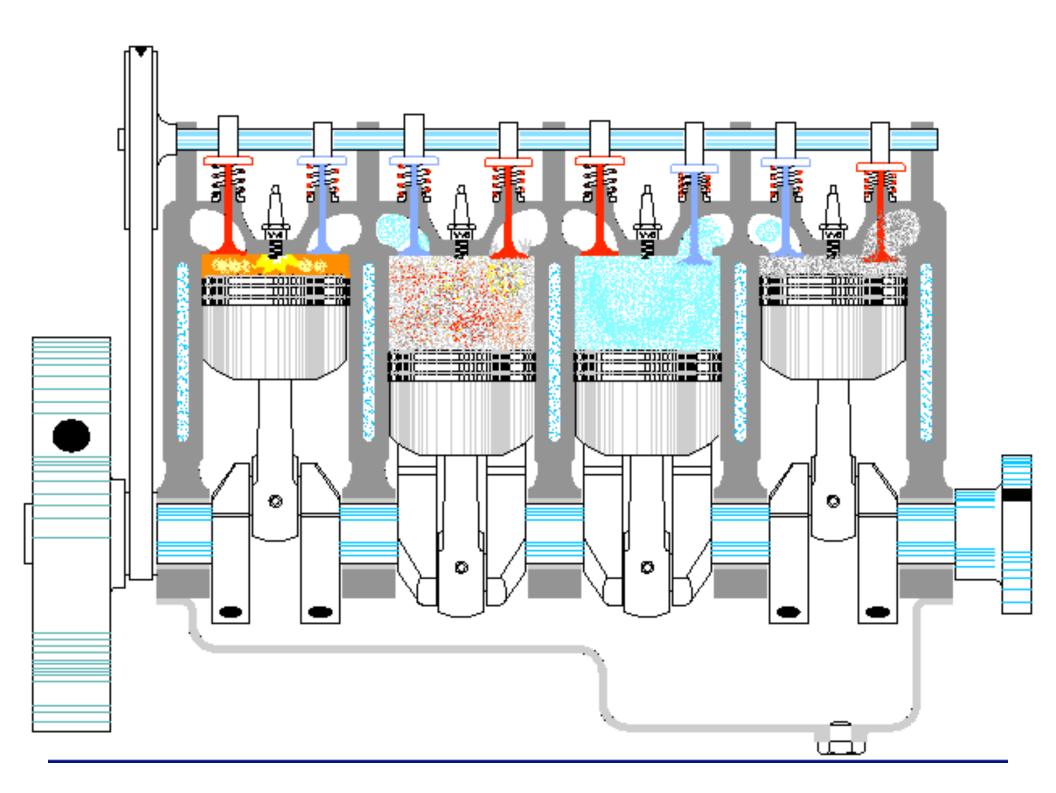
$1806 - 42 = 42^2$

$$x_{i+1} - x_i = x_i{}^2$$

A combination of the first two heuristics
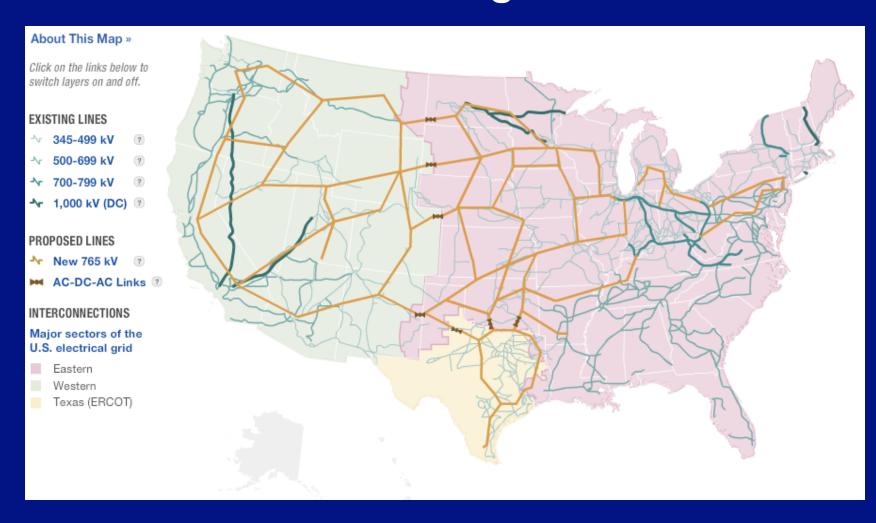
Superadditivity

$$1 + 1 = 3$$

Network + Electrical Engineers

A Test

- Create a bunch of agents with diverse **perspectives** and **heuristics**

- Rank them by their performance on a problem.

- Note: all of the agents must be "smart"
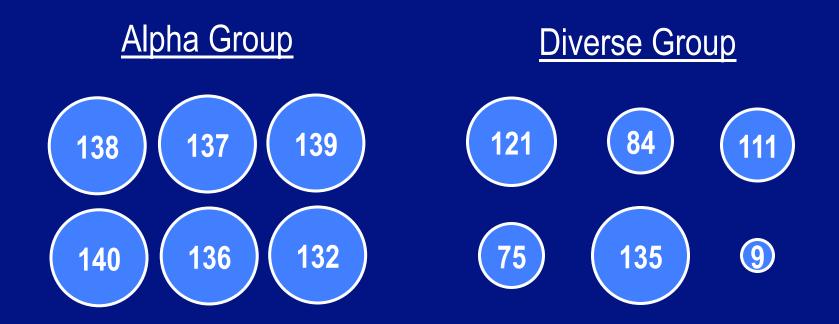
Experiment

Group 1: Best 20 agents

Group 2: Random 20 agents

Have each group work collectively - when one agent gets stuck at a point, another agent tries to find a further improvement. Group stops when no one can find a better solution.

The IQ View

Alpha Group

138 137 139

140 136 132

Diverse Group

121 84 111

75 135 9

The diverse group almost always outperforms the group of the best by a substantial margin.

See Lu Hong and Scott Page

Proceedings of the National Academy of Sciences (2002)

What Must be True?

Calculus Condition: *Problem solvers must all be smart--we must be able to list their local optima*

Diversity Condition: *Problem solvers must have diverse heuristics and perspectives*

Hard Problem Condition: *Problem itself must be difficult*

Case: Netflix Prize

Outline

Netflix Prize: Background

Predictive Models

   Factor Models

Ensembles of Models

Ensembles of Teams

The Value of Diversity

Netflix Prize

November 2006, Netflix offers a prize of $1 million to anyone who can defeat their Cinematch recommender system by 10% of more.

Some Details

Netflix users rank movies from 1 to 5

Six years of data

Half million users

17,700 movies

Data divided into (training, testing)

Testing Data dived into (probe, quiz, test)

Interesting Asides

Lost in Translation and *The Royal Tenenbaums* had the highest variance

Shawshank Redemption had the highest rating

Miss Congeniality had the most ratings.

Singular Value Decomposition

Each movie represented by a vector:

$$(p_1, p_2, p_3, p_{4}...p_n)$$

Each person represented by a vector:

$$(q_1, q_2, q_3, q_{4}...q_n)$$

Rating: $r_{ij} = m_i + a_j + p \bullet q$

Training: choose p,q to minimiize

$$(actual_{ij} - r_{ij})^2 + c(||p||^2 + ||q||^2)$$

BellKor's Initial Models

Approximately 50 dimensions

Best Model: 6.8% improvement

Combination of Models: 8.4% improvement

Two Questions

Q1: Why more than one model?

Q2: Why do more work better than one?
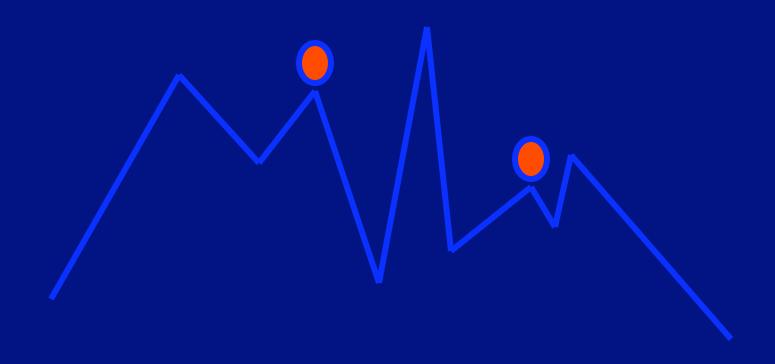
Q1: Why More than one Model

This question has two answers.

A1: they used different variables

A2: their stochastic optimization technique got stuck in different places

Different Tuning Parameters and Initial Points Lead to Different Peaks on a Rugged Landscape

A2: Diversity Prediction Theorem

SqE(**c**) = SqE(**s**) - PDiv(**s**)

$$(c - \theta)^2 = \frac{1}{n}\sum_{i=1}^{n}(s_i - \theta)^2 - \frac{1}{n}\sum_{i=1}^{n}(s_i - c)^2$$

BellKor's Pragmatic Chaos

More is Better: Seven person team created combining top two teams Now over 800 predictor sets (sets of variables).

Difficult be build a "grand" model but possible to build lots of "huge" models

Ensemble Effects

Best Model 8.4%

Ensemble: 10.1%

<u>Rules:</u> Once someone breaks 10%, then the contest ends in 30 days.

Enter ``The Ensemble''

23 teams from 30 countries who blended their predictive models who tried in the last moments to defeat BellKor's Pragamatic Chaos

The Ensemble

"The contest was almost a race to agglomerate as many teams as possible," said David Weiss, a Ph.D. candidate in computer science at the University of Pennsylvania and a member of the Ensemble.

"The surprise was that the collaborative approach works so well, that trying all the algorithms, coding them up and putting them together far exceeded our expectations."

New York Times 6/27/09

And The Winner is…

RMSE for The Ensemble: 0.856714

RMSE for Bellkor's Pragmatic Chaos: 0.856704

By the rules of the competition the scores are rounded to four decimal places so it was a tie.

However, BellKor's Pragmatic Chaos submitted 20 minutes earlier so they won. (and they had the lower error)

Oh, by the way..

BellKor's Pragmatic Chaos	10.06%	
The Ensemble		10.06%
50/50 Blend		10.19%

Takeaways

1. Value of Diversity Depends on Extent of Collaboration.

Holedigging

Boosting

Collective Problem Solving

2. Create Oracles

3. Create Perspectives/Skills Spreadsheets

name	engineer	sales	physics	statistics
A	x		x	
B		x		x
C		x		

4. Listen to Others But Avoid Group Think

Learning

Average individual squared error of seven experts who made forecasts about the NBA draft from May 23rd through June 25th.

May 23rd :	213.17
May 30th :	86.33
June 13th:	114.5
June 18th :	139.67
June 22nd :	109
June 25th:	69.67

Avoiding Group Think

Date	Individual	Diversity	Collective Error
May 23rd :	213.17	168.03	45.14
May 30th :	86.33	81.41	28.57
June 13th:	114.5	70.31	44.19
June 18th :	139.67	113.3	26.34
June 22nd :	109.0	84.0	25.0
June 25th:	69.67	35.58	33.58

Avoiding Group Think
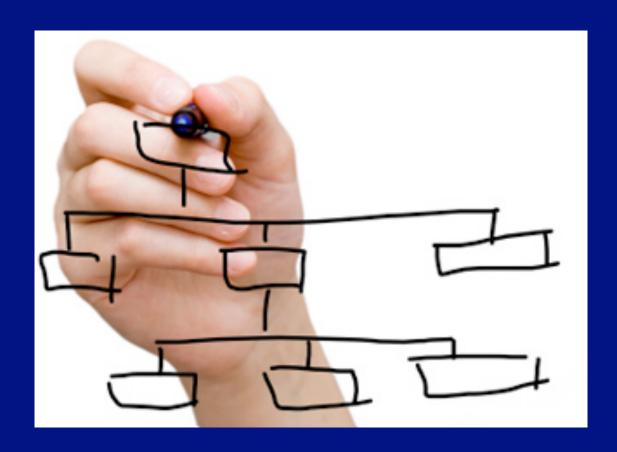
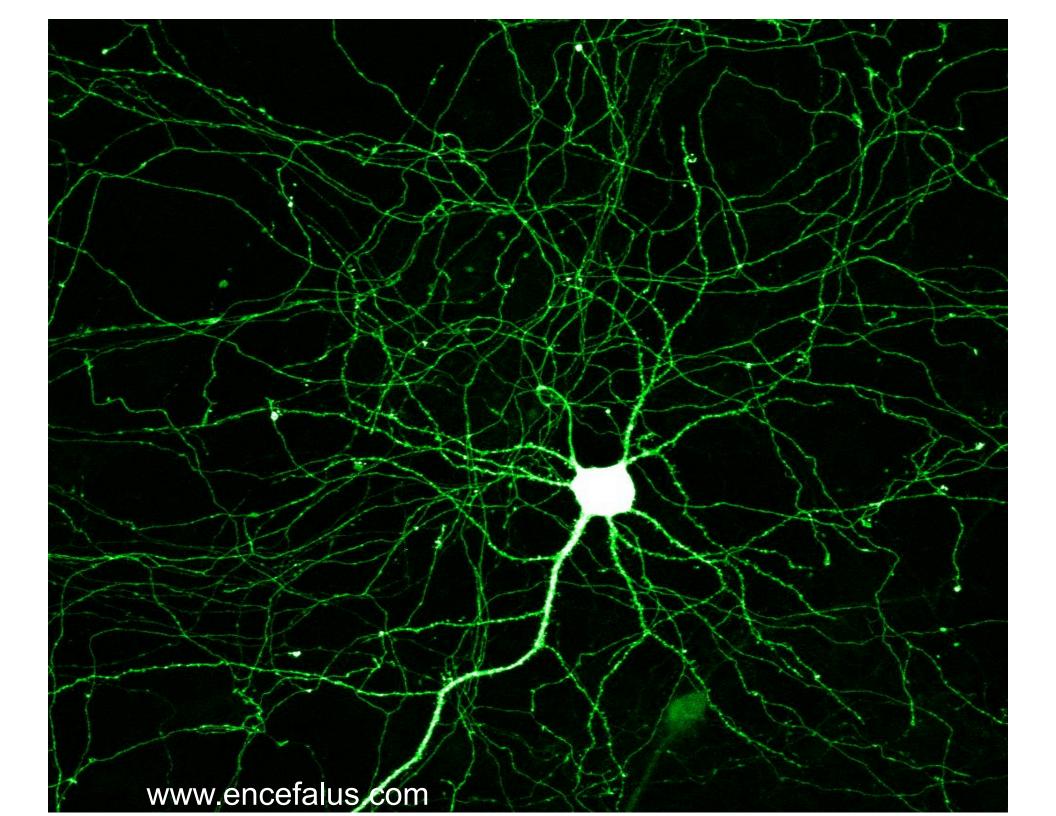| Date	Individual	Diversity	Collective Error
May 23rd :	213.17	168.03	45.14
May 30th :	86.33	81.41	28.57
June 13th:	114.5	70.31	44.19
June 18th :	139.67	113.3	26.34
June 22nd :	**109.0**	**84.0**	**25.0**
June 25th:	**69.67**	**35.58**	**33.58**

Encourage Dissent

If everyone agrees, then either the predictive task was easy and everyone has the correct forecast (in which case the meeting was a waste of time) or the the task was challenging and everyone has the same, wrong forecast.
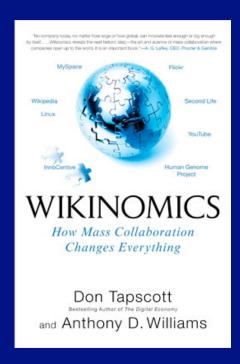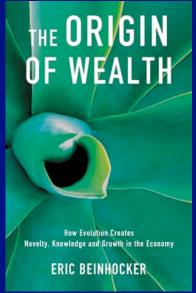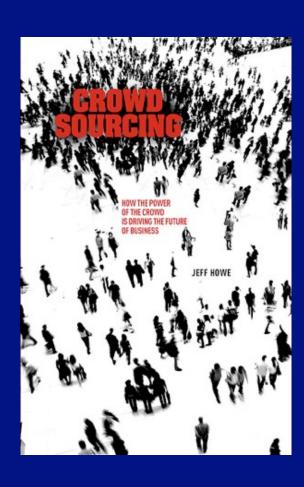
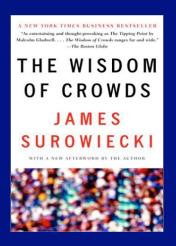# 5. Technology Can Supplement Hierarchy

www.healys.eu

"No company today, no matter how large or how global, can innovate fast enough or big enough by itself. . . . Wikinomics reveals the next historic step—the art and science of mass collaboration where companies open up to the world. It is an important book."—A. G. Lafley, CEO, Procter & Gamble

MySpace   Flickr

Wikipedia   Second Life

Linux

YouTube

InnoCentive   Human Genome Project

WIKINOMICS

How Mass Collaboration Changes Everything

Don Tapscott
Bestselling Author of *The Digital Economy*
and Anthony D. Williams

CROWD SOURCING

HOW THE POWER OF THE CROWD IS DRIVING THE FUTURE OF BUSINESS

JEFF HOWE

A *NEW YORK TIMES* BUSINESS BESTSELLER
"As entertaining and thought-provoking as *The Tipping Point* by Malcolm Gladwell. . . . *The Wisdom of Crowds* ranges far and wide."
—*The Boston Globe*

THE WISDOM OF CROWDS

JAMES SUROWIECKI

WITH A NEW AFTERWORD BY THE AUTHOR

THE ORIGIN OF WEALTH

How Evolution Creates
Novelty, Knowledge and Growth in the Economy

ERIC BEINHOCKER

SMART MOBS

THE • NEXT • SOCIAL • REVOLUTION

TRANSFORMING CULTURES AND
COMMUNITIES IN THE AGE
OF INSTANT ACCESS
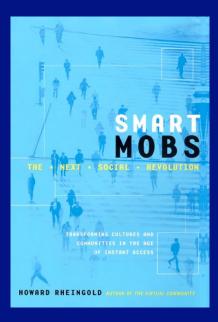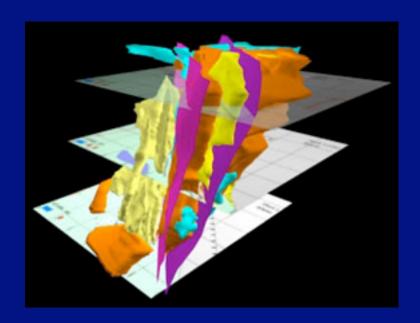
HOWARD RHEINGOLD AUTHOR OF *THE VIRTUAL COMMUNITY*

Goldcorp Challenge

March 6, 2000, Goldcorp offers $575k to participants who would help find gold at its Red Lake Mine in Ontario, Canada

110 targets identified, over 50% were new, over 80% were successful. Company value up from $100 Million to $9 Billion.

Prediction Markets
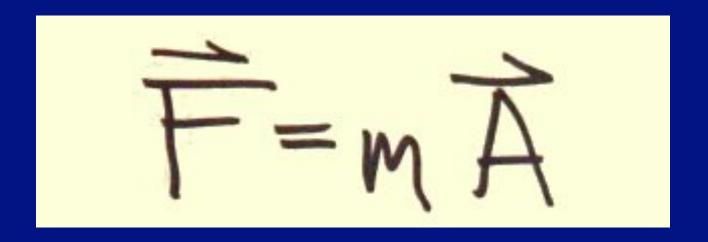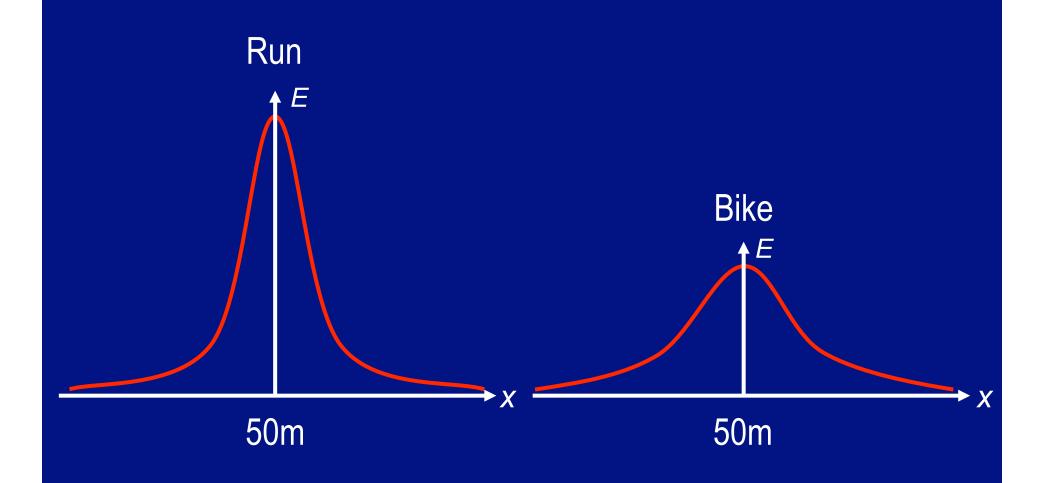
The Math Tells What's Possible

$$\vec{F} = m\vec{A}$$

The Need for Leadership

homogeneous

Cognitively diverse