High Level Strategic Approaches for Conducting Big Data Studies in Assessment Studies

AUSTIN FOSSEY, MA

LIBERTY MUNSON, PhD

MANNY STRAEHLE, PhD, GISF

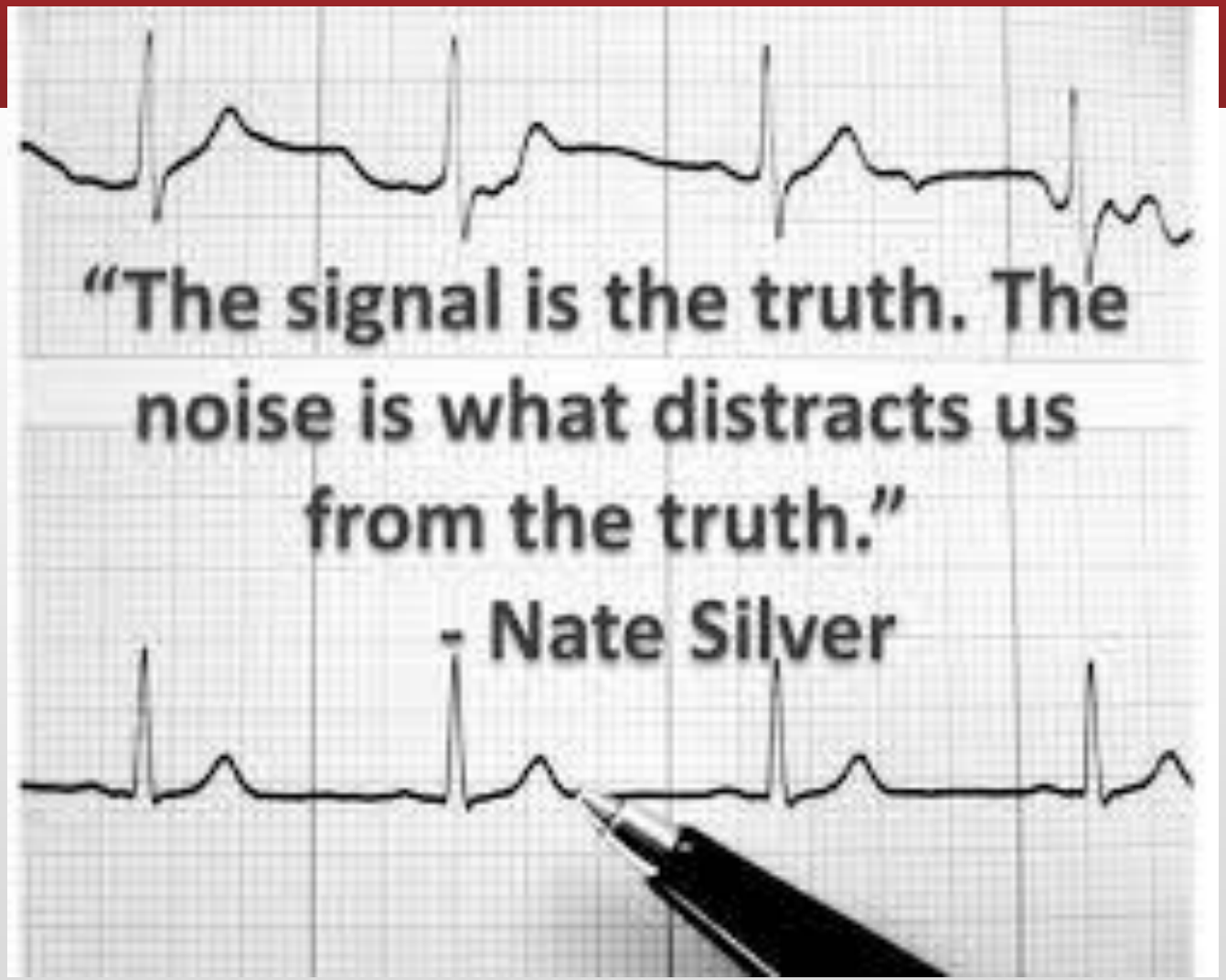
AERE
Assessment, Education & Research Experts

Agenda

BIG DATA 101
- Definition
- Mainstream Examples
- Five Strategic Approaches

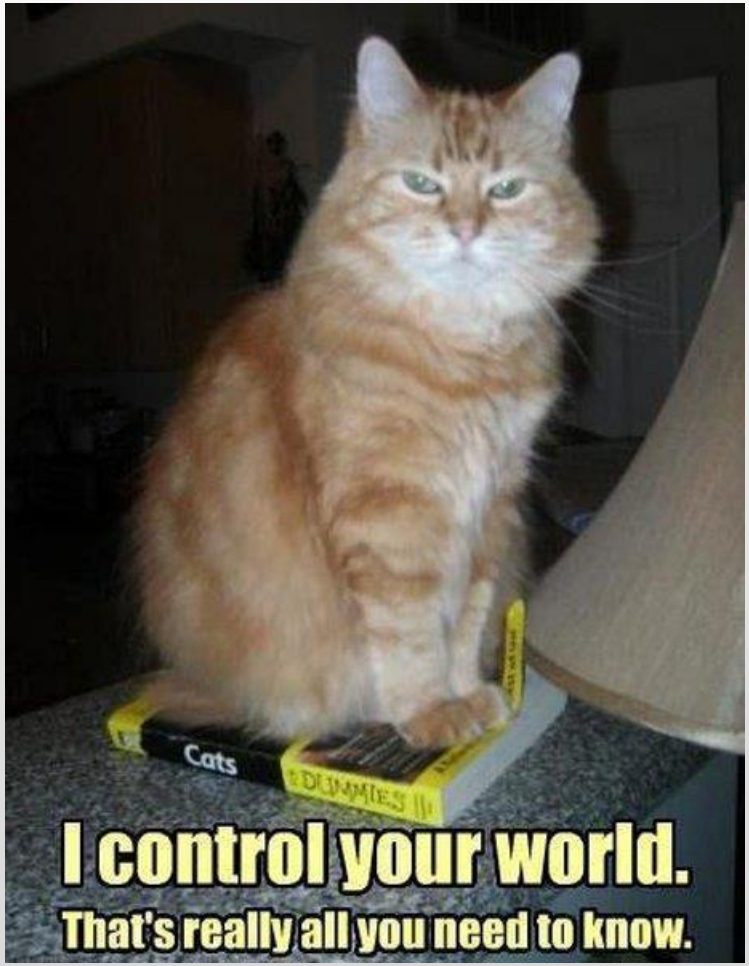
Validation in Big Data
- Using Clustering Algorithms
- Case Study

How to Use Big Data to See the Psychometric World Differently
- Online courses
- Exam Performance
- Forensics

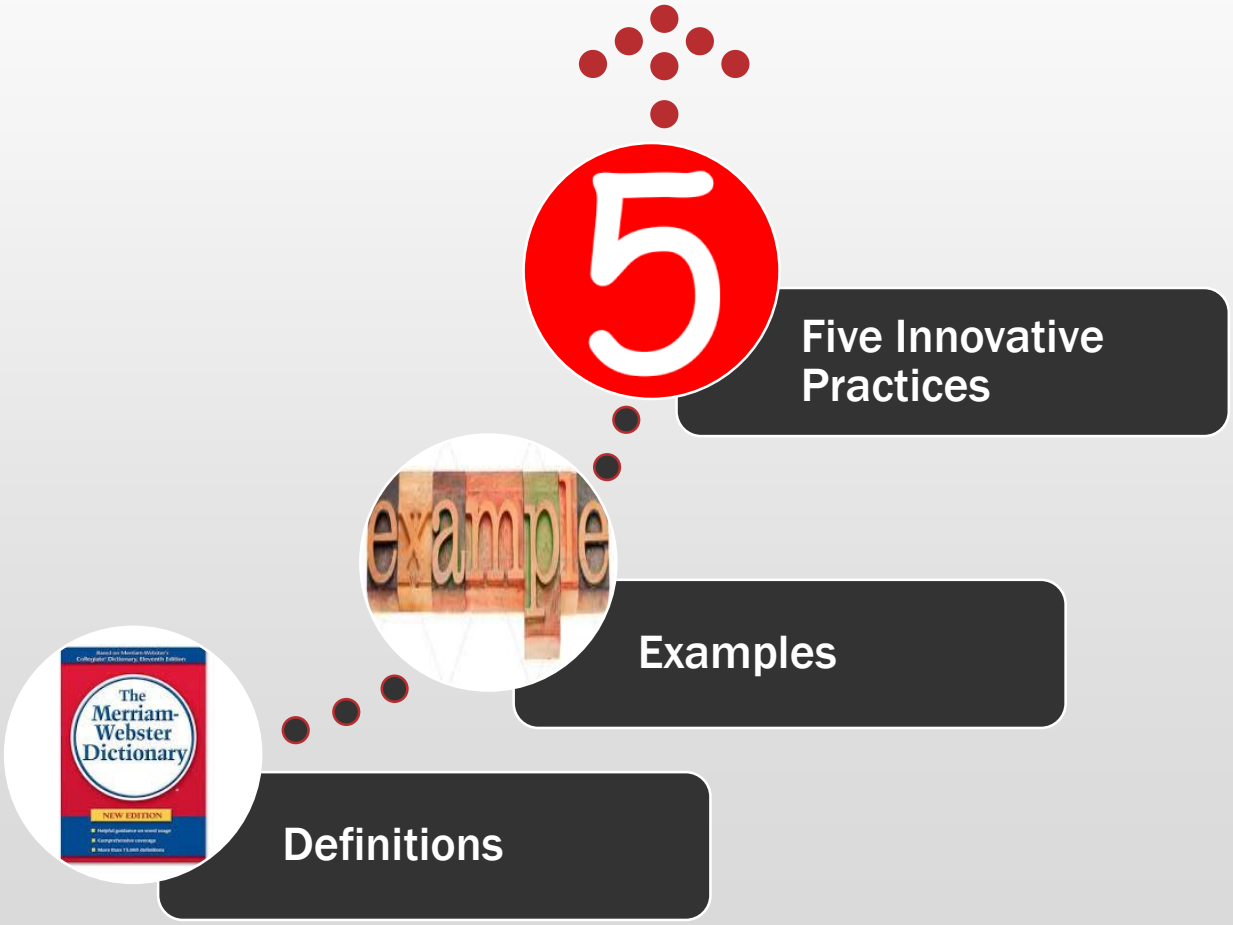

"The signal is the truth. The noise is what distracts us from the truth."
- Nate Silver

The Noise

The Signal



The Journey

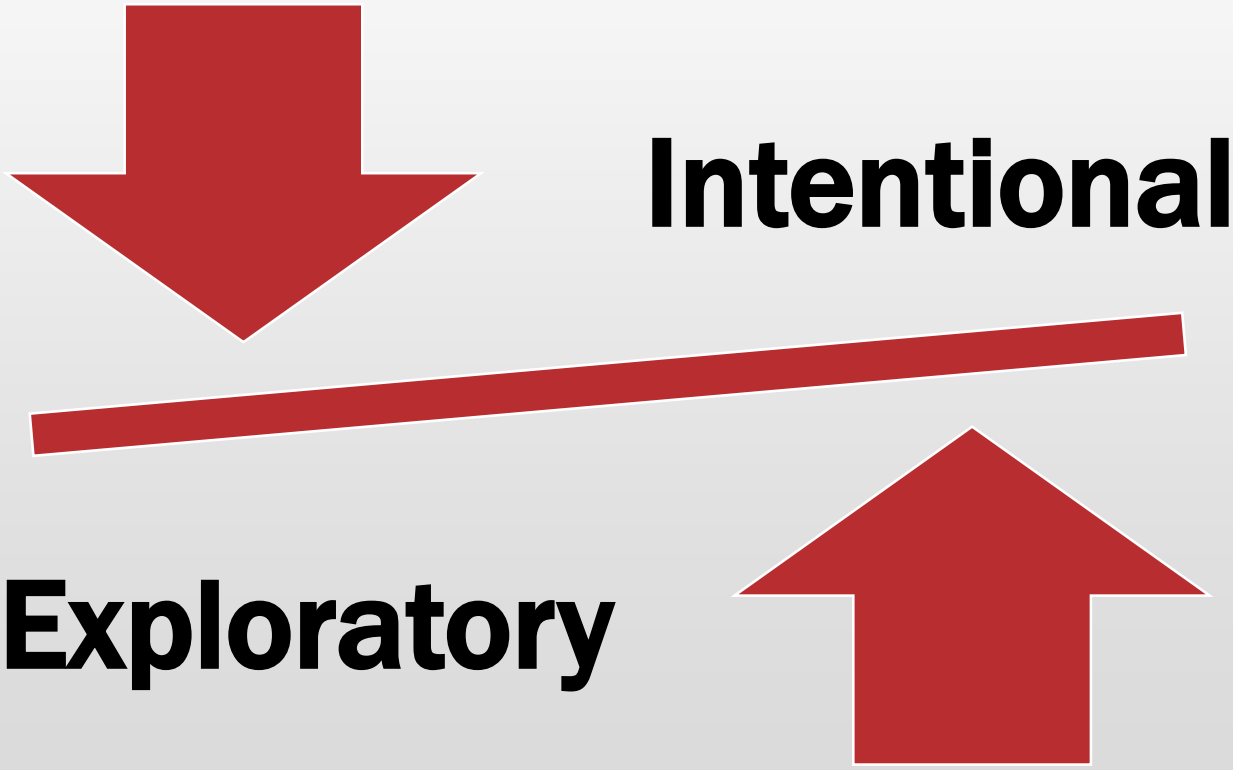

5
Five Innovative Practices
example
Examples
The Merriam-Webster Dictionary
NEW EDITION
Definitions

What is Big Data?

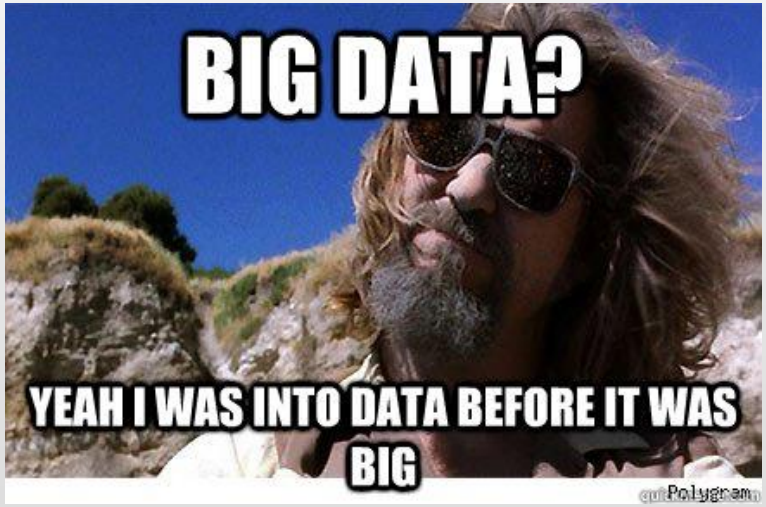
- Big Vs
  - Volume
  - Velocity
  - Variety
  - Veracity

Technical Approach

Strategic Approach



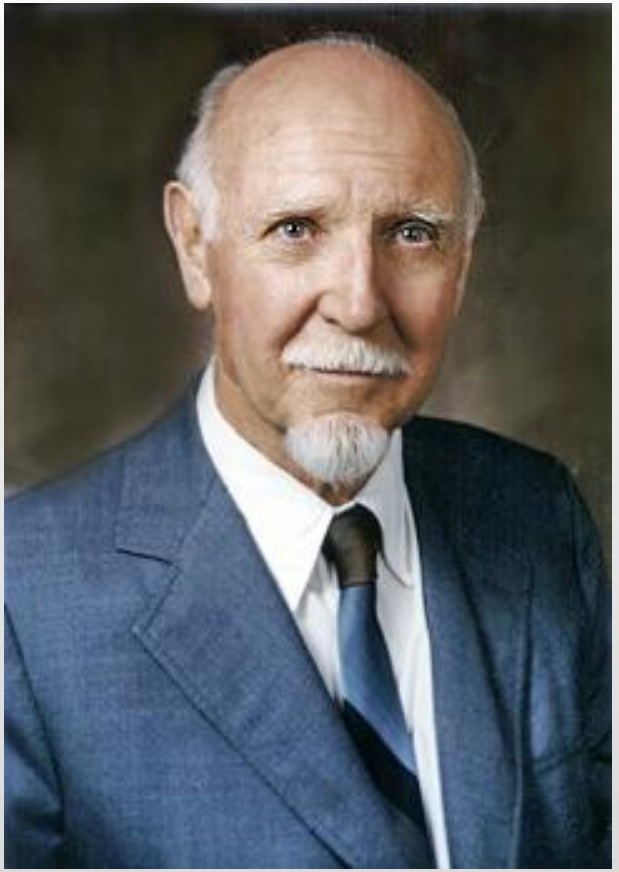
Is Big Data New?



Is Big Data New?

Cattell's Personality Theory

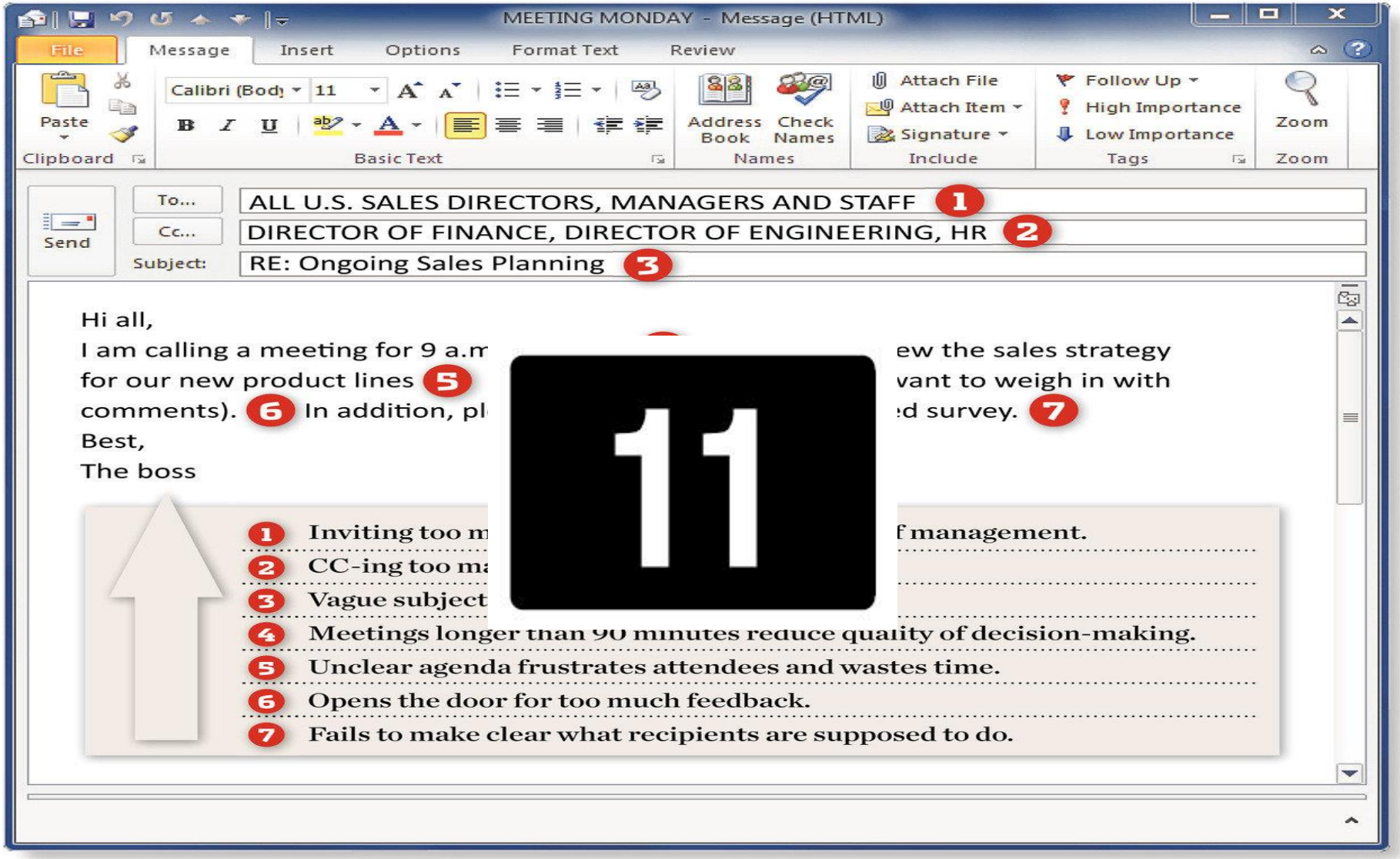
- L-Data - Life Data
- T-Data - Experimental Data
- Q-Data – Questionnaire Data



Netflix

- ✓Kevin Spacey films do well
- ✓British version did well
- ✓Increase in individuals streaming video media

Volometrix


MEETING MONDAY - Message (HTML)
File
Message
Insert
Options
Format Text
Review
Paste
Clipboard
Calibri (Body)
11
Basic Text
Address Book
Check Names
Names
Attach File
Attach Item
Signature
Include
Follow Up
High Importance
Low Importance
Tags
Zoom
Zoom
Send
To...
ALL U.S. SALES DIRECTORS, MANAGERS AND STAFF 1
Cc...
DIRECTOR OF FINANCE, DIRECTOR OF ENGINEERING, HR 2
Subject:
RE: Ongoing Sales Planning 3
Hi all,
I am calling a meeting for 9 a.m
ew the sales strategy
for our new product lines 5
vant to weigh in with
comments). 6 In addition, pl
d survey. 7
Best,
The boss
11
1 Inviting too m
f management.
2 CC-ing too ma
3 Vague subject
4 Meetings longer than 90 minutes reduce quality of decision-making.
5 Unclear agenda frustrates attendees and wastes time.
6 Opens the door for too much feedback.
7 Fails to make clear what recipients are supposed to do.

Big Data in Testing

- Data Sets
  - Registration Data
  - Exam Preparation Data
  - Psychometric Data
  - Recertification Data
  - Marketing/Financial Data (Volume)
  - Other Departmental Data (e.g., Number of Individuals Taking Courses related to exam)
- Data Purposes
  - Volume
  - Satisfaction
  - Competency
  - Validity

The Five Ways

1

Define and operationalize the question your organization seeks to answer

The Five Ways

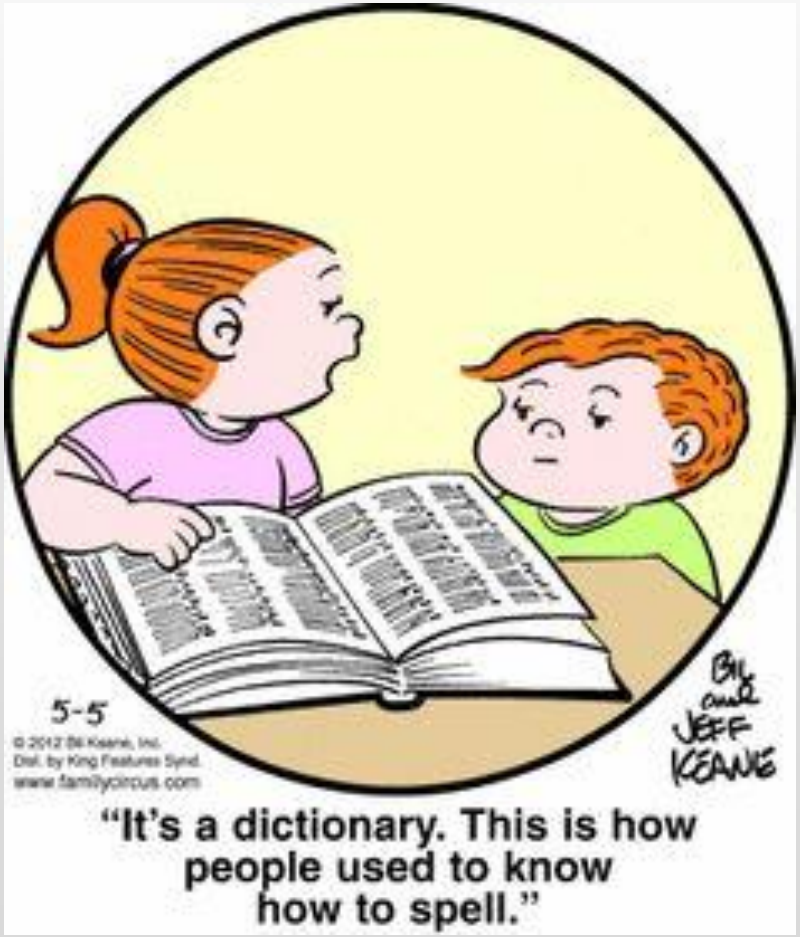
2 *Involve multiple stakeholders when discussing Big Data (e.g., vendors, IT, decision makers, certification holders/SMEs)*

Nobody said Business Analysis will be easy!

The Five Ways

3

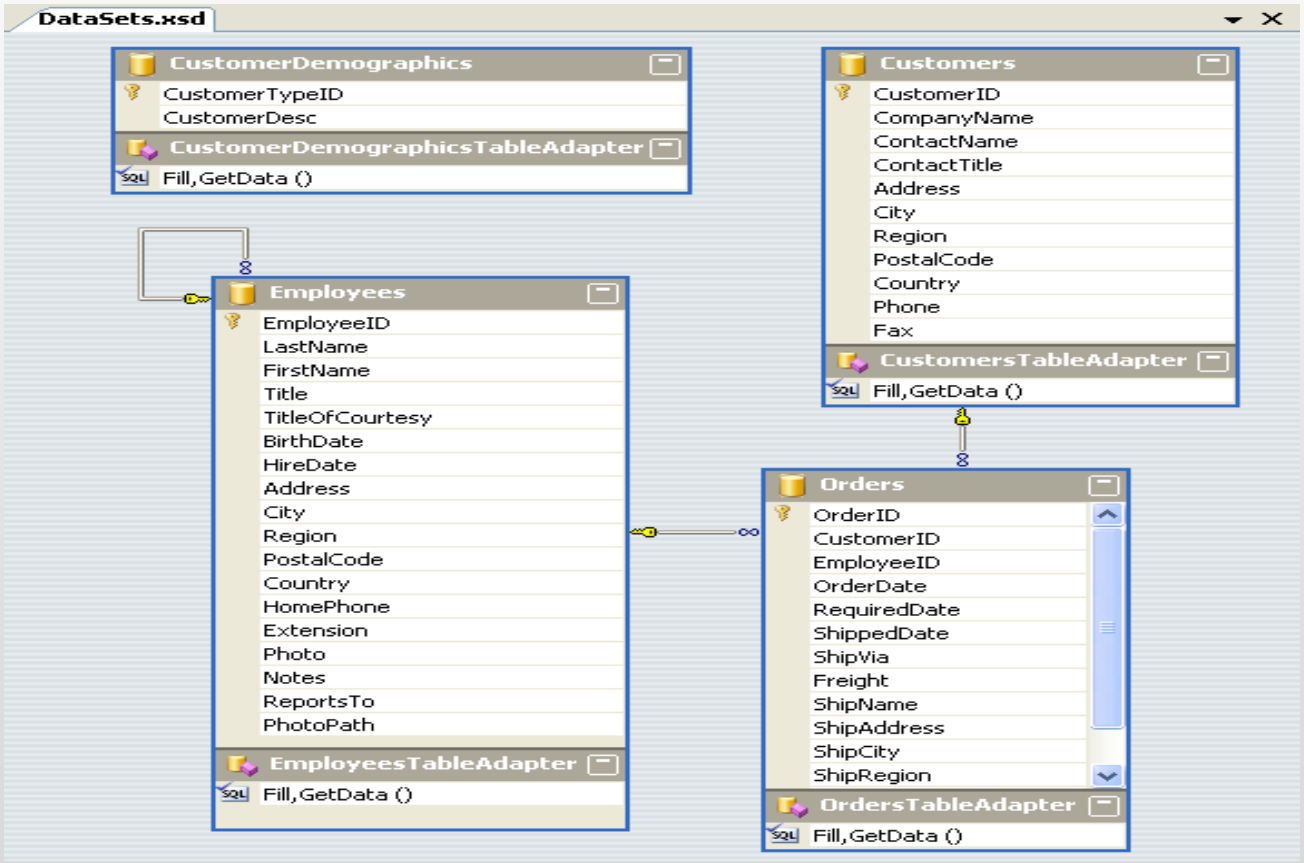
Identify your data sources and develop a data dictionary



The Five Ways

4

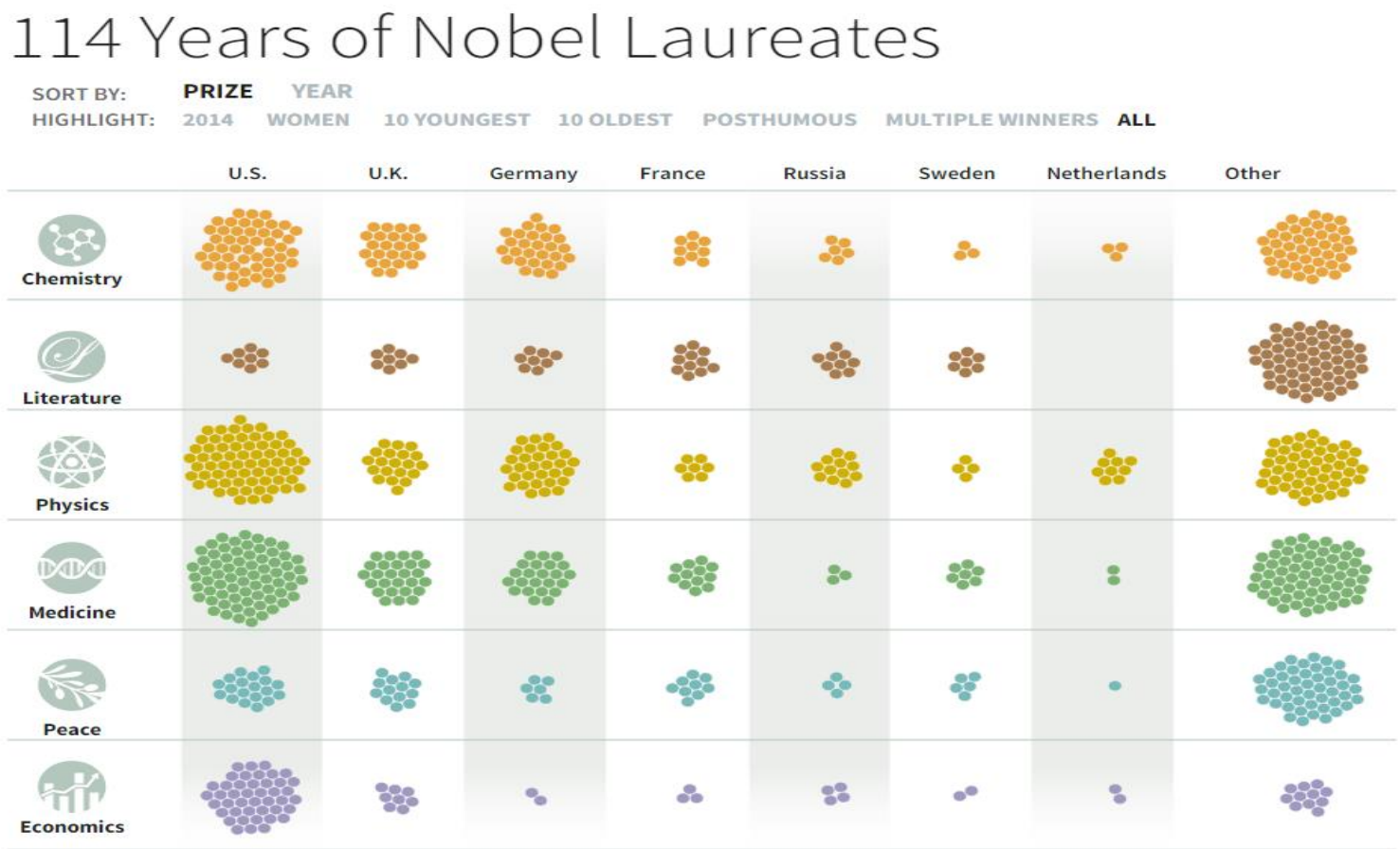
Standardized multiple data sets among common fields



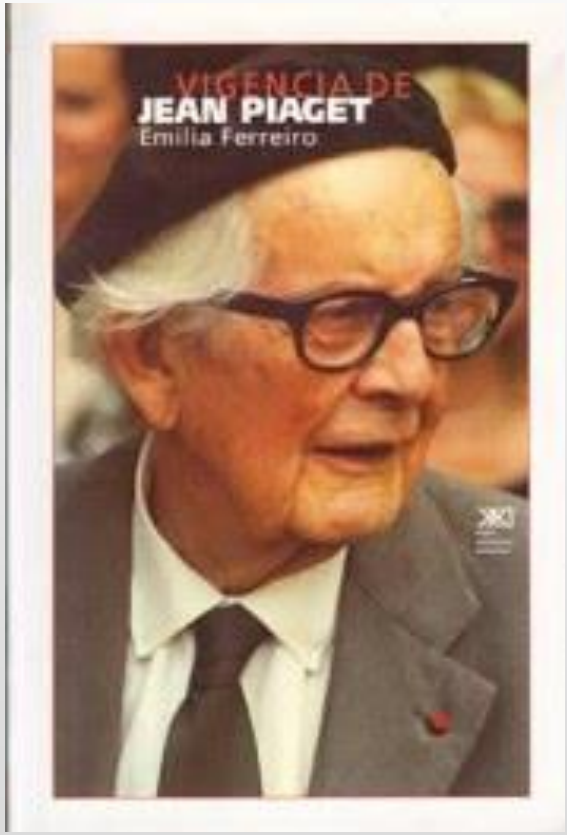
The Five Ways

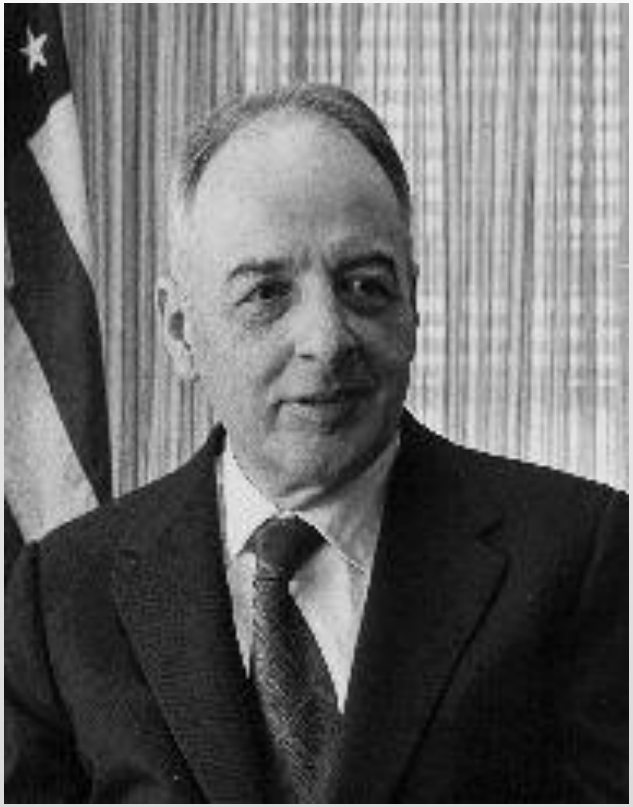
5

Produce simple reports designed to answer the question across multiple stakeholders



Umm, What About The Other Data?





What is Really Important?

Using Big Data to See the Psychometric World Differently

LIBERTY J. MUNSON, PHD

- How are learners interacting with learning resources?
- Identify learning gaps, restructure courses
- What is the relationship between those activities, course completion, and exam success?

https://views.scraperwiki.com/run/gviz_json_output_demo/


```
ousefulHack({"rows":[{"c":[{"v":"A150_1"},{"v":"Speeches and speech-making"},{"v":"Arts and
Humanities"},{"v":"A150"}]},{"c":[{"v":"A200_3"},{"v":"The origins of the wars of the three
kingdoms"},{"v":"Arts and Humanities"},{"v":"A200"}]},{"c":[{"v":"A296_1"},{"v":"Getting started on
Classical Greek"},{"v":"Arts and Humanities"},{"v":"A296"}]},{"c":[{"v":"A300_2"},{"v":"Lewis
Grassic Gibbon: Sunset Song"},{"v":"Arts and Humanities"},{"v":"A300"}]},{"c":[{"v":"B700_1"},
{"v":"The market-led organisation"},{"v":"Business and Management"},{"v":"B700"}]},{"c":
[{"v":"L203_1"},{"v":"Lebenszyklen"},{"v":"Languages"},{"v":"L203"}]},{"c":[{"v":"LDT101_10"},
{"v":"Giving presentations"},{"v":"Skills"},{"v":"LDT101"}]},{"c":[{"v":"MST209_6"},{"v":"Modelling
with differential equations: oscillations"},{"v":"Mathematics and Statistics"},{"v":"MST209"}]},
{"c":[{"v":"MST209_7"},{"v":"Systems of differential equations"},{"v":"Mathematics and
Statistics"},{"v":"MST209"}]},{"c":[{"v":"MST209_8"},{"v":"Modelling with systems of differential
equations"},{"v":"Mathematics and Statistics"},{"v":"MST209"}]},{"c":[{"v":"MST209_9"},
{"v":"Fourier series"},{"v":"Mathematics and Statistics"},{"v":"MST209"}]},{"c":[{"v":"SK183_1"},
{"v":"Nutrition: Proteins"},{"v":"Science"},{"v":"SK183"}]},{"c":[{"v":"T152_1"},{"v":"Welcome to
saving energy"},{"v":"Engineering and Technology"},{"v":"T152"}]},{"c":[{"v":"Y156_1"},
{"v":"Understanding children: Babies being heard"},{"v":"Education"},{"v":"Y156"}]},{"c":
[{"v":"Y157_2"},{"v":"Understanding society: Families"},{"v":"Social Sciences"},{"v":"Y157"}]},
{"c":[{"v":"Y158_1"},{"v":"Understanding health - taster materials"},{"v":"Health and Social
Care"},{"v":"Y158"}]},{"c":[{"v":"Y159_3"},{"v":"Understanding management: I'm managing thank
you!"},{"v":"Business and Management"},{"v":"Y159"}]},{"c":[{"v":"Y160_2"},{"v":"Making sense of
art history"},{"v":"Arts and Humanities"},{"v":"Y160"}]},{"c":[{"v":"Y161_2"},{"v":"Introducing the
environment: Ecology and ecosystems"},{"v":"Science"},{"v":"Y161"}]},{"c":[{"v":"Y163_2"},
{"v":"Starting with psychology"},{"v":"Psychology"},{"v":"Y163"}]}],"cols":
[{"type":"string","id":"unitcode","label":"Unit Code"},{"type":"string","id":"name","label":"Unit
name"},{"type":"string","id":"topic","label":"Topic"},
{"type":"string","id":"parentCourseCode","label":"Parent Course"}]})
```


Understanding how Learners Learn

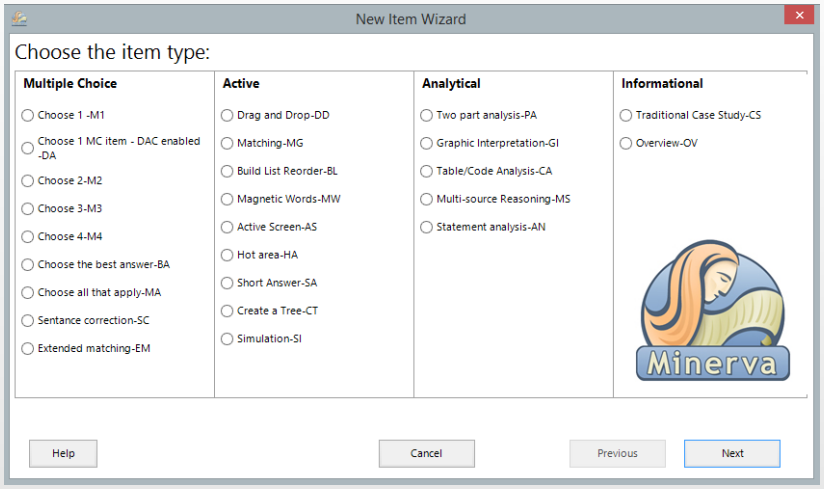
We can track every move the individuals make in online courses and exams, assessments, and other learning activities

“Click Stream Data” will provide valuable insights into learning and assessment…

Understanding how Test Takers Take Exams

- Big data, exams, and items
  - Understand how test takers are interacting with item types
  - Are they following instructions?
  - Do they understand what's required?
  - Where are they struggling and why?

Item Type Performance



KEY INSIGHTS

- Which item types are easiest? Most difficult?
- Which item types are the most discriminating?
- Does this vary by audience?
- Do certain variations of different item types perform better than others?
  - What is the “sweet spot” for number of answer choices for build lists, drag and drop, etc.
- How do item sets differ (e.g., case studies vs. non-case study items)?

Understanding Behavior During the Exam

You have content that people want!

Data forensics

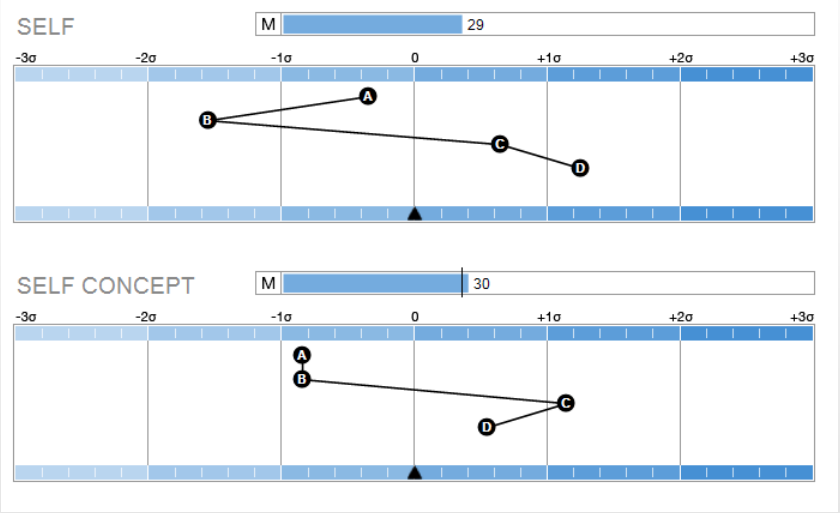
What are learners doing during the exam?
Is it aberrant?

Validation in Big Data

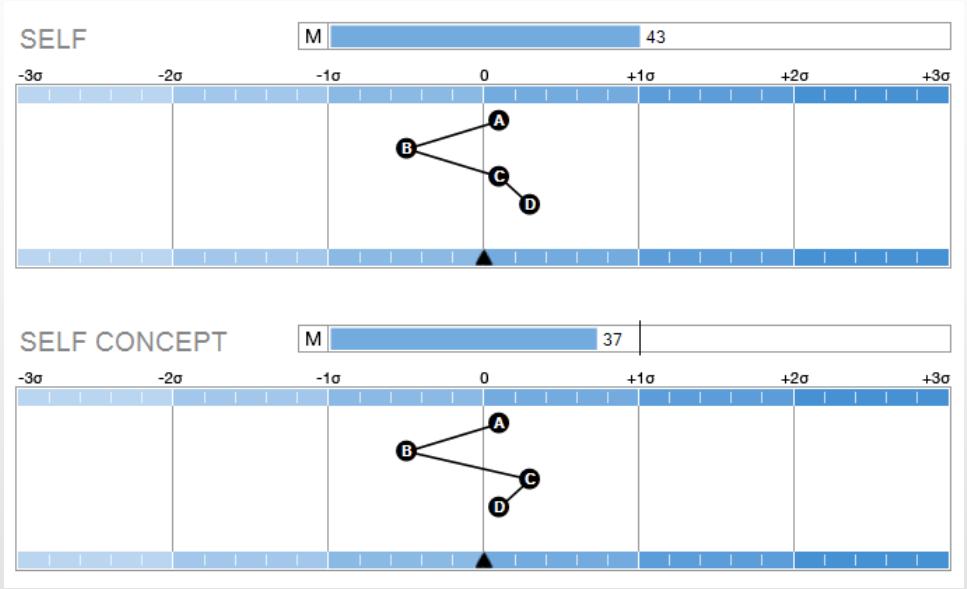
AUSTIN FOSSEY, MA

Background

- PI Behavioral Assessment
  - Released in 1955
  - Measures four behavioral drives
  - Free-choice adjective checklist
  - Configural interpretation
  - Workforce use
  - Approximately 2 million administrations per year



THE PREDICTIVE INDEX

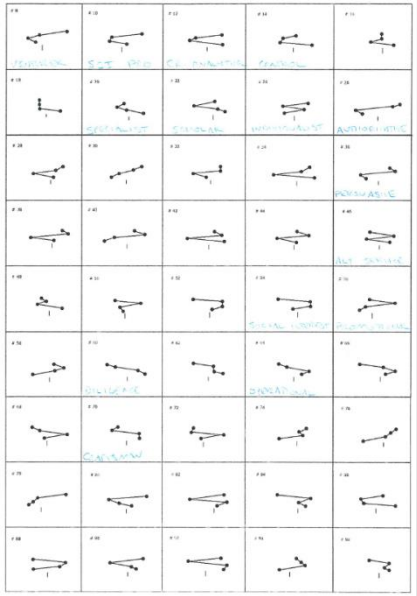


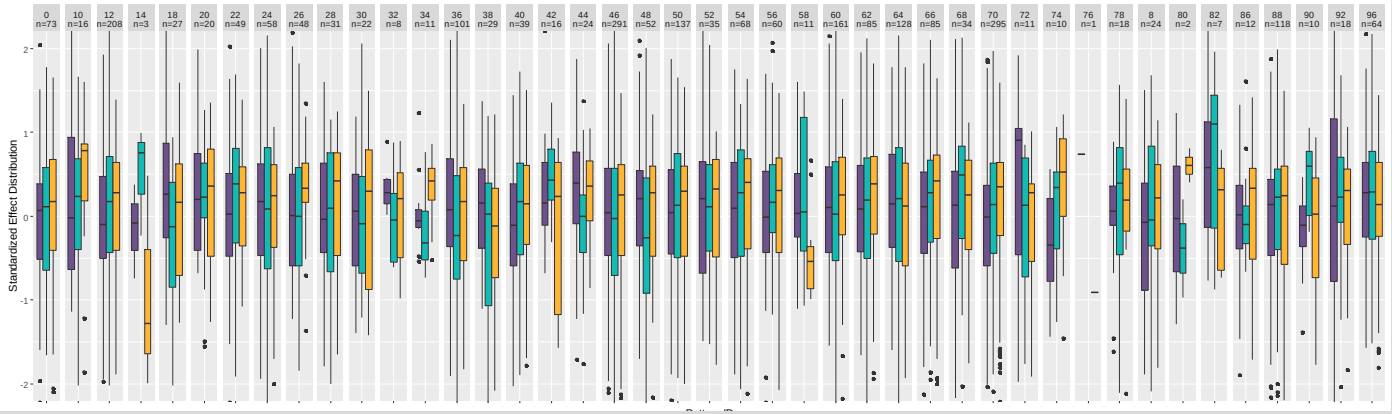
AN EXAMPLE: AUSTIN FOSSEY

Low Extraversion: Awkward Public Speaker

1. Define and Operationalize the Question

- What is the best way to describe similar personality configurations?
- Help improve:
  - Reporting
  - Valid interpretations
  - Competitiveness



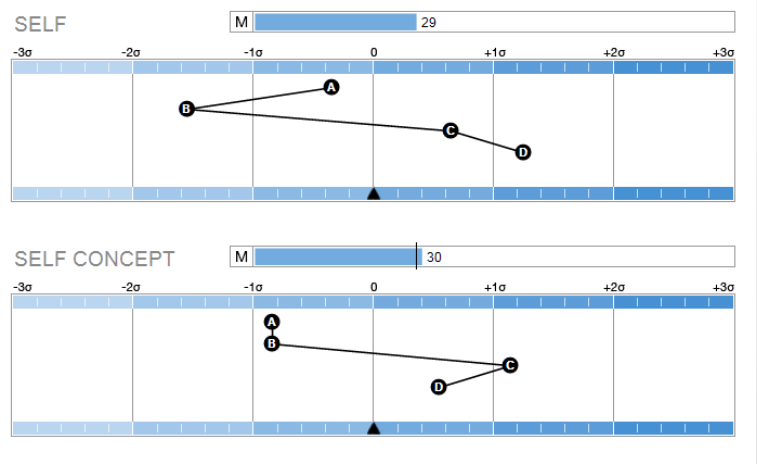


2. Involve Multiple Stakeholders

- What are the key defining elements that make two people similar?
- Stakeholders
  - Test users (hiring managers/consultants)
  - Trainers
  - Product designers
  - Psychometricians
  - Business leaders/policy makers
- Surveys, interviews, round tables

3. Identify Data Sources, Develop Data Dictionary

- What are the key defining elements that make two people similar?
- Evidence Model Candidates
  - Score magnitudes/distances
  - Overall distance
  - Relative placement (shape)
  - Position relative to center
  - Direction (away from or towards center)
  - Different weightings
- Potential to identify groups with clustering algorithm



3. Identify Data Sources, Develop Data Dictionary

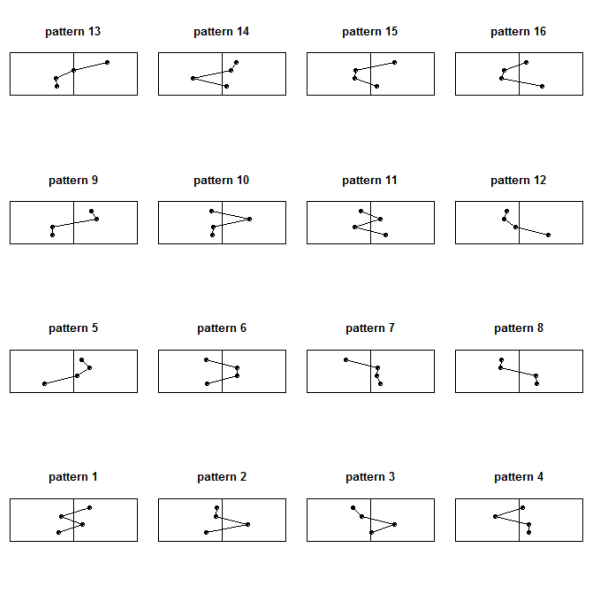
- How to evaluate measurement model?
- Find out how SMEs think about fit
  - Survey of comparisons
  - Ratings by factor score and overall patterns
  - Collect comments on their thought processes/strategies
  - Model SME fit ratings using evidence model candidates

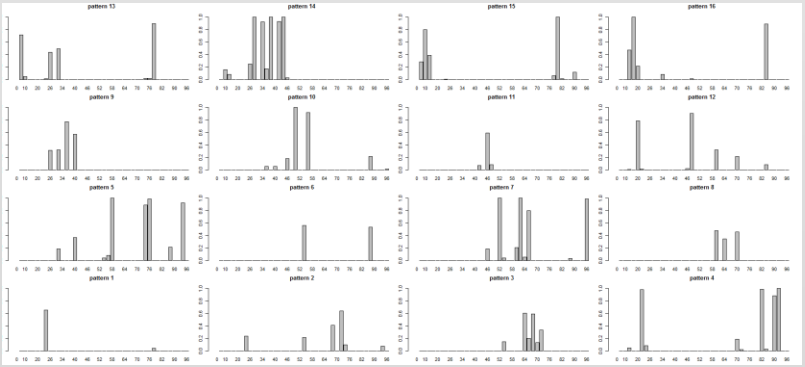
3. Identify Data Sources, Develop Data Dictionary

- Want to compare SME logic tree to two competing measurement models
- Need to select algorithms that are appropriate for the data
- Selected Self-Organizing Maps (SOM) and Fuzzy Clustering
  - Have been used with assessment data by other researchers
  - Appropriate for the data
  - Provide output that could help with reporting, describing similarities

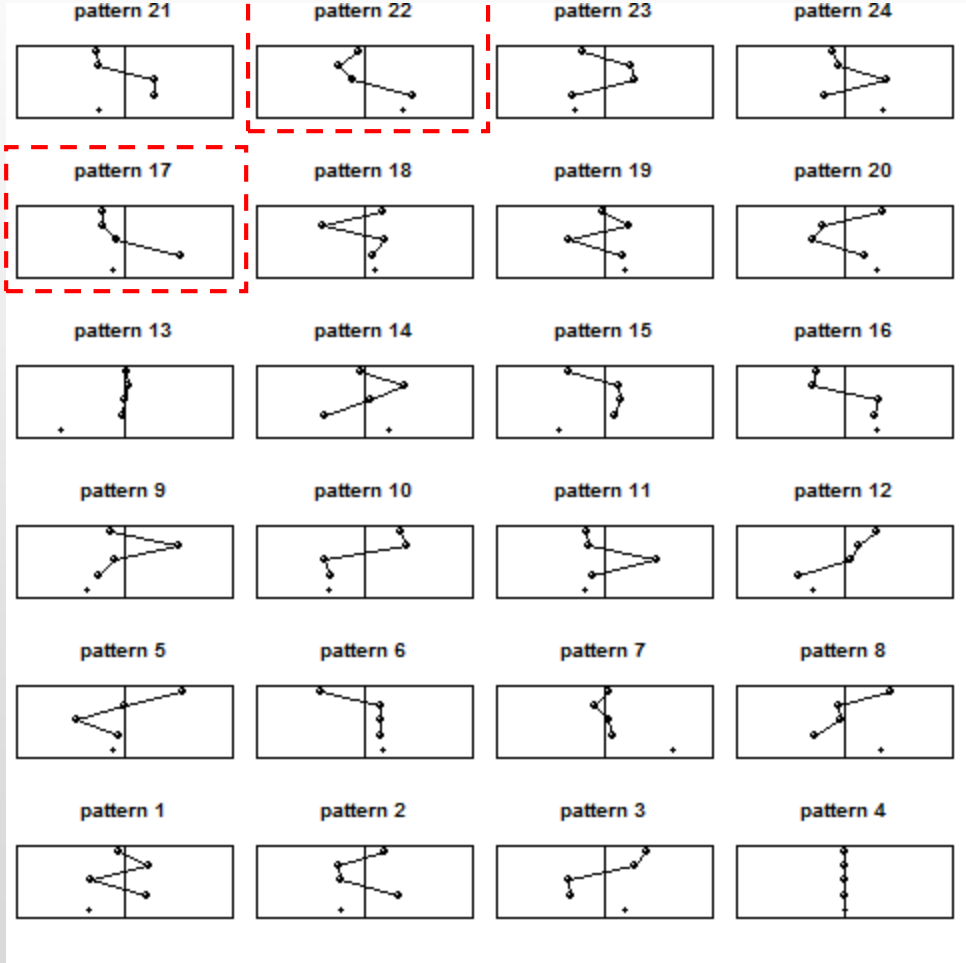
4. Standardize Multiple Data Sets

- Run clustering algorithms and SME logic tree against sample data
- Vary parameters
  - Number of groups (clusters)
  - Algorithm-specific parameters
  - Evidence data
- Standardize cluster identifiers across solutions
- Evaluate potential solutions
  - Interpretability for SMEs
  - Frequency distributions
  - Davies-Bouldin index
  - Cohen's Kappa/% Agreement

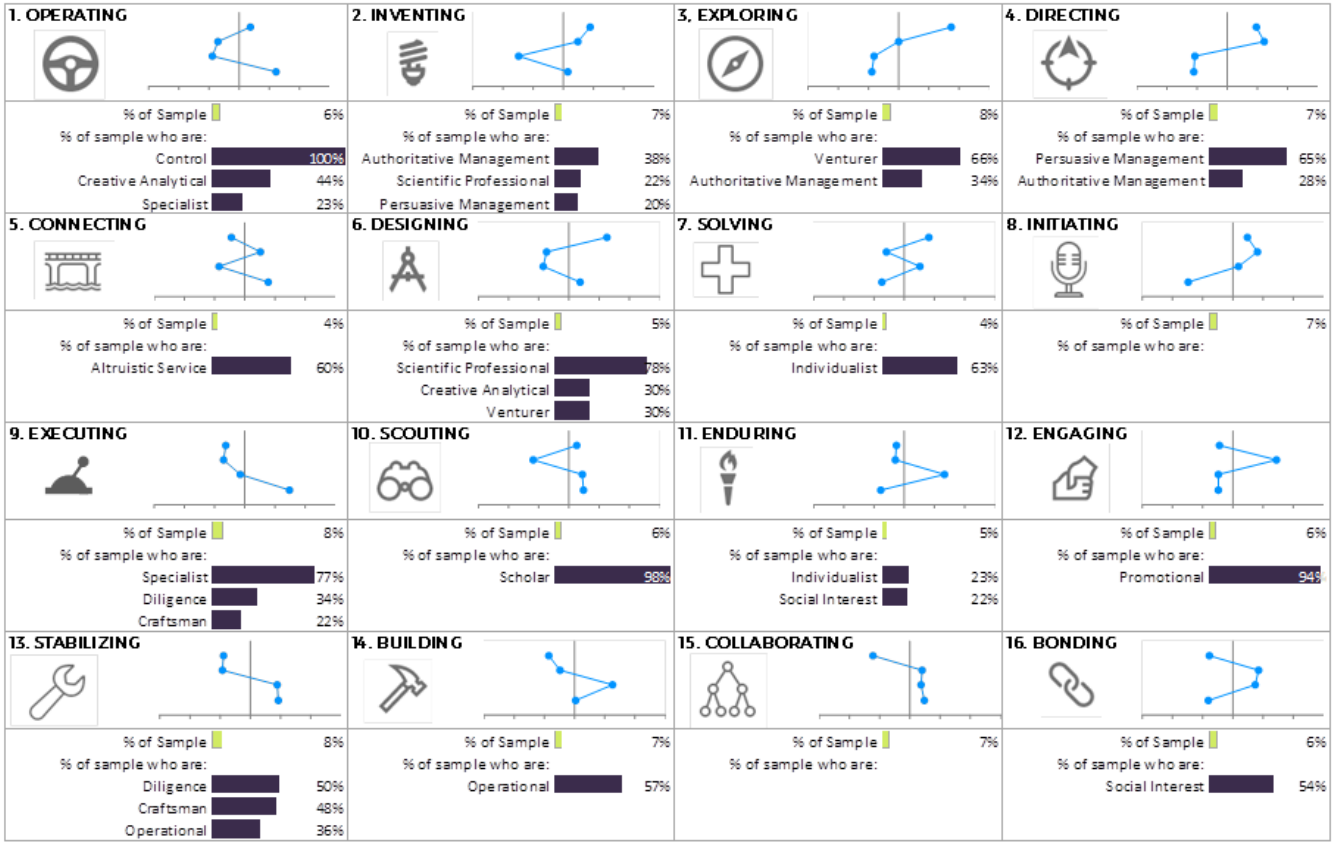




4. Standardize Multiple Data Sets



5. Produce Simple Reports



Next Steps

- Select proposed solution
- Check against existing criterion validity studies
- Discuss with stakeholders
- Create transition plan
- Update technical documentation
- Product implementation

Thank you!

manny@aerexperts.com

President and Founder