# Performance evaluation

Examples
Teach for America: optimize hiring process
- Rigorously tracks the performance of all teachers, comparing it to evaluations when they were hired.
- Helps refine the most productive steps in the hiring process

Google
- Systematically track interview predictions about new hires to figure out how good they were at it
Answer Not very
- So dramatically reduced the number of interviewers.

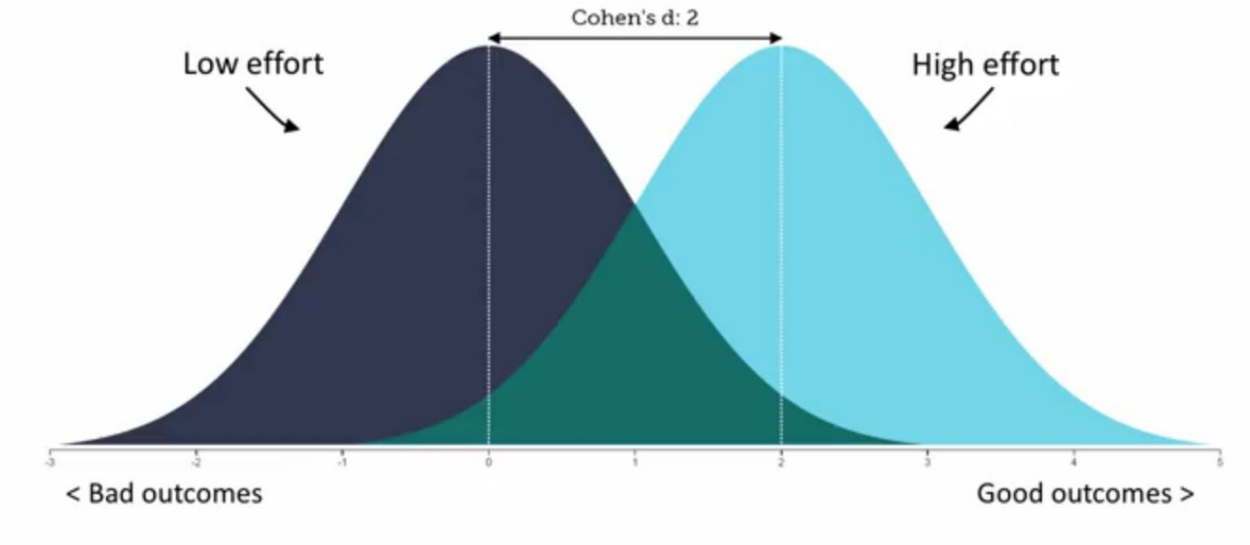
Credit Suisse
- 3-year study: changing jobs internally increases employee stickiness
- Increased internal postings of open jobs from <50% to >80%

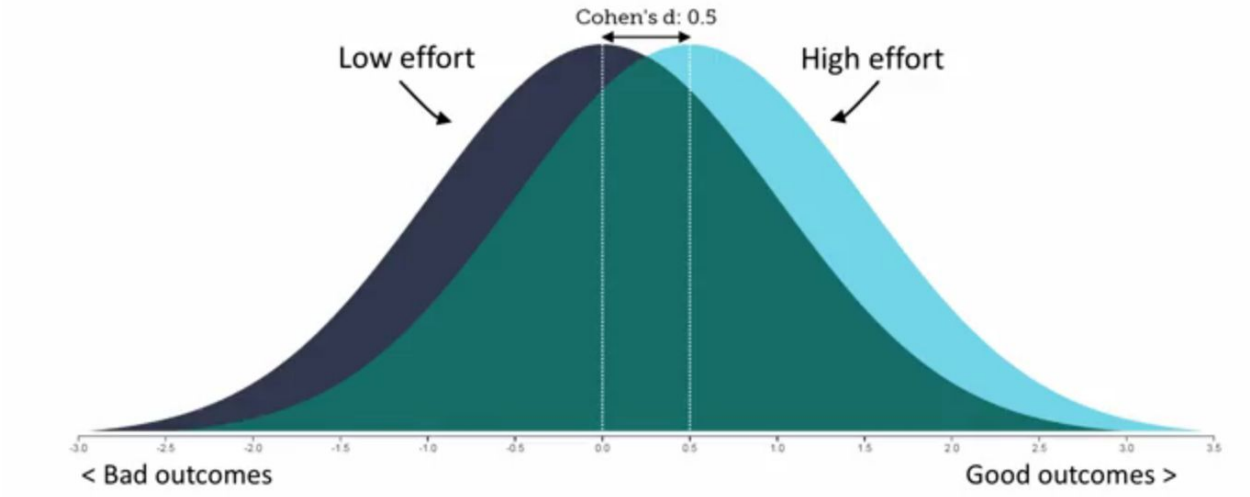
Many firms are now finding:
- better levers for retaining key employees
- More diagnostic methods for hiring
- Who their most valuable employees are
- How to compose the most productive teams

- Purpose of performance evaluation
    - feedback
    - Rewards / punishment
- Performance valuation, not talent management. Tough to compare employees if not in identical situations.
- Helpful starting place. Begin by assuming all employees have equally ability

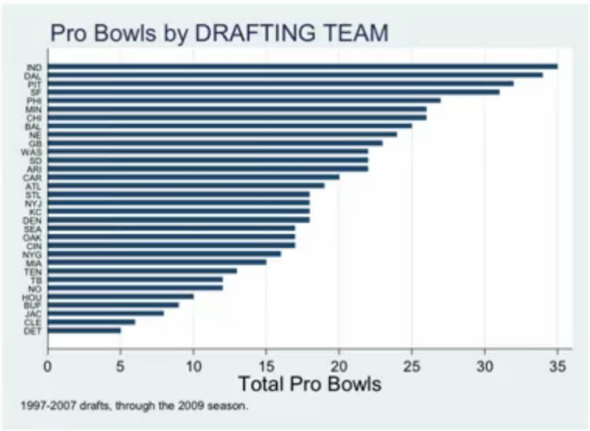
Performance measures are noisy

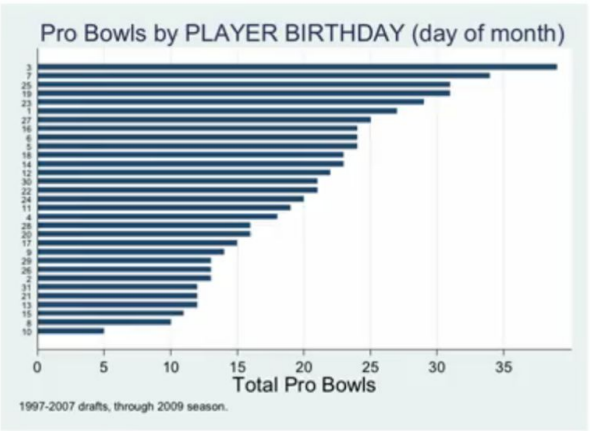
- For any given level of effort, a range of outcomes can occur due to factors outside the employee's control
    - Competitors, team members, the boss, the economy
    - The challenge: Separating skill from luck
Separate low/high efforts




Chance vs Skill:
Example NFL (Moneyball): Separate signal from noise
Draft outcomes are completely due to random chance


Skill and chance in the Draft
- Clearly there is skill involved
- Are there differences in skill?
- Are some teams better than others at picking players?

Robust pattern
- Performance statistic
    - Games started vs pro bowls vs compensation
- Player's career stage
    - 1st 5 year vs free-agent year (5th)
- Additional norming
    - Player position
- Decision-making unit

    - actual individual in charge of a team's draft
- Draft stage
    - 1st 3 rounds vs last 4 rounds

A simple model
there are two components to performance:
- In informal terms: Real tendency + Luck
- In more formal terms: y = x +3
    - x = true ability
    - e = error

Regression to the mean
- Anytime you sample based on extreme values of one attribute, any other attribute that is not perfectly related will tend to be closer to the mean.
    - Attributes can be
    Performance at different points in time
    Different qualities within the same entity

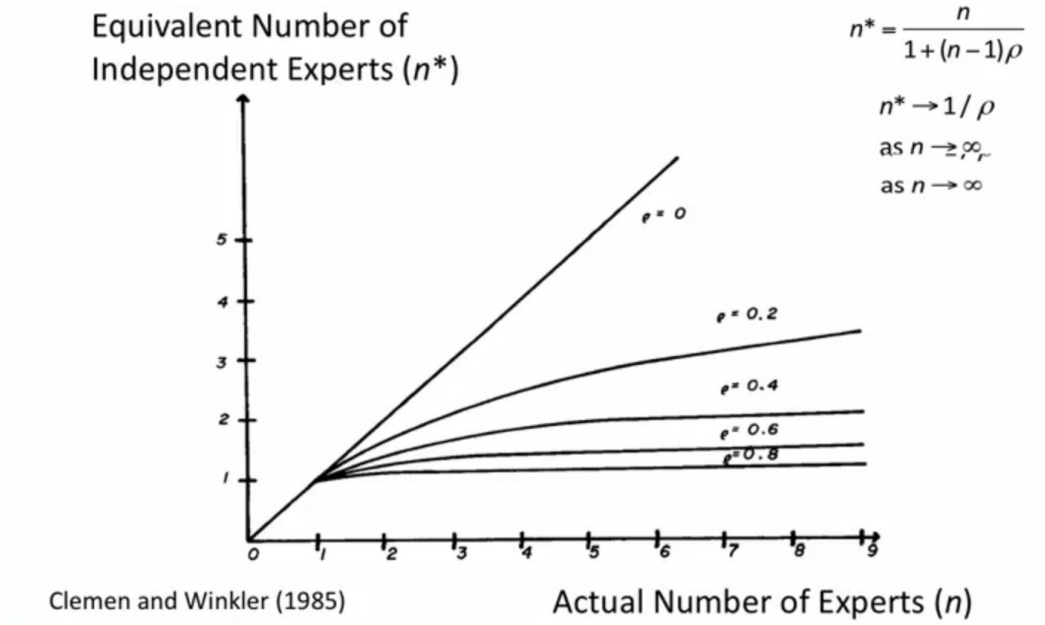
What gets in the way of seeing this?
- Outcome bias
- Hindsight bias
- Narrative seeking

Make sense of the past
- We find a story that connects all the dots
- Chance plays too small a roel in these stories

Extrapolating from small samples
- Sample size
    Principle: Sample means converge to the population mean as the sample size increases. (Central limit theorem)

Signal independence
- The wisdom of crowd: The average of a large number of forecasts reliably outperforms the average individual forecast.
- But the value of the crowd critically depends on the independence of their opinions.
- Independent means uncorrelated.
- If correlated, the value of additional opinions quickly diminishes.

Impact of Correlation




Sources of correlation between two opinions
- They've discussed it already
- They talk to the same people
- The have the same background - from the same place, trained the same way, same historical experiences.

Process vs Outcome
Consider broader set of objectives
- Organizations generally care about how a person goes about his/her job
    - Most important: impact on others
- People consider too few objectives
    - Systematically omit nearly half of the objectives they later identify as personally relevant
- Leads many firms to rely on too narrow a set of performance measures.

Example
- Dell changed their performance evaluations
    - Before change: 100% results
    - After change:
    - 50% what an employee accomplished
    - 50% how he/she accomplished it, as judged by those affected

- The more uncertainty in the environment, i.e. the less control an employee has over exact outcomes, the more a firm should emphasize process in their evaluations.

Focus on process
Use analytics to better understand, and focus on, the processes that tend to produce desired outcomes.
- Key issue: identify the fundamental drivers of value.
    Example: ice hockey: possession>shots>goals

Critical questions
- Are the differences persistent or random? How do we know this isn't just good/bad luck?
- Is the sample large enough to draw strong conclusions? How can we make it larger?
- How many different signals are we really tapping into here? How can we make them as independent as possible?
- What else do we care about? Are we measuring enough? What can we measure that's more fundamental?

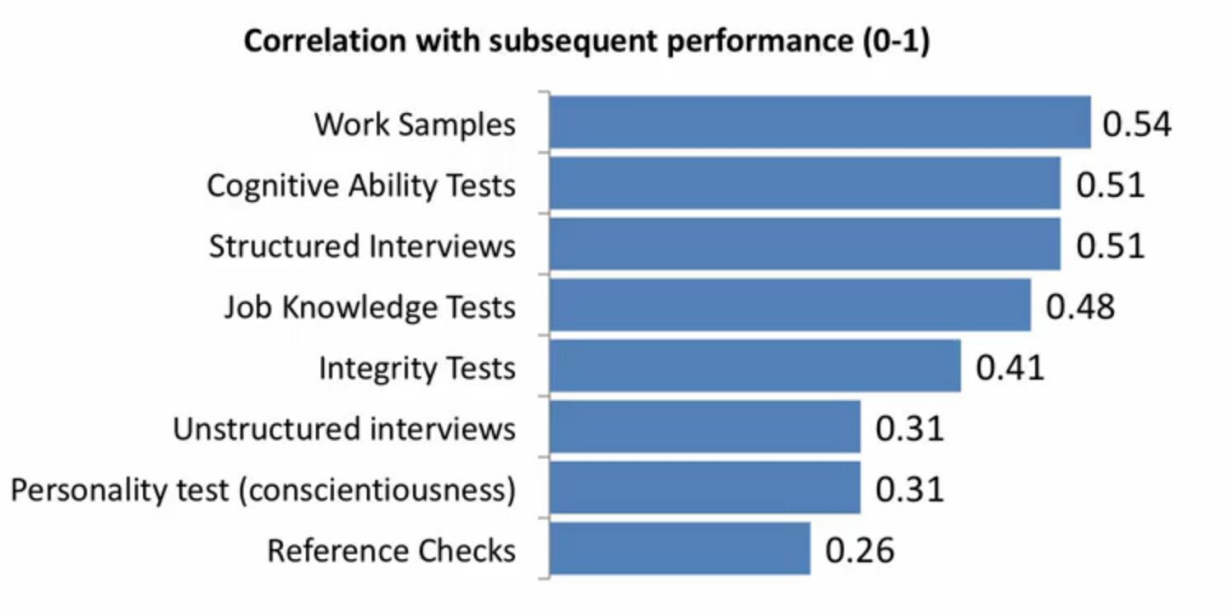
# Staffing analytics

The staffing cycle graph
- Basic facts about staffing processes
- The value of analysis
- Possible analytic approaches


Hiring 1: Predicting Performance
With a lot people applying for the position, we want to know who will perform the best for the job.

Methods to evaluate job candidates is most effective in the job
- Job knowledge tests
- Cognitive ability test
- Personality tests

- Reference checks
- Structured interviews
- Unstructured interviews
- Work samples
- Integrity tests




Reference test:
The predictor for the quality of a reference is the person writing the reference, rather than the person the reference is about.

Hiring 2: Fine-tuning predictors
Performance
- Performance evaluation
- Objective performance metrics
    - Sales
    - Productivity
    - Customer satisfaction
- Attrition
- Rate of promotion

Predictors
- Background/experience
- Test scores
- Interview performance
    - Specific questions
    - Specific interviewers

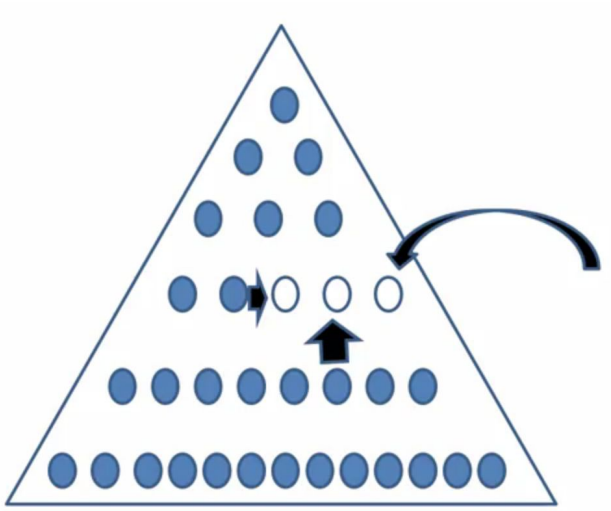
Google: think out of the box

e.g. How many call boxes are there in manhattan?
How many golf ball could you fit into a jumbo jet?

GPAs no for google but yes for investment bank
Flight attendants: helpful>friendly
Use the predictors

Predicting new hire performance: hints, tips, issues
1. Comparing Apples with Apples
- work
- location
- manager/unit
- level
- time in the job
2. disentangling influences
- beware spurious correlations
- apply common sense/understanding
3. Accounting for selection
- who got hired
- who stayed

Hiring 3: Using Data Analytics to predict performance
Approaches to predicting hire performance
Okay:
- Compare characteristics of best and worst performers
- Test for statistical significance
Better:
- Compare characteristics of best and worst performers within same cohort and job
Good:
- Use multivariate regression to separate out influences of different characteristics
Best:
- Use multivariate regression to separate out influences of different characteristics
- Apply selection correction to account for who was hired and attrition from sample



Comparing Data Analysis with human judgement
Bad news
- Combination of various tests and selection methods leaves much of performance unexplained
Worse news
- Implementation of algorithms reduced turnover in call centers
- Turnover was lower the less often managers over-rules the algorithm

Bringing data to internal mobility: beyond the Peter principle

Internal mobility 1:
The Peter principle
"In time, every post is occupied by an employee who is incompetent to carry out its duties"
OR
How well does success in the current job predict performance in a higher level job?
OR
What does predict success in higher level jobs?

The question is,
to what extent does success in a current job actually predict performance in the next job?
Or,
What does predicts success in higher level jobs?

Analyzing promotability
Requirements
- Multiple dimensional
    - Output measures
    - Competence
    - Assessments

Potential

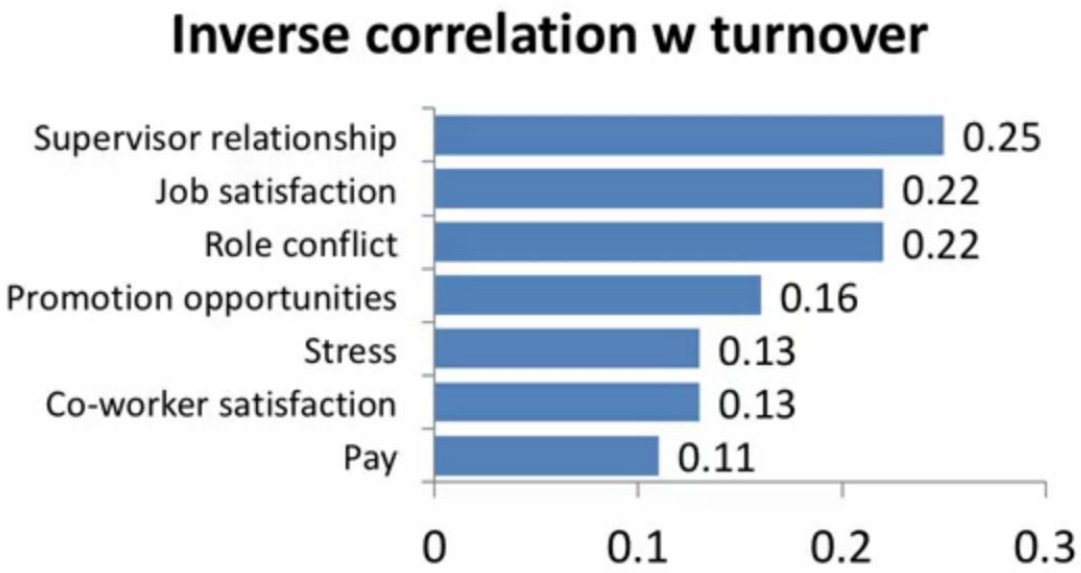
- Which dimensions of lower level performance best predict performance in the higher level job?

Good salesperson: good sales
Good salesperson: good sales, but also good team sales
Promotion to manager based on sales is not guaranteed for success.

Internal mobility 2:
Optimizing movement within the organization
Evaluating staffing options
- Which routes lead to better performance
- What is the effect on cost?
Trade offs:
- should I be doing more to promote internal development and promotional mobility?
- Is that money well spent?
- would it be better of spending that money to try and beef up our recruitment efforts?

A practical example
- Personnel data from large investment banking division
- Wide variety of functions
- Annual snapshots of employees from 2003-2009
    - Performance evaluations
    - compensations
    - job
- Focus on effect of how workers entered their current job (hired versus promoted)
- Use very detailed job controls to compare workers entering similar jobs by different routes
- study only jobs that
    - can be entered by promotion
    - observe being filled

Some results
Performance
- Hires performed substantially worse than similar promotes
- Takes 3 years to acquire similar performance to those promoted into the job
Pay
- New hires receive 18% more compensation than promotes

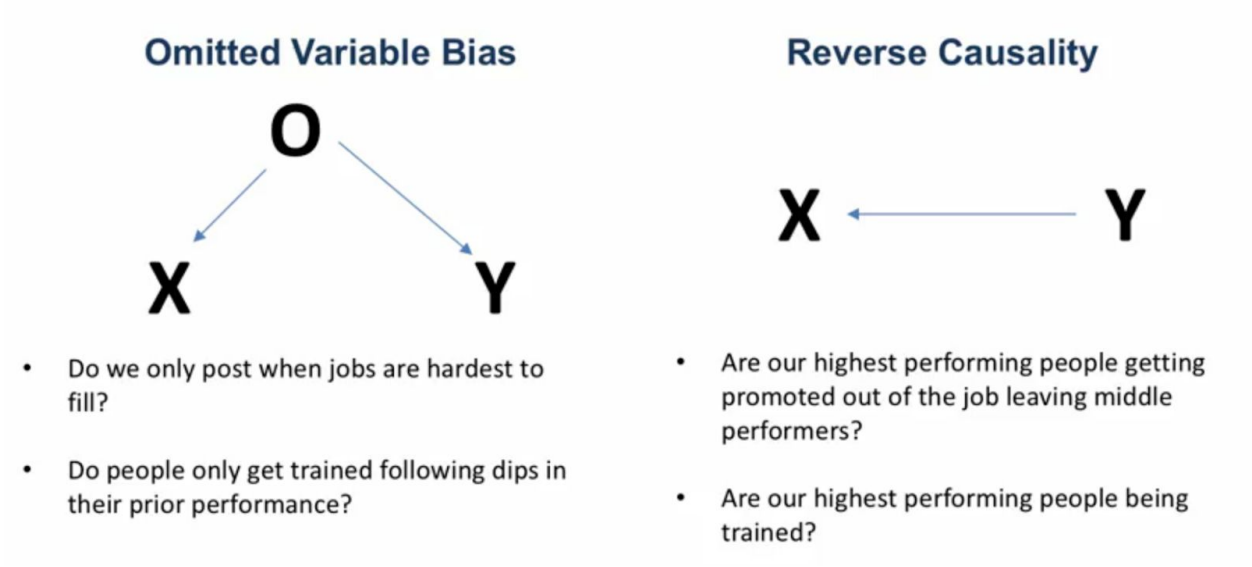



Comparing hiring inside firms
internal posting, internal market
- manager posts job & invites interested candidates to apply
- creates competition for jobs within the firm
Sponsorship, social network
- manager identifies candidates through their personal network
- Appoints preferred candidate to the job

Internal posting
- creates unconventional career paths
- leads to higher performance ratings:
    - larger pool of candidates
    - disciplines decision-making
- associated with higher salaries (3%-6%)

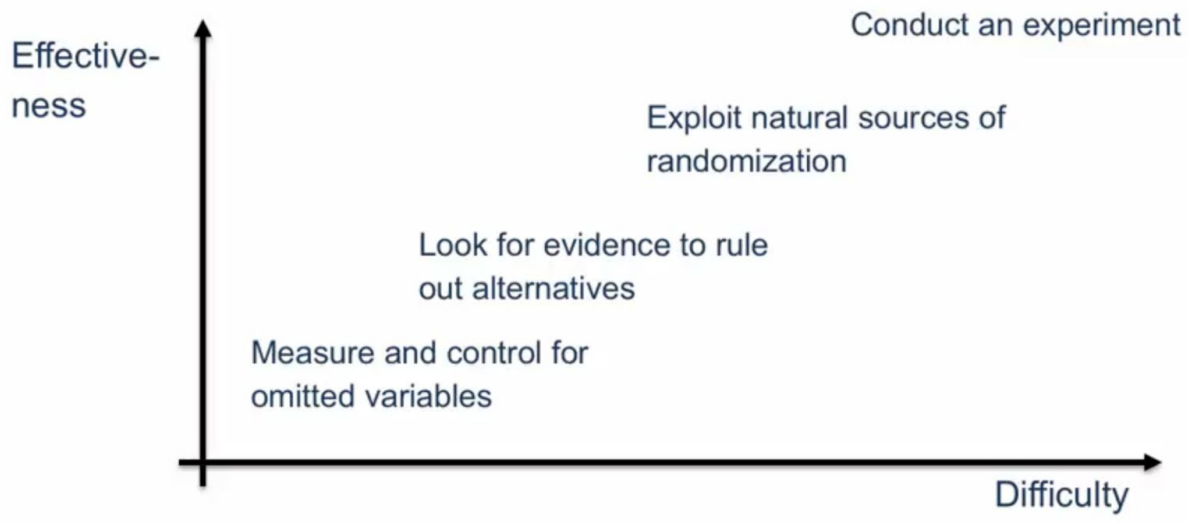
Causality
Why we care about causality
- People who enter jobs through formal posting perform worse
Should we avoid posting?
- People who have been in the job longest have lower performance?
Should we move people around more
- People who have taken a training program perform better
Should we send more people to training
- People who have taken a training program show greater performance improvements
Should we send more people to training?

Two types of causality problems




The central, underlying question
What is leading to difference in our main predictor variable?

Approaches to addressing causality




Measure and control for omitted variables
- collect data on possible omitted variables and
    - include in regressions
    - create matched pairs with similar values

- examine within person changes to hold person constant
But
- not everything can be measured

Look for evidence to rule out alternatives
- What would be some implications of alternative explanations?
- Can you find evidence for or against those explanations in the data?

Exploit natural sources of randomization
- Natural experiments change your X variable in ways that should not also affect Y
- Mimics assignment to treatment vs control group in genuine experiment
- allows for assessment of causal effects
But
you need to be lucky

Conduct an experiment
- Randomly assign individuals/jobs to treatment and control groups (ensuring balanced characteristics of each group)
- Test whether results in two groups are different
But
- You need to persuade people to let you do it
- Very time-consuming

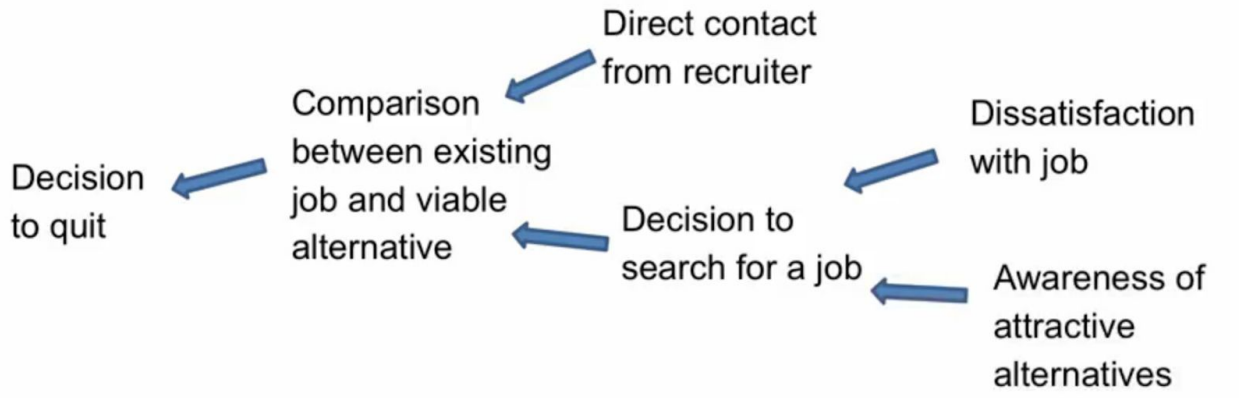
Attrition: understanding and reducing turnover
Problems
- Hiring costs
- training costs
- loss of critical knowledge
- impact on customer relationships

Levers
- inform hiring strategy
- target interventions
    - improve conditions
    - address unmet needs
    - train managers
    - focus retention efforts

Understanding attrition
People leave their jobs because there is something else that they would rather be doing
- attractiveness of outside opportunities
    - demand for skills
    - industry / regional growth

- planned career evolution
vs
- satisfaction with current job situation
- perception of future opportunities / trajectory in organization

Turnover as a search process

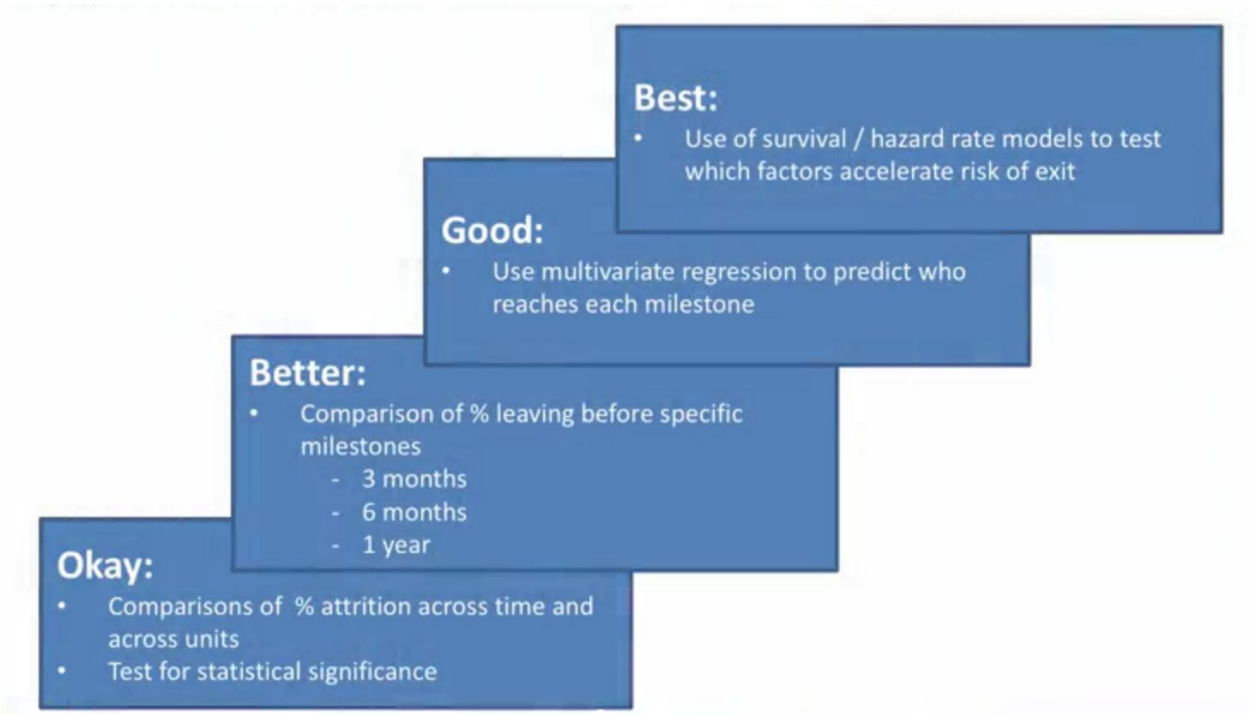



Assumptions:
- We will enjoy and be better at jobs that are a better fit with our abilities and preferences
- We can only assess fit once we are actually in the job
- If we turn out to be a poor fit, we will quit

Implications
- Probability of turnover decreases the longer people have spent in the job
    - Have learned whether it is a good fit or not
- Rate of turnover falls as workers get older
    - More likely to know what fits and what doesn't

Some common predictors of turnover
- manager
- pre-hire background
- type of work / project / function
- performance evaluations
- geography
- social network behavior

Approaches to predicting attrition




survival model
look at the continuous time

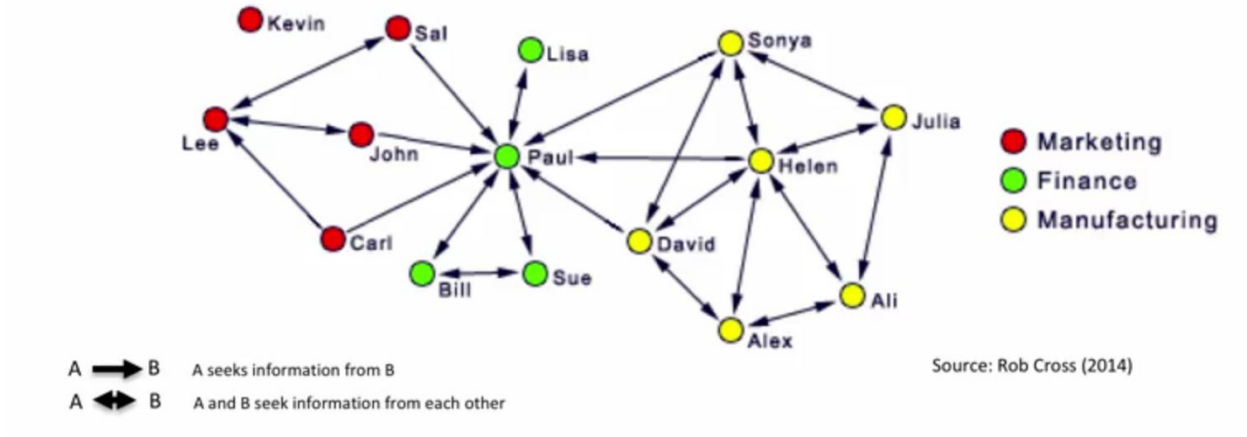
## Collaboration

Collaboration is the action of working with others to produce or create something.
Our focus:
Collaboration between employees inside an organization.

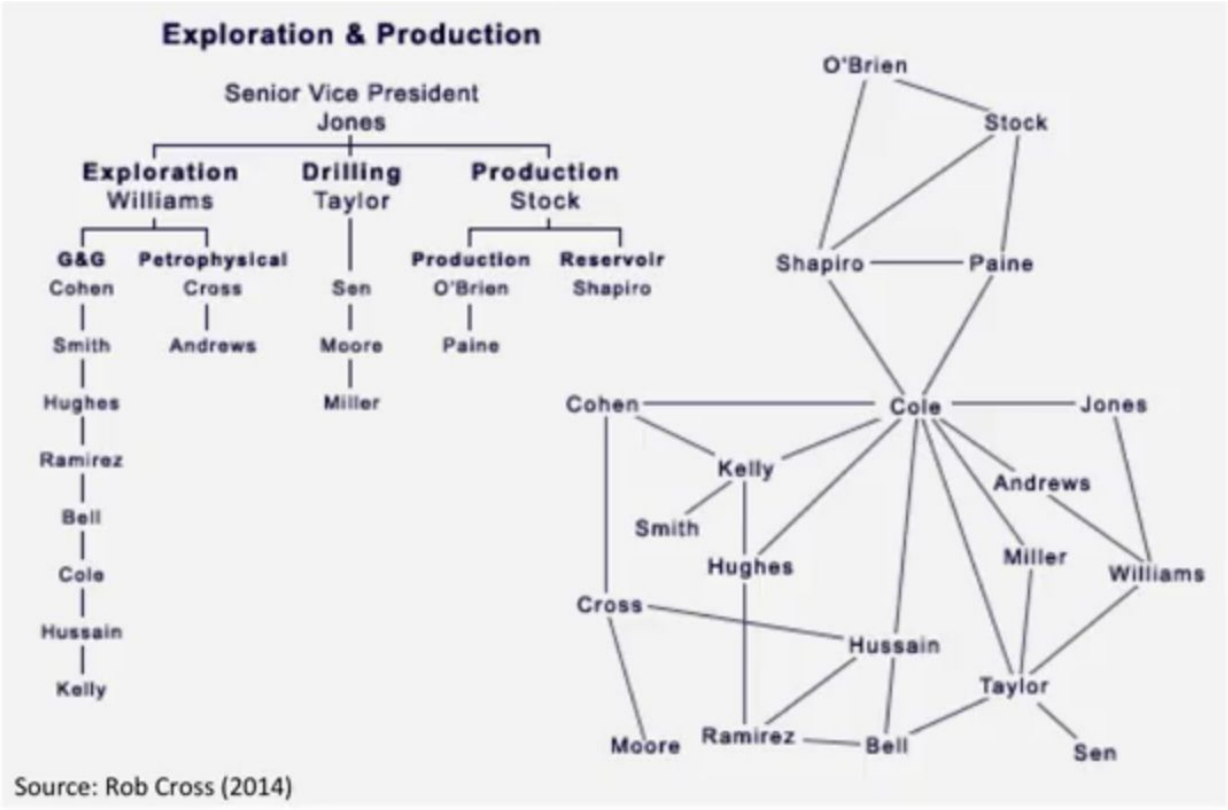
Main question
How can we improve collaboration inside corporation?

Analyzing collaboration
How can we describe collaboration patterns between employees?
How can we map these collaboration patterns?
How can we evaluate these collaboration patterns?
How can we improve these collaboration patterns?

Organizational Network Analysis (ONA)




Describing collaboration networks
Organization chart




Types of organization networks
- Collaboration networks: information flows, knowledge sharing
- Communication networks

- Friendship networks
- Advice networks
- Trust networks

5 Building blocks
Network size Network strength, Network range, Network density, Network centrality


Two ways to collect network data
- Survey
- Other sources

Identify sample
- Sample boundaries
    - formal units
    - locations
    - communities
    - cohorts
    - teams, etc
- Sample size

N = 25-300

Create survey
- Opening statement
    - purpose
    - confidentiality
- network questions
- additional questions
- order & format
- Test & refine
- 10-15min max

Administer & monitor
- Cover note from senior sponsor
- Timing
- Incentives

- High response rate is critical

Clean & enter data
- Collate, clean, and enter
- Visualize and analyze data using customized software packages: UCINET, netdraw

Some issues about surveys
Pros
Customized, detailed information from target sample
Cons
- High response rates are critical
- Network cannot be too large
- Survey cannot be too long
- Questions must be worded and interpreted with care
- Confidentiality is critical
- Relatively costly method of data collection

Other sources
- Big data
    - interactions via email, phone calls, computer conferencing, bulletin boards, social media
- Archival records
    - corporate databases - e.g. info on shared project assignments, work histories, event
- Attendance
    - public databases - e.g. info on co-patenting, co-authorship ,co-citations
- Fieldwork
    - observations, diaries, electronic tags

Other sources
Pros
- information on larger networks may be available
- may be less invasive
- may be less expensive
- may provide more objective measures
Cons
- Privacy concerns
- What do available measures actually capture
- large datasets can generate statistically significant but unimportant findings

Evaluating collaboration networks
How do collaboration patterns vary?
Simple descriptive statistics
- Compare across individuals

- Compare changes over time

Implications for managing employees
- Performance assessment
- Roles & responsibilities
- Pay & promotions
- Training & mentoring

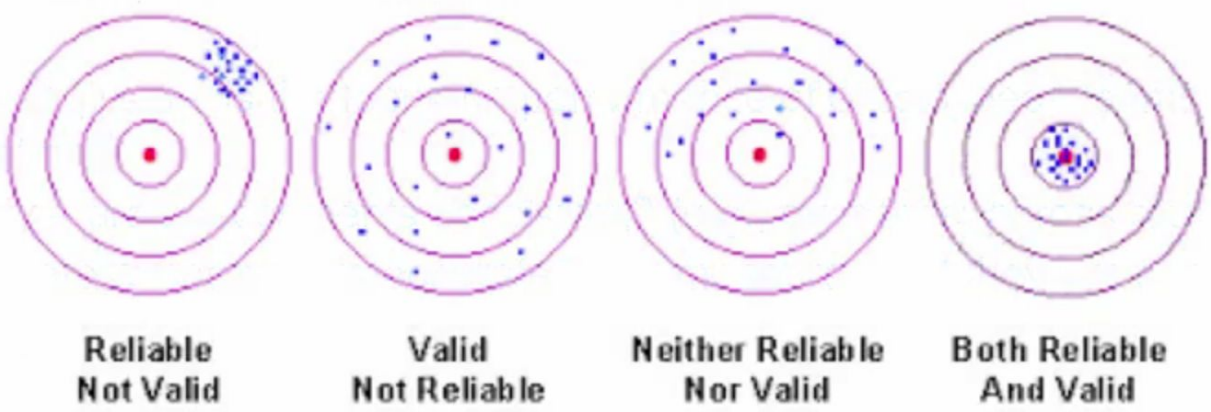
How do collaboration patterns matter for important outcomes?
Individual outcomes
- Performance
- Satisfaction
- Commitment
- Burnout
- Turnover etc

Correlation & multivariate analysis
- identify relationships between network variables and outcomes

Implications for managing employees
- Performance assessment
- Roles & responsibilities
- Pay & promotion
- Training & mentoring
- Job rotations & career development
- Retention

NOTE
There is no one "best" collaboration network for every organization in every situation!
To understand what's best for your particular organization in your particular situation, you will need to collect and analyze the data.

Measuring outcome
Useful input is necessary!
What is a strong measure of performance?
- Level of analysis: the focus on performance of employees? teams? organizations?
- Reliability: are assessments consistent? e.g. over time, across raters
- Validity: are assessments accurate? i.e. measure what they are supposed to measure




- Comparability - consistently measured and meaningful for all units in the dataset
- comprehensiveness - available for all or most units in the dataset
- Cost effectiveness - not too expensive to collect
- Causality - defensible as an outcome variable

Individual outcomes:
- Performance
sales per quarter? cost saving? self-reported 1-3 ratings? manager-reported 1-3 ratings? Bonus? etc

How can we improve collaboration patterns?
Is more collaboration needed?
- More is not always better
Where is more collaboration needed?
- Build ties strategically
How to increase collaboration?
- Provide motivation to build ties
    - Emphasize & promote collaboration
    - Recognition & reward collaboration
- Provide opportunities to build ties
    - Cross-functional meetings, conference calls, job rotations, site visits, events, etc

Where is more collaboration needs?
- Build ties strategically
How to increase collaboration?
- Provide motivation to build ties
    - emphasize & promote collaboration
    - recognize & reward collaboration
- Provide opportunities to build ties
    - cross-functional meetings, conference calls, job rotations, site visits, events, etc

Intervening in collaboration networks: five examples
- Reducing employee overload by rebalancing collaboration demands
Example: in a financial services organization, a network analysis shows that 5% of people accounted for up to 35% of the value-added collaborations; these valuable people often felt very overloaded.
Solution: reduce the overload of these people
- Improving resiliency of global teams by connecting peripheral members
Problem:
A company found that its global IT teams often relied on only a few key people to connect their members across the world; if a few key people left, these teams were vulnerable to breakdown.
solution: identify a small number of new connections that would have the biggest positive impact on team connectivity, and shift responsibilities more evenly across the members.
- Reducing collaboration inefficiencies by through targeted coaching
Problem: A major utility company asked employees how much time they spent interacting with each other and how useful those interactions were, the analysis shows some employees who were very highly regarded, but also a small number of employees who were much less effective than the rest.
Solution: focused personalized coaching efforts on collaborative issues unique to each of the low performers.
- Eliminating organizational silos by building cross-divisional ties
Problem: A fortune 500 conglomerate had grown by acquisition, but analysis of collaboration among the top 126 executives revealed that some divisions were much less integrated than others.
Solution: identify and target network connections that hold most strategic relevance for the firm, and track changes to these ties over time to assess the impact of interventions.
- Enhancing career paths through better performance management processes
Problem: a global consulting firm mapped the network of about 80 partners, and found two types of collaboration that were very valuable for the firm but not recognized at all in its performance management processes, which focused on individual revenue production:
    - collaborating to win clients
    - collaborating to serve clients

solution: revise performance evaluation systems to recognize contributions of partners who help others to win new clients or serve current clients

# Talent management

- Many different perspectives / definitions
- Talent assessment and development
    - identify differences in ability
    - develop so that everyone's ability is maximized
- Deeper than performance evaluation
    - essentially: employee evaluation

Motivation case: promotion

Challenges:
- Data is good. A lot of data is, typically, better. But they can also be misleading.
- If you are doing talent analytics you will be crunching all kinds of numbers - performance evaluation,s test scores, 360 feedbacks, sales figures, employee morale etc.
- Before you draw inferences from these numbers, it is critical to navigate a few challenges

challenges:
context, interdependence, self-fulfilling prophecies, reverse causality

Context
- We tend to neglect context when evaluating performance
    - Over-attribute performance to personal traits (personality, skill, etc.)
    - under-attribute performance to the situation the person was in (easy vs difficult task, helpful vs hurtful colleagues, favorable vs unfavorable economy, etc)
- The "fundamental attribution error"
There is a saying on Wall street: Don't confuse brains and a bull market.
- When using data to compare employees you must find ways to put them on an even playing field. "Apples to apples"
- Think of performance relative to expectations, as driven by team, product, industry, economy, boss, etc.

Interdependence
- A humbling amount of our work depends on other people
e.g. hiring a start investment analyst resembles an organ transplant. The analyst usually under-perform, and the intaking group suffers too.
- Means performance evaluation is often best done at the group level.
- Reliable individual evaluation typically requires seeing them with multiple teams.
- New and improved performance measures designed to assess contribution to team performance are on the way.
e.g. network analysis

Self-fulfilling prophecies
- People tend toward performing consistent with expectations. High expectations increase performance, low expectation decreases.
    - can occur because we treat them differently as a result of our own expectations.
    - can also occur because our expectations literally change their behavior
- "The matthew effect". The rich get richer and the poor get poorer.
Early advantages often accumulate
    - Consumer goods
    - Education

    - Careers
- where experience and recognition matter, thse with early advantage will be increasingly privileged over time.

Reverse causality
- When we see two correlated factors, we tend to believe one caused the other. Especially when there is an intuitive direction.
E.g. are charismatic leaders more successful?
A study shows that charismatic CEOs did not have more future success, but sucessful CEOs were perceived as more charismatic.
- We are driven to make sense of the world we live in, so we build causal stories from what we observe.
- But this leads us to see things that don't exist, and this can lead to giving people credit, or blame, they don't deserve.
"The normative lesson is.. that the stories we tell each other about success and failure in top management, like the stories we tell about success and failure in gambling are in large part fictions intended to reassure us about justice and encourage the young."

Tests and algorithms
Times: How high is your XQs?
NY Times: Can an algorithm hire better than a human?

Pros:
- Processing efficiency
- Broader search
- Unbiased
Cons
- Hyper-focused
- Low explanatory power

- Prescriptions
    - Do the science
    - Rigorous testing in the relevant setting
    - Provide human oversight
    - Program, test, error-check
    - Use multiple tools
    - Draw on as many diverse signals as possible

Prescription: navigation the challenges of talent analytics
1. Broaden sample
- Allow good performance evaluation, expand the sources/signals
    - Additional opinions
    - Additional performance metrics

    - Additional projects, assignments
    - remember: from maximally diverse sources
- Second chances. And thirds
    - e.g. a new employee's boss has a huge impact, but is completely outside his/her control

2. Find/create exogenous variation
- The only truly valid way to tease out causation is to manipulate an employee's environment
- Trade-off: you still have to run a business
- But can and should be willing to trade off a bit of operational efficiency for greater insight into the abilities of the employees.
- This is major motivation for rotational programs
    - Wide variety of environments, people
    - Pre-committed to changing at preset times, often in preset patterns.
    - Almost an experiment, which is the gold standard
- Lesser versions: change teams, direct reports, projects, offices.

3. Reward in proportion to signal
- Match the duration and complexity of rewards to the duration and complexity of past accomplishments
- For short, noisy signals, better to give bonuses rather than raises, praise than promotions
    - Note 1: Most signals are noisy, and we are prone to underestimate the noise. Suggest we typically over-react in talent management
    - Note 2: Of course you also have to retain people, so must factor in external labor market. But accurately valuing employees should be an advantage.
- Drawing major distinctions, and granting major rewards, should only follow major signals.
    - e.g. consulting/law firm partnerships typically involve a multi-year, up-or-out partnership track
    - e.g. academic tenure is practically irrevocable, so is granted to relatively few and only after 5-10+ years of performance.

4. Emphasize development
- talent analytics is not all about selection
- even in a field as selection-oriented as venture capital, firms spend considerable resources developing people within their portfolio firms.
- Testing and assessment is at least as valuable as development tools as selection tools. And more palatable.

5. ask the critical questions
- Are we comparing "apples to apples"? i.e. have we sufficiently adjusted for context?
- What impact have other people had on this person's work? How interdependent are these measures?
- How have expectations colored our evaluations? To what extent have successes and failures been influenced by the way we've trated people, the situations we've put them in?
- Are the factors we believe lead to success (adn failure) truly causal?

Major conclusions

Organization challenges

- Claims: effective people analytics is more of an organizational challenge than an analytics challenge
- From this flow many prescriptions, but one dominant theme: no black boxes
- major theme: no black blocks
- Specific prescriptions:
    - Be transparent:
    - embed yourself: analytics in t-shirt for NBA, build trust
    - share control: algorithm aversion