

Analytics, optimization, etc are all about using quantitative methods to understand PEOPLE, and why we do the stupid stuff we do Today we're going to talk about some classic studies, and why they matter (and how they help us understand human behaviour)





Stephen Few.



"A dashboard is a

#### visual display

of the most important information

needed to achieve one or more objectives;

consolidated and arranged

## on a single screen

so the information can be monitored

at a glance."

[Stephen Few]

Stephen Few defined a dashboard as...





"A dashboard is a

visual display

of the **most important information** 

needed to achieve one or more objectives;

consolidated and arranged

## on a single screen

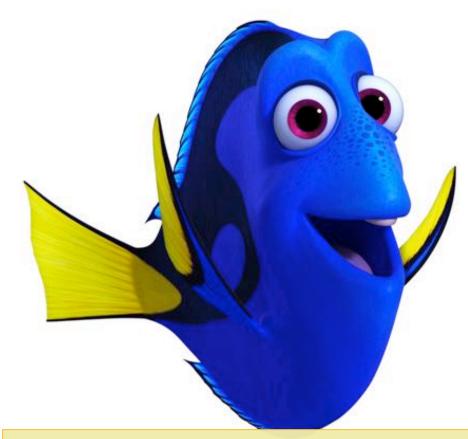
so the information can be monitored

at a glance."

[Stephen Few]

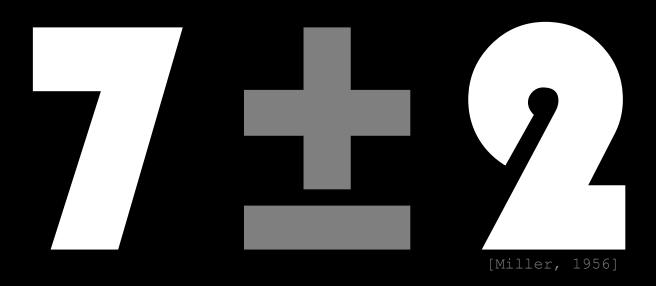
To me, the most important words are "on a single screen"





The reason why we need the information to be on one page or screen? The limits of our memory.





In 1956, George A. Miller published one of the most cited papers in psychology, where he spoke of the magic number 7. For our purposes here, Miller found that the amount of information that we can retain our working (or short term) memory is seven, plus or minus two.

# RXUWBJP

However, this doesn't mean that we can only recall 7 +/- 2 individual letters, or numbers (for example.) The 7 +/- 2 items can be 7 +/- 2 letters, or 7 +/- 2 words, or ...



Take my hand / Take my whole life too / For I can't help / Falling in love with you

You may say I'm a dreamer / But I'm not the only one

We are young / Heartache to heartache we stand

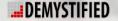
I fell into a burning ring of fire / I went down down down and the flames went higher

All you need is love / Love is all you need

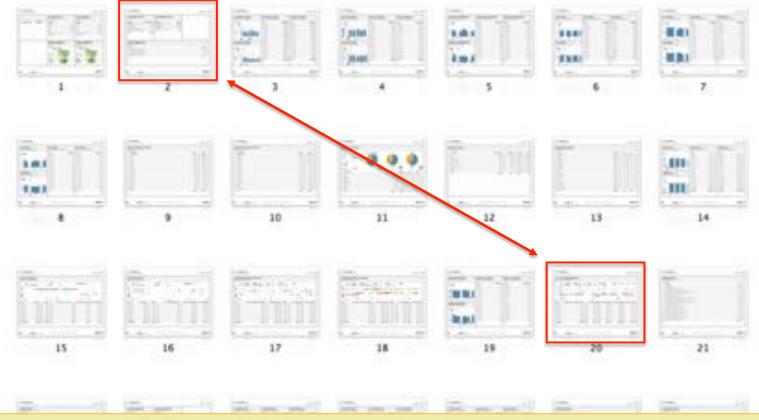
Every new beginning / Comes from some other beginning's end.

... or 7 + / - 2 song lyrics.

We are capable of "chunking" information, so that each piece of information only occupies one "item" within our memory system.







Because, back to Stephen Few's definition, if you are presenting information, you need to work WITHIN the limits of our perceptual and memory systems. Expecting users to draw connections between two data points six pages apart, or thirty slides ago, is a recipe for failure.





# "A man with a conviction is a hard man to change.

Show him facts or figures and

### he questions your sources.

Appeal to logic and he fails to see your point.



"Suppose that he is presented with evidence,

### unequivocal and undeniable evidence,

that his belief is wrong: what will happen?



"The individual will frequently emerge, not only unshaken, but

# even more convinced of the truth of his beliefs

than ever before."





This might happen especially if:

- Beliefs are deeply held;
- Action has been taken, that is difficult to undo; (for example, the Doomsday cult had sold all their possessions)
- Believer has social support (the cult continues to reinforce each others' beliefs)



# cognitive dissonance

Festinger explains this by the theory of cognitive dissonance. This theory tells us that we don't like the feeling of inconsistency (for example, of our beliefs and our actions.) We seek to reduce this uncomfortable feeling by justifying our beliefs, and avoiding information that might further conflict. For example, smokers who know smoking is bad for them may choose to tell themselves, "It's not so bad, it's not as bad as crack cocaine." We will change our beliefs, to match our behaviour!

In the case of the Doomsday cult, the idea that they had sold all their possessions and believed something so asinine is an uncomfortable disconnect, so they choose to believe they were "spared", which justifies their actions.





It also explains why you can never convince your racist uncle of anything...



However, Chanel, Luchini, Massoni, Vergnaud (2010) found that if we are given an opportunity to discuss the evidence and exchange arguments with someone (rather than just reading the evidence and pondering it alone) we are more likely to change our minds, in the face of opposing facts. (So... maybe there's some value in talking to your racist uncle.)





Why this matters for analysts: Even if your data seems self-evident, if you come in with "breaking news", that goes against what the business has known, thought, or believed for some time, you may need more data to support your contrary viewpoint. You may also want to allow for plenty of time for discussion, rather than simply sending out your findings, as those discussions are critical to getting buy-in for this new viewpoint.



We know now that "the facts" may not persuade us, even when brought to our attention. However, Confirmation Bias tells us that we intentionally seek out information that continually reinforces our beliefs, rather than searching for all evidence and fully evaluating the possible explanations.



In a 1967 study by Brock & Balloun, subjects listened to several messages, but the recording was staticky. However, the subjects could press a button to clear up the static. They found that people selectively chose to listen to the message that affirmed their existing beliefs. For example, smokers chose to listen more closely when the content disputed a smoking-cancer link.

Wason (1960) conducted a study where participants were presented with a math problem: find the pattern in a series of numbers, such as "2-4-6." Participants could create three subsequent sets of numbers to "test" their theory, and the researcher would confirm whether these sets followed the pattern or not.



"Even Numbers"

Rather than collecting a list of possible patterns, and using their three "guesses" to prove or disprove each possible pattern, Wason found that participants would come up with a single hypothesis, then seek to prove it. (For example, they might hypothesize that "the pattern is even numbers" ...

and check whether "8-10-12", "6-8-10" and "20-30-40" correctly matched the pattern.)



Increasing umbers!

When it was confirmed their guesses matched the pattern, they simply stopped. However, the actual pattern was "increasing numbers" – their hypothesis was not correct at all!



Do you do this with your analysis? When you start analyzing data, where do you start? With a hunch, that you seek to prove, then stop your analysis there? (For example, "I think our website traffic is down because our paid search spend decreased.") Or with multiple hypotheses, which you seek to evaluate one by one?

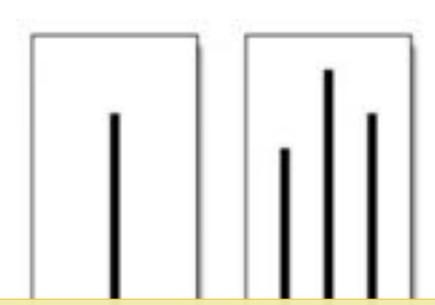
### **Analysis of Competing Hypotheses**



An alternative – Analysis of Competing Hypotheses. Analysis of Competing Hypotheses (or "ACH" was developed in the 1970's, for use by intelligence analysis. It is an attempt to overcome the natural biases that we all have. The idea is, you come up with multiple hypotheses that could explain the data you are seeing, and seek to evaluate each one in turn. Which ones are more or less likely to be true? What data do you have, and which hypotheses does it support? This attempts to avoid our analysis simply being a self-fulfilling prophecy.







In 1951, Asch found that we conform to the views of others, even when they are flat-out wrong, surprisingly often!

He conducted an experiment where participants were seated in a group of eight others who were "in" on the experiment ("confederates.") Participants were asked to judge whether a line was most similar in length to three other lines. The task was not particularly "grey area" – there was an obvious right and wrong answer.

Each person in the group gave their answer verbally, in turn. The confederates were instructed to give the incorrect answer, and the participant was the sixth of the group to answer.



# 76%

gave the wrong answer at least once!



# 5%

<u>always</u> conformed to the incorrect answer



### Only **25%**

never followed the group's incorrect answer







As mentioned previously, if new findings contradict existing beliefs, it may take MORE than just presenting new data.

# challenge the status quo

However, these conformity studies suggest that efforts to do so may be further hampered if you are presenting information to a group. It is less likely that people will stand up for your new findings against the norm of the group. In this case, you may be better to discuss your findings slowly to individuals, and avoid putting people on the spot to agree/disagree within a group setting.



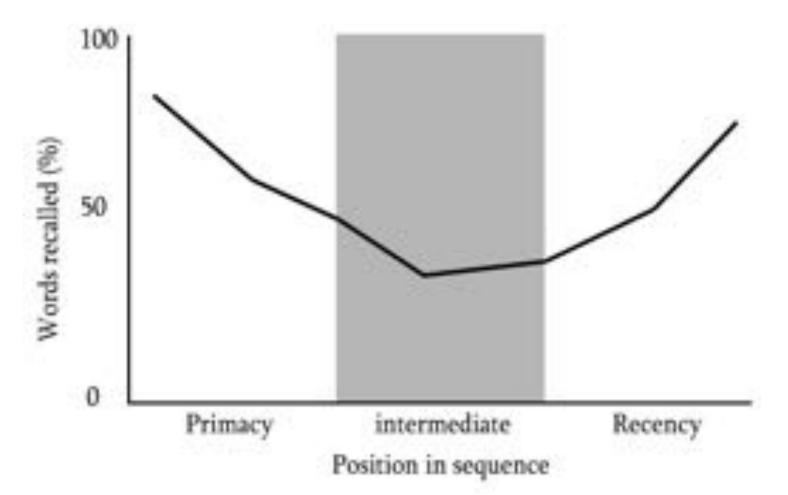
Similarly, this argues against jumping straight to a "group brainstorming" session. Once in a group, Asch demonstrated that 76% of us will agree with the group (even if they're wrong!) so we stand the best chance of getting more varied ideas and minimising "group think" by allowing for individual, uninhibited brainstorming and collection of all ideas first.



#### PRIMACY AND RECENCY EFFECTS

The serial position effect (so named by Ebbinghaus in 1913) finds that we are most likely to recall the FIRST and LAST items in a list, and least likely to recall those in the middle.















For example, let's say you are asked to recall pear, orange, banana, apple and pineapple.



The serial position effect suggests that individuals are more likely to remember pear (the first item; primacy effect) and pineapple (the final item; recency effect) and less likely to remember the items in the middle.





Primacy
Long term memory

The theory here is that the last item is retained in short term memory ...













Recency
Short term memory
Remembered in context

... while the first item is more likely to have been processed through to long term memory. Also, the last item is being remembered "in context" – the situation in which they're recalling it is CLOSER to when they learned the last item, than the first.





In 1975, Godden & Baddeley conducted experiments where they had scuba divers memorize information, and found their recall was better when the context in which they were recalling it was the same as when they were remembering it. (For example, they recalled better when they memorized underwater, and recalled underwater, than when they tried to recall on land.)

(This is also why we can't remember what the heck we were thinking until we get up, and go back to the spot we first had the thought. Then bam! It comes to us.)





The longer the list, the less likely the primacy effect is to apply.





The recency effect can be reduced, by requiring attention be switched to another task, and then return to the recall. (That way, the information was forced out of the short term memory.)





#### Steve

"is smart, diligent, critical, impulsive, and jealous"



Asch (1964) found participants told "Steve is smart, diligent, critical, impulsive, and jealous" had a positive evaluation of Steve,



#### Steve

"is jealous, impulsive, critical, diligent, and smart"



Whereas participants told "Steve is jealous, impulsive, critical, diligent, and smart" had a negative evaluation of Steve. Even though the adjectives are the exact same – only the order is different!



This is related to the recency bias – the idea that we give more weight to more recent information. For example... when shopping for a car: 'I want the sporty convertible. Last winter wasn't "that bad" (And you quickly forget the ten winters before that, where you needed a four wheel drive to get out of our blacked-out neighborhood, due to terrible snowstorms.)



Why does this matter for analytics?

When you present information, your audience is unlikely to remember everything you tell them. So choose wisely. What do you lead with? What do you end with? And what do you prioritize lower, and save for the middle?

These findings may also affect the amount of information you provide at one time, and the cadence with which you do so. If you want more retained, you may wish to present smaller amounts of data more slowly, rather than rapid-firing with constant information. For example, rather than presenting twelve different "optimisation opportunities" at once, focusing on one may increase the likelihood that action is taken.

This is also an excellent argument against a 50-slide PowerPoint presentation – while you may have mentioned something in it, if it was 22 slides ago, the chance of your audience remembering are slim.

#### why matters





#### Primacy



#### **Novelty**

Related to optimization effects:

Primacy: The idea that people are familiar with the old experience, so there's an advantage to the control over the new variation. Novelty: The idea that the new experience is something cool and different, so there is a (short term) life, that won't sustain over time.





However, sometimes the idea of primacy and recency can be taken too far, and be used to explain results that are simply not driven by them. A/B test results often show dramatic differences in the early days, that seldom continue on. (This is due to the high variance and lower sample size at the start of the test.)

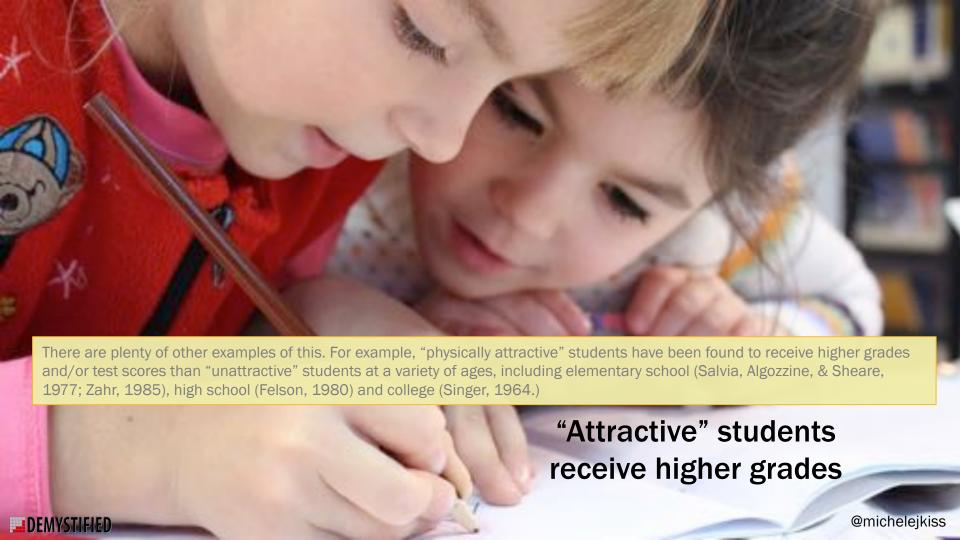
However, if the product team has put their sweat and soul in to a feature, it's easy to ascribe an actual reason to the early results, outside of statistical fluctuation. For example, "Oh, people are just used to the old design, and the variation will perform AMAZINGLY once they get used to it." However, it's rare for these early effects to completely flip the other way. So, while the low performance may not continue, it's also unlikely that the effects will completely reverse.

I've even seen the "primacy effect" be argued as the reason for the control to perform better (after a reasonable amount of time) in sites that have little to no return visitation!





In 1977, Nisbet and Wilson conducted an experiment with university students. The two students watched a video of the same lecturer deliver the same material, but one group saw a warm and friendly "version" of the lecturer, while the other saw the lecturer present in a cold and distant way. The group who saw the friendly version rated the lecturer as more attractive and likeable.





Thorndike (1920) found similar effects within the military, where a perception of a subordinate's intelligence tended to lead to a perception of other positive characteristics such as loyalty or bravery.

The Halo Effect tells us that we extrapolate from one positive effect to another. For example, because someone is intelligent, they are also loyal and brave. Because someone is attractive, they are also smart.

**DEMYSTIFIED** [Thorndike, 1920] @michelejkiss



Why this matters for analysts: The appearance of your reports/dashboards/analyses, the way you present to a group, your presentation style, even your appearance may affect how others judge your credibility and intelligence.



The Halo Effect can also influence the data you are analysing! It is common with surveys (especially in the case of lengthy surveys) that happy customers will simply respond "10/10" for everything, and unhappy customers will rate "1/10" for everything – even if parts of the experience differed from their overall perception.

For example, if a customer had a poor shipping experience, they may extend that negative feeling about the interaction with the brand to all aspects of the interaction – even if only the last part was bad! (And note here: There's a definite interplay between the Halo Effect and the Recency Effect!)





# "37 WHO SAW MURDER DIDN'T CALL THE POLICE"

- The New York Times

March 27, 1964

THE BYSTANDER EFFECT





In 1964 in New York City, a woman name Kitty Genovese was murdered. According to papers, the attack had lasted an hour, and almost forty people witnessed the attack. Yet, no one called the police. (Later reports suggested this was an exaggeration – that there had been fewer witnesses, and that some had, in fact, called the police.)

However, this event fascinated psychologists, and triggered several experiments. This became known as the "Bystander Effect", which proposes that the more bystanders that are present, the less likely it is that an individual will step in and help.



Darley & Latane (1968) manufactured a medical emergency, where one participant was allegedly having an epileptic seizure, and measured how long it took for participants to help. They found that the more participants, the longer it took to respond to the emergency.

(This is actually why, if you go to CPR training, the FIRST thing you're instructed to do is to point to a specific person, and tell them "YOU! Go call 911. YOU! Go get me the AED machine." Because, if you leave it up to an anonymous group, everyone may assume someone else will call, and no one calls for help!)

## Diffusion of Responsibility

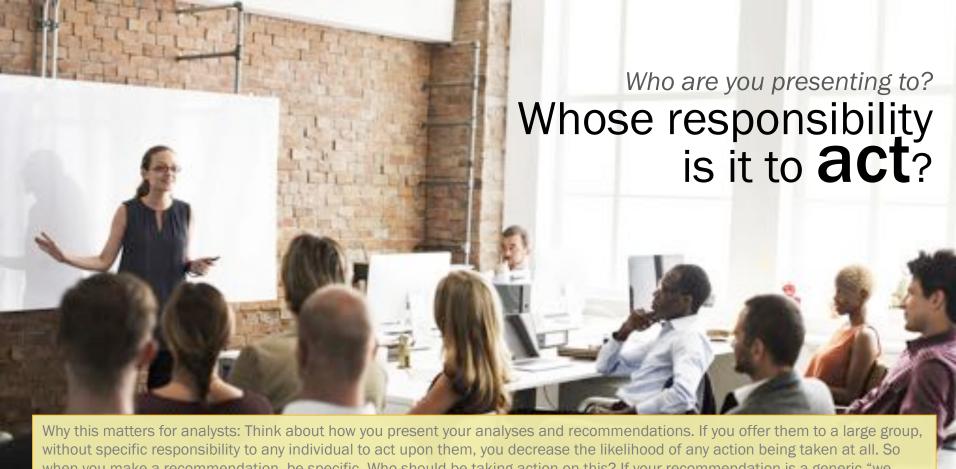




Think about how hard an INDIVIDUAL works in a tug of war with 20 people on one side, versus 2.







when you make a recommendation, be specific. Who should be taking action on this? If your recommendation is a generic "we should do X", it's far less likely to happen.



#### **SELECTIVE ATTENTION**



In 1999, Simons and Chabris conducted an experiment in awareness at Harvard University. Participants were asked to watch a video of basketball players, where one team was wearing white shirts, and the other team was wearing black shirts. In the video, the white team and black team respectively were passing the ball to each other. Participants were asked to count the number of passes between players of the white team.





Did you see it??

During the video, a man dressed as a gorilla walked into the middle of the court, faced the camera and thumps his chest, then leaves (spending a total of 9 seconds on the screen.) Amazingly? Half of the participants missed the gorilla entirely! Since then, this has been termed "the Invisible Gorilla" experiment.

Participants were SO FOCUSED on counting the passes of the white team that they completely missed it. They were selectively attending to some information in the video, at the exclusion of other information.



# Don't get so bogged down in the details you miss critical points

Why this matters for analysts: As you are analyzing data, there can be huge, gaping issues that you may not even notice. When we get SO FOCUSED on a particular task (for example, counting passes by the white-shirt players only, or analyzing one subset of our customers) we may overlook something significant at a higher level.

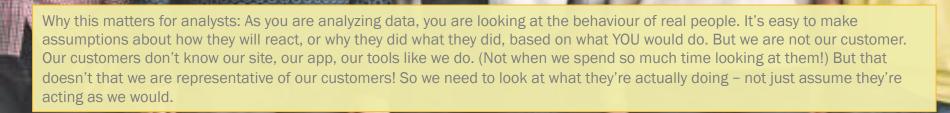
Take time before you finalize or present your analysis to think of what other possible explanations, or variables there could be (what could you be missing?) — or invite a colleague to poke holes in your work.

#### **FALSE CONSENSUS**



Experiments have revealed that we tend to believe in a false consensus: that others would respond similarly to the way that we would. For example, Ross, Greene & House (1977) provided participants with a scenario, with two different possible ways of responding. Participants were asked to explain which option THEY would choose, and guess what OTHER PEOPLE would choose. Regardless of which option they actually chose, participants believed that other people would choose the same one.





# you are not your customer

**E**DEMYSTIFIED

@michelejkiss



There is a related effect here: the Homogeneity of the Outgroup. (Quattrone & Jones, 1980.)

In short, we tend to view those who are different to us (the "outgroup") as all being very similar, while those who are like us (the "ingroup") are more diverse.

For example, all women are chatty, but some men are talkative, some are quiet, some are stoic, some are more emotional, some are cautious, others are more risky... etc.

We do this a lot in marketing! "All millenials are spoiled, entitled little brats. All Baby Boomers are old fuddies who don't know how to use a computer."

#### HOMOGENEITY OF THE OUTGROUP





Why this matters for analysts: Similar to the False Consensus Effect, where we may analyse user behaviour assuming everyone thinks as we do, the Homogeneity of the Outgroup suggests that we may OVERSIMPLIFY the behaviour of customers who are different to us, and fail to fully appreciate the nuance of varied behaviour.

This may seriously bias our analyses!

For example, if we are a large global company, an analysis of customers in another region may be seriously flawed if we are assuming customers in the region are "all the same." To overcome this tendency, we might consider leveraging local teams or local analysts to conduct or vet such analyses. Or, add qualitative data to give context (and maybe an opposing viewpoint, to our assumptions!)

# "everyone in [countryX] behaves like Y"

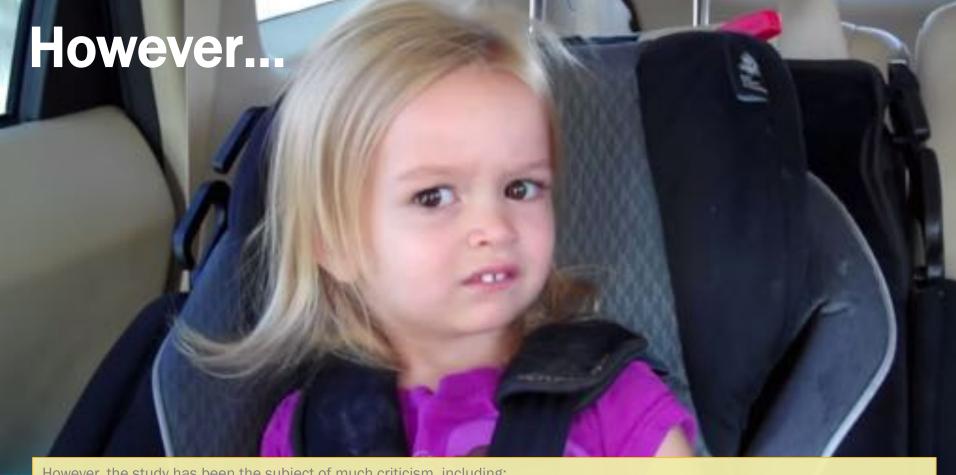


#### THE HAWTHORNE EFFECT



The story of the Hawthorne Effect goes as follows - In 1955, Henry Landsberger analyzed several studies conducted between 1924 and 1932 at the Hawthorne Works factory outside Chicago. These studies originally intended to discover whether the level of light (among other changes) within a building changed the productivity of workers. They found that the changes did have an impact, however they were short-term improvements, AND they seemed to happen due to the fact that a change was made – not the result of the actual change. Aka, it wasn't the fact that the light was brighter, but the fact that the light had been changed.





However, the study has been the subject of much criticism, including:

### "Glorified anecdote"

Once you've got the anecdote, you can throw away the data.

Dr. Richard Nisbett







One possibility is that the effects were due to "demand characteristics", and the fact that the employees knew they were being observed. (Martin Orne) (This can go either way – participants may be looking to "please" the researchers, or "correctly" complete the study, or they may seek to sabotage studies)

### **Demand Characteristics**





But in 2011, Levitt & List conducted a re-analysis of the original data. They actually had microfilm, so they were able to go right to the source. What they found... all of the changes were made on a SUNDAY, as it was the only day the factor was closed. So by...









People may not agree

People may not agree with your conclusions. USE THIS to produce a more comprehensive analysis. Take those viewpoints in to account. Respond to them, disprove those hypotheses.





Demand characteristics suggest that observing people may change their behaviour. We need to be aware of this possibility when analyzing behaviour. For example, what data would you get from an anonymous survey, versus one that requires a person's name and phone number? (Or their employee ID!) How do user testing and a/b testing results differ, based on people's level of awareness that they are being observed?

## Observation may change behaviour



#### Be aware of your biases



Be aware of your biases, and attempt to control for them.

- Are you just using the data to "prove" one particular hypothesis, rather than looking for all the possible explanations?
- Are you overgeneralizing to a certain population?
- Is your interpretation of the data being influenced by conforming to your company's norms, and the way things have always been done?







Get someone to keep your biases in check! Have someone else to vet your assumptions. (Especially someone unfamiliar/who hasn't been involved with your analysis or experiment so far.)







#### **Questions?**



@MicheleJKiss

michele@analyticsdemystified.com





January 2019 **Discount Code: DAASF** 

10% off Conference or Training

