Why Good Engineers Give Bad Cost Estimates:

Results of Psychological Research

Jordan Garner
UC Davis (JPL Summer Hire)

Arthur B. Chmielewski
Jet Propulsion Laboratory
California Institute of Technology

February 22, 2012

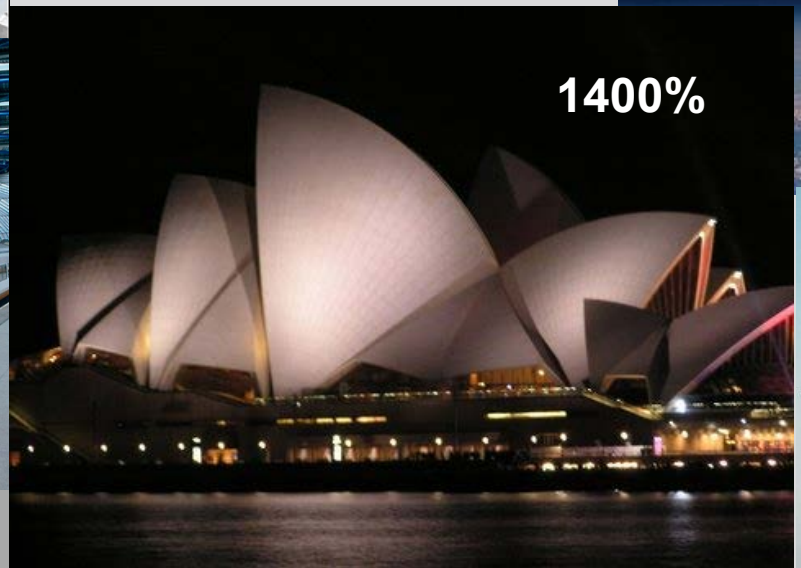
Special Thanks to

- Dr. David Ullman of Robust Decisions and Oregon State University for his assistance with the web experiment and continued support of this novel research.

- Prof. Don Forsyth of the University of Richmond for his expert consultation on socio-psychological effects in decision making.

- Joseph Mrozinski of JPL for his expertise in statistics and mathematical data interpretation.

Great Engineering - Bad Estimates



Bad Estimates - as Old as Human History

Pharaoh Snefru's engineers were some of the first humans in the recorded history to register a bad estimate.

Their ruler's tomb had 600% schedule overrun.

Medium Pyramid collapsed

Bent Pyramid cracked open

He died before his third (Red Pyramid) was finished

Bad Costing – Universal Problem

- Bad estimates are everywhere:
  - Construction
  - Transportation
  - Military
  - Film making
  - Family vacation to Hawaii
- Bad costing knows no borders or time period.

Causes of Overruns

- Commonly blamed for overruns:
  - Scope changes
  - Poor communication
  - Technical issues
  - New technology
  - Acts of god
- But are these the *main* culprits?

- **Thesis:** Bad cost estimating is universal because psychological effects critically distort human judgment.

- We conducted a simple experiment to measure the power of psychology on estimators.

Dishwashing Experiment

- Participants in the on-line experiment were asked to estimate the time needed to perform a simple task – washing the dishes shown below.
- Random participants were asked slightly different questions to test psychological effects.

- 5 psychological effects were tested:
  1. Anchoring Bias
  2. Q&A Mismatch
  3. Decomposition
  4. Reserve Fallacy
  5. Optimism

- 507 volunteers participated: 142 NASA employees, 305 college students and 60 other adults.
- ~2300 data points were collected to determine quantitative strength of effects.

Effect #1: Anchoring Bias

Anchoring Bias occurs when the question suggests a wrong answer.

"You have an allocation of $1M, what is your estimate?"

"I have a bogey of $400k for your system. Give me your own estimate by tomorrow."

1. Anchoring Results

- Unanchored question: “How long will it take to clean this kitchen?”
- **30 min**
- Anchored question: “George took 10 minutes. How long will it take to clean this kitchen?”
- **24 min**

- **Psychological effect:** Managers frequently anchor the initial estimate to the wrong value which is later very difficult to change.

Effect #2: Q&A Mismatch

Q&A Mismatch happens when the intention of the question is missed by the answer.

- For example, is the question “What is your cost estimate for this robotic arm?” a request for a mean, 90% confidence or something else?

2. Q&A Mismatch Results

- (A) "How long will it take you to clean the kitchen?"
  30 min
- (B) "How long will it take you 90% of the time?"
  39 min
- (C) "How long will it take you 50% of the time?"
  31 min

- **Psychological effect:** When managers ask a general question type (A), they implicitly expect a high confidence answer like (B), but usually receive a 50/50 answer like (C).

Effect #3: Decomposition

Decomposition is breaking the estimate into smaller parts which people feel will improve accuracy.

“Break down your robotic arm task into level 5 WBS elements and cost them by Monday.”

3. Decomposition Results

No decomposition: "How long will it take to clean this kitchen?"

30 min.

Decomposition: "Give estimates for cleaning the bowls, silverware, pots and pans, and add together."

31 min.

- **Effect:** Decomposition did not change the result
  - Decomposition was more time consuming than helpful.
  - Deep decompositions often miss the cost of interfaces (for example, walking to the drawer to put away the dishes.)

Effect #4: Reserve Fallacy

Reserve Fallacy describes the fact that people are comfortable with unfounded intuitive reserve levels.

"This proposal includes a robust 30% budget reserve."

4. Reserve Fallacy Results

(A) “How long will it take to clean this kitchen?”

30 min

(B) “I am 90% sure that the time is no more than plus or minus __ min.”

9 minutes = 30%

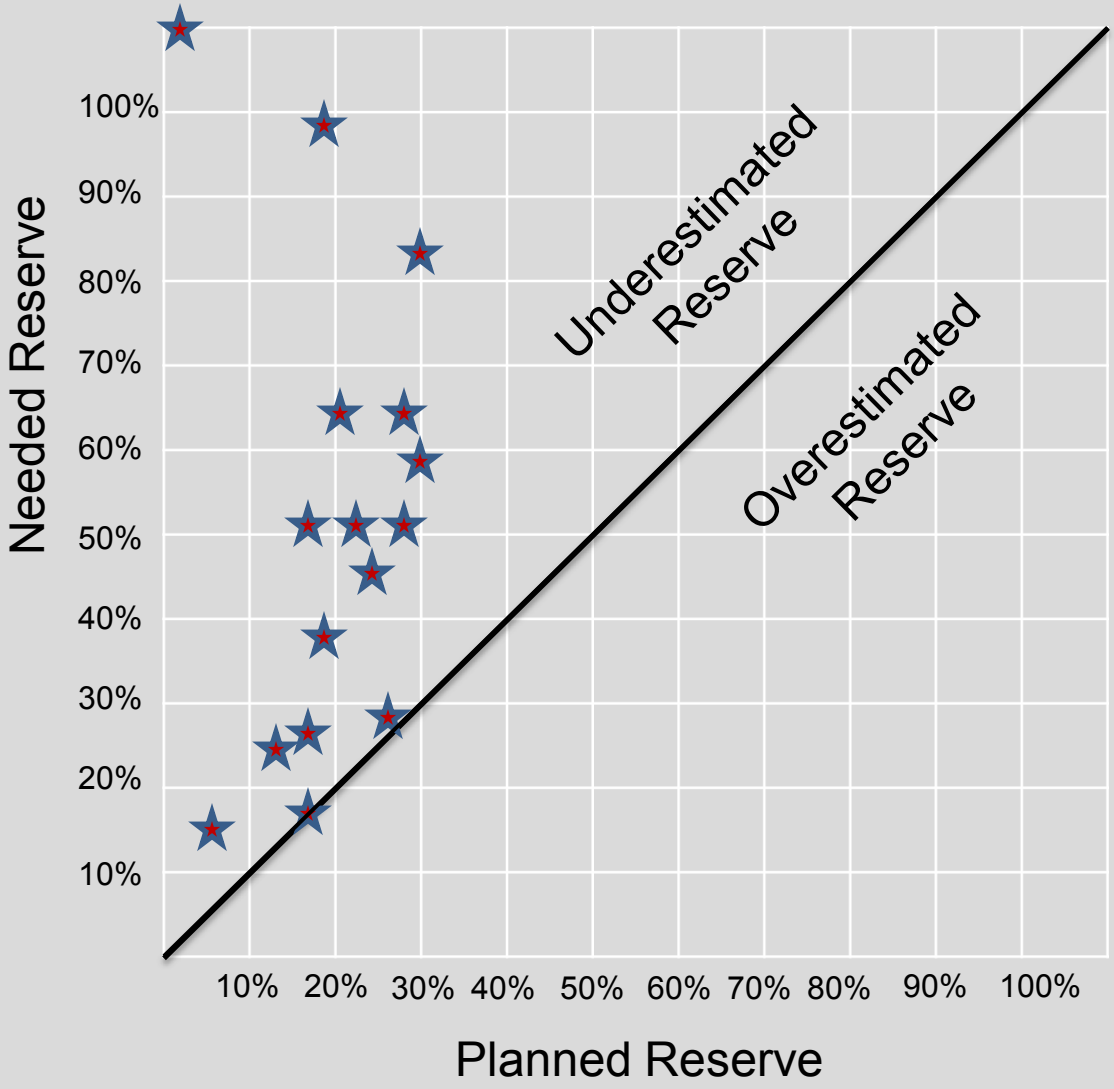
(C) “I am 99% sure that it will take me no more than__ minutes.”

45 min (15 minutes = 50%)

Psychological Effect:

People display high confidence in a 30% reserve. Such high confidence would have been more appropriate for a 50% reserve.

Large Projects Reserve Comparison



The average amount of budget reserve required for 18 large projects surveyed is 52%.

Effect #5: Optimism

People tend to be overly optimistic in planning.

“This time, there will be no delays.”

“All the risks have been mitigated by the last project.”

5. Optimism Results

How long will it take to clean the kitchen:

30 min

(A) “Best case scenario - nothing goes wrong.”

27 min

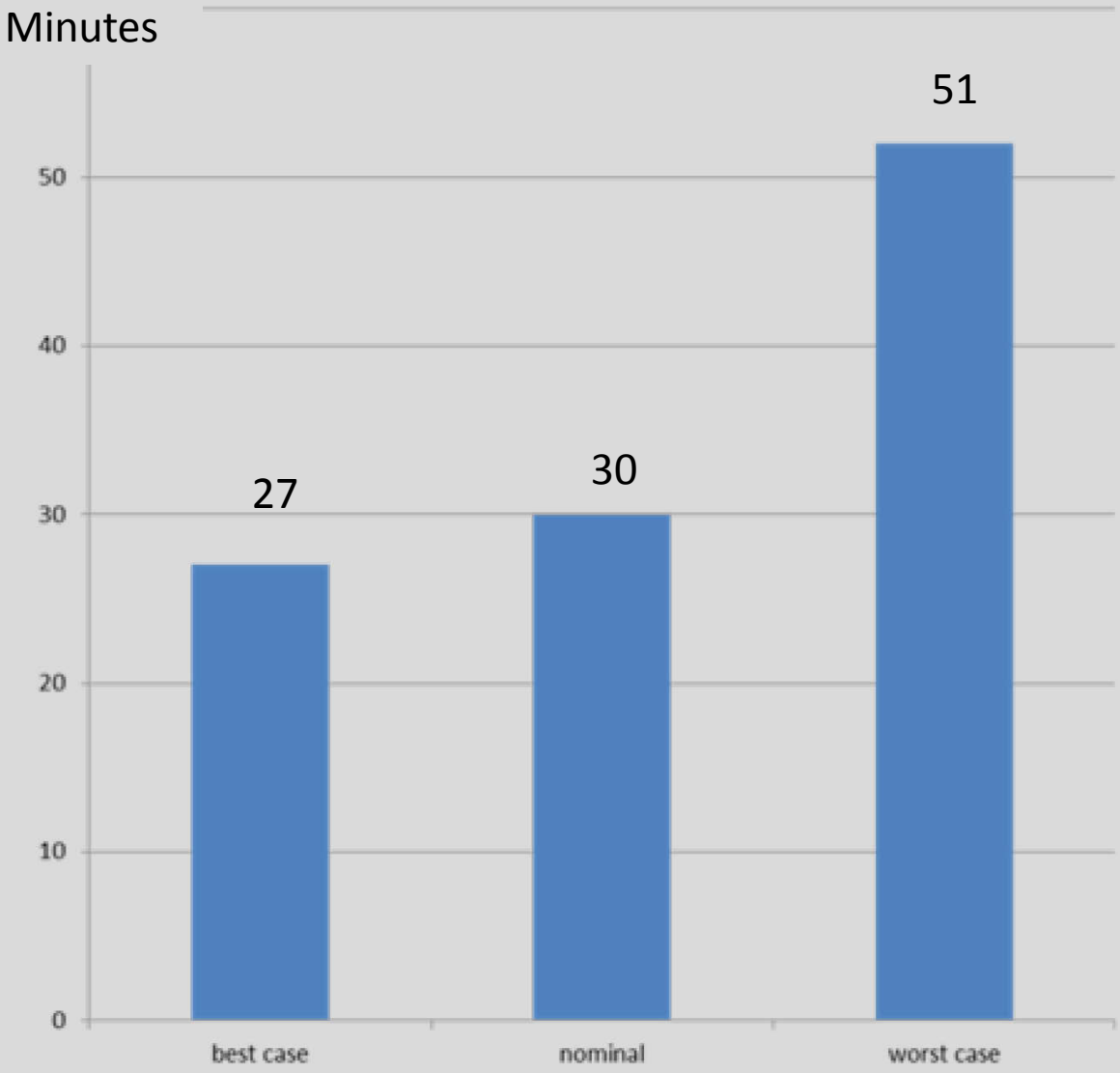
(B) “Worst case scenario - everything goes wrong.”

51 min

(C) “George took 40 minutes. How long will it take you”

30 min

5. Optimism Results



- The expected time to clean the kitchen was almost identical to the best case scenario.
- Optimistic anchoring worked very well, pessimistic anchoring failed.

Conclusions

- **Psychology distorts cost estimates**.
- To counter:
  - Train the managers **not to anchor.**
  - Use **Financial Language** not common language.
  - Deep **decompositions do not improve accuracy.**
  - **Calculate** the **reserve based on risk.**
  - Combat **optimism** by baselining historical and likely risks.