Eliciting Expert Opinion in Acquisition Cost and Schedule Estimating

Gregory E. Brown, Maj.
School of Systems and Logistics
Air Force Institute of Technology

Contact:
Gregory.Brown@AFIT.edu

Background

- Despite analytics push, expert opinion will likely remain relevant
  - Limited ACAT II/III data collection…databases focus on ACAT I data
  - New technologies and production methods
  - Parametric models often require subjective inputs (e.g. estimated lines of code)
- Why is this a problem?
  - Formal cost analyst training/education focuses on parametric modeling, underemphasizes elicitation techniques
  - No structured elicitation model in DoD:
    - "[E]licitation methodologies are largely ad hoc…seldom based on elicitation literature" (Galway, 2007)
  - Poor Documentation & Unstated Assumptions within DoD elicitations
    - "It is especially hard to go back to finished projects and get historical information about the elicitations that were done." (Galway, 2007)
  - Inconsistent adjustment for expert overconfidence

Research Questions

- Research Questions

Q1. Does research support the Joint Agency Cost Schedule Risk Uncertainty Handbook (JA CSRUH) heuristic of adding 30% additional uncertainty to the expert's range?

Q2. Does a formal elicitation model exist that could be adapted for DoD cost and schedule estimating?

Q1: Overconfidence Heuristic

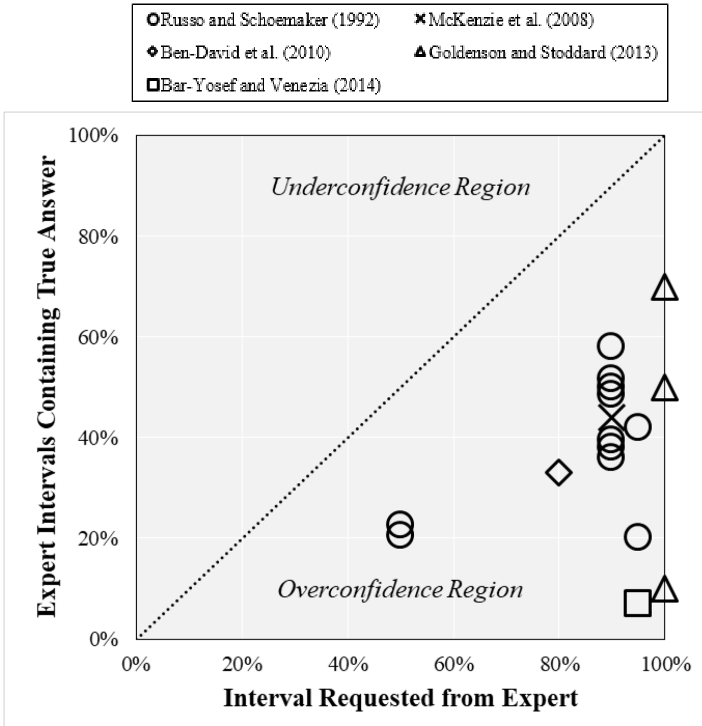
- Per JA CSRUH, add 30% additional uncertainty to expert's low/high bounds

- Heuristic based on Capen (1976); shortcomings:
  - A single study with 1,200 participants x 10 questions each: 12,000 observations
  - Research is 40+ years old
  - Capen quizzed petroleum engineers with encyclopedia questions

- Does contemporary research exist that supports/refutes Capen's finding? Inclusion criteria:
  - Must utilize experts with industry experience (<u>not</u> undergraduate college students)
  - Must involve business or engineering domain (closest equiv. to DoD acquisition)
  - Experts must estimate continuous values (<u>not</u> probabilities of event)

Q1: Overconfidence Heuristic

Author (Year)	Interval Requested	Intervals Containing Truth	Number of Predictions	Expert	Estimating Task
Russo & Schoemaker (1992)					
Advertising 1	90%	39%	750	Corporate Managers	Advertising Industry Knowledge
Advertising 2	50%	22%	750	Corporate Managers	Advertising Industry Knowledge
Computers 1	95%	20%	1,290	Corporate Managers	Computer Industry Knowledge
Computers 2	95%	42%	1,290	Corporate Managers	Computer Firm Knowledge
Data Processing 1	90%	58%	252	Corporate Managers	Data Processing Industry Knowledge
Data Processing 2	90%	38%	261	Corporate Managers	General Business Knowledge
Money Management 1	90%	50%	480	Corporate Managers	Financial Industry Knowledge
Petroleum 1	90%	50%	850	Corporate Managers	Petroleum Industry and Firm Knowledge
Petroleum 2	50%	21%	850	Corporate Managers	Petroleum Industry and Firm Knowledge
Pharmaceutical 1	90%	51%	390	Corporate Managers	Pharmaceutical Firm Knowledge
Security Analysis 1	90%	36%	497	Corporate Managers	Security Industry Knowledge
McKenzie et al. (2008)	90%	44%	1,720	IT Professionals	IT Industry Knowledge
Ben-David et al. (2013)	80%	33%	11,600	Chief Financial Officers	Stock Market Return (S&P 500)
Goldenson & Stoddard (2013)					
Battery 1	100%	10%	140	Graduate Students	Software Development Effort
Battery 2	100%	50%	80	Graduate Students	Software Development Effort
Battery 3	100%	70%	70	Graduate Students	Software Development Effort
Bar-Yosef & Venezia (2014)	95%	7%	30	Brokerage Analysts	Financial Forecasts



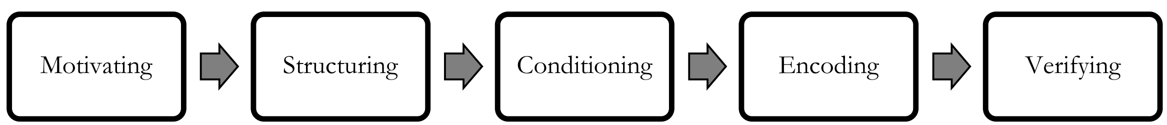
Analysis:

- Experts are overconfident—on average, never identified more than 70% of the range
- Magnitude of overconfidence was inconsistent across studies
  - Would adjusting expert’s range to meet a minimum coeff. of variation be superior to utilizing standard 30% rule-of-thumb?

Q2: Elicitation Model

- Based on Stanford Research Institute (SRI) model
  - First proposed by Spetzler & Stael von Holstein (1975)
  - Morgan & Henrion (1990) cite SRI model as most influential elicitation model
  - Model further adapted by this author for cost and schedule estimating

- Five Phases
  1. Motivating
  2. Structuring
  3. Conditioning
  4. Encoding
  5. Verifying



Q2: Elicitation Model

Motivating
- Motivate expert
- Educate analyst…beware "hypocognition"

Structuring
- Break down estimate
- Units of Measure

Conditioning
- Explain features of distribution to be utilized
- Explain order of questions to follow
- Emphasize importance of verbalizing rationale & assumptions

Encoding
- Ask for low, high, and then most likely estimates
- What is the probability of outcome greater than high?
- What is the probability of outcome less than low?
- Ask why, why, why… and document

Verifying
- Adjust to 100% confidence interval, if necessary
- Are there any scenarios that could cause the outcome to fall outside range? (consider the opposite)
- Compare coeff. of variation against established ranges

Conclusion/Recommendations

- Meta-analysis confirms JA CSRUH heuristic—on average, experts never identify more than 70% of uncertainty range
  - Magnitude of overconfidence is inconsistent—consider adjusting expert's range using coeff. of variation (CV), and not just 30% heuristic
  - Recommended research: produce accuracy and CV ranges specific to estimating modality (parametric, scaled analogy, expert opinion)…underway in AFLCMC
- Recommend adoption of structured elicitation model. Potential benefits:
  - Provides a guide for new analysts
  - Will encourage better documentation
  - More consistent/repeatable/examinable results

Back-up

Works Cited

Air Force Cost Analysis Agency. (2007). *Air Force Cost Risk and Uncertainty Analysis Handbook (CRUH)*. Washington, DC.

Bar-Yosef, S. & Venezia, I., 2014. An Experimental Study of Overconfidence in Accounting Numbers Predictions. *International Journal of Economic Sciences*, 3(1), pp.78–89.

Ben-David, I., Graham, J., & Harvey, C. (2010). *Managerial Miscalibration*.

Boyd D. W., & Regulinski S. G. (1979). *Characterizing Uncertainty in Technology Cost and Performance*. Decision Focus, Inc. Report 1114, Los Altos, California.

Brown, Bernice B. (1968). *Delphi Process: A Methodology Used for the Elicitation of Opinions of Experts*. RAND Corporation: Santa Monica, CA.

Burgman, M. A. (2016). *Trusting judgements: how to get the best out of experts*. Cambridge, United Kingdom: Cambridge University Press.

Capen, E. (1976). The Difficulty of Assessing Uncertainty (includes associated papers 6422 and 6423 and 6424 and 6425). *Journal of Petroleum Technology*, 28(08), 843-850. doi:10.2118/5579-pa

Carney, S.T. (2013). *Investigation into Risk and Uncertainty: Identifying Coefficient of Variation Benchmarks for Air Force ACAT I Programs* (Master's thesis). Air Force Institute of Technology.

Works Cited

Cohn, M. (2012). *Agile estimating and planning*. Upper Saddle River, NJ: Prentice Hall PTR.

Elicit [Def. 1]. (n.d.). Merriam-Webster Online. In Merriam-Webster. Retrieved October 19, 2017, from http://www.merriam-webster.com/dictionary/elicit.

Flyvbjerg, B. (2011). Over Budget, Over Time, Over and Over Again. *Oxford Handbooks Online*.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19-30. doi:10.1037/1040-3590.12.1.19

Galway, L. A. (2007). *Subjective Probability Distribution Elicitation in Cost Risk Analysis: A Review*. RAND Corporation: Santa Monica, CA.

Goldenson, D.R. & Stoddard R.W. (2013). *Quantifying Uncertainty in Expert Judgment: Initial Results*. Carnegie Mellon Software Engineering Institute.

Herzog, S.M., & Hertwig R. (2009). The Wisdom of Many in One Mind. *Psychological Science, 20*(2), 231-237.

U.S. Government (2015). *Joint Agency Cost Schedule Risk and Uncertainty Handbook (JA CSRUH)*. Washington D.C.

Jorgensen, M. (2007). Forecasting of software development work effort: Evidence on expert judgement and formal models. *International Journal of Forecasting, 23*(3), 449-462. doi:10.1016/j.ijforecast.2007.05.008

Works Cited

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kitchenham, B., Pfleeger, S. L., McColl, B., & Eagan, B. (2002). An empirical study of maintenance and development estimation accuracy. *Journal of Systems and Software*, 64(1), 57−77.

Levy, R. I. (1973). *Tahitians: Mind and experience in the Society Islands*. Chicago, IL: University of Chicago Press.

Li, Q., Wang, Q., Yang, Y., & Li, M. (2008). Reducing biases in individual software effort estimations. *Proceedings of the Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement - ESEM 08*. doi:10.1145/1414004.1414041

McKenzie, C.R.M., Liersch, M.J. & Yaniv, I. (2008) Overconfidence in interval estimates: what does expertise buy you? *Organizational Behavior and Human Decision Process*, 107, 179-191.

Morgan, M., Henrion, M., & Small, M. (1990). *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge, UK: Cambridge University Press.

Nashef, S., Roques, F., Michel, P., Gauducheau, E., Lemeshow, S., & Salamon, R. (1999). European system for cardiac operative risk evaluation (EuroSCORE). *European Journal of Cardio-Thoracic Surgery, 16*(1), 9-13.

Nichols, T. (2017, March & April). How America Lost Faith in Expertise. *Foreign Affairs*, *96*(2), 60-73.

Works Cited

Patterson, R. E., Pierce, B. J., Bell, H. H., & Klein, G. (2010). Implicit Learning, Tacit Knowledge, Expertise Development, and Naturalistic Decision Making. *Journal of Cognitive Engineering and Decision Making, 4*(4), 289-303.

Russo J.E. and Schoemaker P. (1992), Managing Overconfidence, *Sloan Management Review*, vol. 33, no. 2, pp. 7-17.

Sanders, N. R., & Ritzman, L. P. (1991). On Knowing When to Switch from Quantitative to Judgemental Forecasts. *International Journal of Operations & Production Management, 11*(6), 27-37. doi:10.1108/01443579110005523

Soll, Jack & Klayman, Joshua. (2004). Overconfidence in Interval Estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 30(2), 299-314.

Tversky, A. & Kahneman, D. (1974). Judgement under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124-1131.

Shen, L., Chen, X., Gu, J., & Xue, S. (2018). Validation of EuroSCORE II in Chinese Patients Undergoing Coronary Artery Bypass Surgery. *The Heart Surgery Forum, 21*(1), 036.

Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G., & Burgman, M. (2010). Reducing Overconfidence in the Interval Judgements of Experts. *Risk Analysis*, 30(3), 512-523.

Spetzler C.S., & Staël Von Holstein C.S.E., (1975). Probability Encoding in Decision Analysis. *Management Science, 22*(3):340-358.

Works Cited

Teigen, K. H., & Jorgensen, M. (2005). When 90% confidence intervals are 50% certain: on the credibility of credible intervals. *Applied Cognitive Psychology*, 19(4), 455–475.

Valentine, S. (2019, March 13). Personal interview.

Watern, K. (2016). Cost Assessment Data Enterprise (CADE). *Air Force Comptroller Magazine*, 49(1), 25.

Wilson, T.D., Houston, C.E., Etling, K.M., & Brekke, N. (1996). A New Look at Anchoring Effects. *Journal of Experimental Psychology*: General, 125(4): 387-402.

Winkler, R. L. (1967). The Assessment of Prior Distributions in Bayesian Analysis. *Journal of the American Statistical Association*, 62(319), 776-800.

Strategies

Strategies for decreasing overconfidence:

- Avoid anchoring bias—ask for the high and low outcomes prior to asking for the most likely (e.g. Soll and Klayman, 2004; Speirs-Bridge et al., 2010).
- Allow expert to self-select the confidence level
- If expert struggles, manually walk the expert through the creation of the prediction interval (Teigen & Jorgenson, 2005). For example:
  - "Could the requirement exceed 1,000 hours?", or
  - "What is the probability that the requirement exceeds 1,000 hours?".
- Ask "why?"….a lot
- Ask expert to consider why they may be wrong (Herzog & Hertwig, 2009):
  - "Imagine that the true requirement is greater than the upper bound of your given prediction interval. What are a few reasons this could be? What assumptions or considerations may be wrong?"

Elicitation Checklist

1. Motivating
   - ❑ Analyst: Familiarize yourself with the requirement needing expert elicitation. Begin formulating questions, and gather data that may be relevant to the expert.
   - ❑ Tell the Expert: The purpose of this cost estimate is to estimate _____ in support of _____.

2. Structuring
   - ❑ *Ask the Expert*: Should we break down the estimation of the requirement into smaller components?
   - ❑ *Ask the Expert*: Would you feel most comfortable estimating the unknown quantity in person-hours, full-time equivalents (FTEs), SLOC, or another unit?
   - ❑ *Ask the Expert*: What ground rules and assumptions are you making about the requirement being estimated?

Elicitation Checklist (cont.)

1. Motivating
   - ❑ Analyst: Familiarize yourself with the requirement needing expert elicitation. Begin formulating questions, and gather data that may be relevant to the expert.
   - ❑ Tell the Expert: The purpose of this cost estimate is to estimate _____ in support of _____.

2. Structuring
   - ❑ *Ask the Expert*: Should we break down the estimation of the requirement into smaller components?
   - ❑ *Ask the Expert*: Would you feel most comfortable estimating the unknown quantity in person-hours, full-time equivalents (FTEs), SLOC, or another unit?
   - ❑ *Ask the Expert*: What ground rules and assumptions are you making about the requirement being estimated?

Elicitation Checklist (cont.)

3. Conditioning
   - ❑ Tell the Expert: Today I will ask your assistance in constructing the triangular or Beta-PERT distribution that best represents your state of knowledge. I will begin by asking for your low outcome, followed by your high outcome, and lastly I will ask for the most likely outcome.
   - ❑ Tell the Expert: Subsequently, I will ask you for the probability (or likelihood) that the costs will be lesser/greater than your estimated low and high.
   - ❑ Tell the Expert: When providing your response for low/high/most likely, please explain the assumptions, rationale, mental model, or analogy used to estimate each outcome. This will help us defend the estimate to decision makers, and will be useful if the estimate is later revisited.

Elicitation Checklist (cont.)

4. Encoding
   - ❑ Ask the Expert: What is the low outcome? Why?
   - ❑ Ask the Expert: What is the high outcome? Why?
   - ❑ Ask the Expert: What is the most likely outcome? Why?
   - ❑ Ask the Expert: Could an outcome be less than your low estimate? If so, what is the probability? What scenario could cause this to happen?
   - ❑ Ask the Expert: Could an outcome be more than your high estimate? If so, what is the probability? What scenario could cause this to happen?

Elicitation Checklist (cont.)

5. Verifying
   - ❑ Analyst: If the expert responded that the interval had a confidence interval of less than 100 percent, adjust the expert's low and high using JA CSRUH Table 2-8 so that a 100 percent confidence level is reached. These values are the distribution's absolute min and max.
   - ❑ Ask the Expert: Are there any conceivable scenarios that could cause the outcome to be less than the minimum? If so, what is the probability? What scenario could cause this to happen?
   - ❑ Ask the Expert: Are there any conceivable scenarios that could cause the outcome to be more than the maximum? If so, what is the probability? What scenario could cause this to happen?
   - ❑ Ask the Expert: Does the distribution require any further adjustments? Does it best represent your current state of knowledge?
   - ❑ Analyst: The elicitation is complete. Thank the expert for their time. Compute the elicited distribution's coefficient of variation (CV), and consider adding 30 percent additional uncertainty if the CV is low (less than 0.25). Note that the expected CV will vary depending on the requirement being estimated and the milestone that the estimate is supporting.