



# Building a Credible SCAMPI Appraisal Representative Sample

Bob Moore, Business Transformation Institute, Inc.
Will Hayes, Software Engineering Institute


# What is Design of Experiments?

- Design of Experiments (DOE) is a mathematical statistics technique used to help understand the influence that different experimental factors have on the response from a system.
  - DOE allows us to understand the interaction between factors, as opposed to experimentation that changes just one factor at a time.
  - DOE provides a means for maximizing the information gained from each experiment, thus reducing the number of experiments that we need to conduct.


# DOE in SCAMPI

- DOE has two applications for SCAMPI A, B, and C appraisals:
  1. **Appraisal Planning**: DOE can help to construct an appropriate representative sample of the organizational unit (OU) to be appraised.
  2. **Appraisal Execution**: DOE can help to choose which personnel should be interviewed and which questions should be asked in collecting affirmations.


# DOE and Appraisal Planning


# Why Should We Care About A Good Representative Sample?

- A well-constructed representative sample leads to a superior appraisal return by:
  - Selecting for examination the set of instantiations that provide the greatest potential for verifying process institutionalization per each member of the examined set of instantiations.
    - This provides the most information gained per appraisal resource invested.
    - Other sets of instantiations could be examined, but would be inferior with respect to insights gained on process institutionalization.
  - Enhancing the credibility of the appraisal by providing defensible reasoning that led to the selection of some instances to be included in the appraisal while excluding others:
    - A representative sample that excludes some instantiations without clear reason invites suspicion that the appraisal results may not reflect OU process institutionalization because instantiations detrimental to the OU's case for institutionalization are being avoided.
    - Likewise, a representative sample that insists on including some instantiations without justification might raise questions about the appraisal results again, only this time because instantiations that reflect atypical "good" institutionalization effort are being included.


# How are Representative Samples Constructed Now?

- The SCAMPI Method Description Document does not give us much advice!
  - "Upon determining that sufficient coverage of the reference model and the organizational unit has been obtained, appraisal findings and ratings may be generated." (SCAMPI MDD, p. I-11.)
  - Coverage is said to imply:
    - "(a) the collection of sufficient data for each model component within the CMMI reference model scope selected by the sponsor, and"
    - "(b) obtaining a representative sample of ongoing processes) spanning the life-cycle phases that the appraised organization is using in the development and delivery of its products and services."
  - The lead appraiser is further cautioned to construct a "valid sample of the organizational unit to which results will be attributed" based on organization size, scope, and geographic dispersion.
  - The lead appraiser and sponsor are reminded that all statements should be accurate in describing the organization to which results may be attributed.


# Does The MDD's Guidance Work?

- Given this guidance, how is a lead appraiser to construct a "valid sample" that can withstand rigorous, independent examination?
- The current typical practice of using no more than four projects in an appraisal, no matter the size of the appraised organizational unit, may entirely miss information that characterizes how well or poorly the OU is doing with its processes.
- Unfortunately, increasing the number of projects examined doesn't help!
  - Very large samples of projects from a large OU soon become cost prohibitive without providing analytically defensible insight into process performance
  - Although saying "we looked at 10 projects and 10,000 artifacts" sounds impressive—even if it isn't!


# But What Else Can We Do?

- Since a SCAMPI A appraisal is meant to provide a benchmark of an OU's process performance, we need some technique that:
  - Seeks to maximize information received,
  - Minimizes cost, and
  - Provides appropriate rigor to justify our appraisal planning choices to an independent examiner.
- DOE provides exactly these capabilities!


# DOE Language and SCAMPI

- Experiment = an appraisal
- Experimental factors = characteristics of the OU as they are observed across different parts of the organization where work is underway
- Experimental design = the list of instantiations from which we will examine artifacts, based on:
  - The experimental factors present in the OU,
  - The budget available for the appraisal, and
  - The amount of confounding between factors we are willing to accept.
- Response variables = weakness and strengths of process area specific or generic practices and satisfaction of goals.
- Factors effects = the influence that different factors have on the response variables under consideration.
- Confounding = our inability to distinguish between the influence on the response variables of one or more factors with respect to another set of factors. Confounding is undesirable, but may be managed through choice of designs.


# DOE Language and SCAMPI, Continued

- Replication = examining more than one instantiation corresponding to a particular set of experimental factors in our chosen design—which provides better insight into institutionalization by having additional instantiations to confirm observed responses.
- Balanced design = a fractional factorial design in which an equal number of trials (at every level state) is conducted for each factor.
- Block & Blocking = When structuring appraisals, blocking may be used to account for some unknown that one wishes to avoid; a block may be a dummy factor that does not interact with the real factors.
- Orthogonal = An appraisal is orthogonal if the effects of any factor balance out (sum to zero) across the effects of the other factors.


# Experimental Resolution

- Experimental resolution helps us to understand the degree of our "known unknowns" in an appraisal.
  - Resolution I = we gain no insight from an appraisal
  - Resolution II = we cannot tell the difference between the influence of main factor effects (why bother?)
  - Resolution III = Main factor effects are confounded (aliased) with two-factor interactions.
  - Resolution IV = No main factor effects are aliased with two-factor interactions, but two-factor interactions are aliased with each other.
  - Resolution V = No main effect or two-factor interaction is aliased with any other main effect or two-factor interaction, but two-factor interactions are aliased with three-factor interactions.


# Example OU Experimental Factors

- The factors that influence process institutionalization in an OU depend on that OU. Some typical factors to be considered:
  - The size of the project:
    - Projects that are large or small with respect to the OU's typical project mixture may influence how processes are used.
  - Project age:
    - New or existing projects for the OU may have different understanding or maturity of processes.
  - Project geographic location:
    - Projects performed at a core location or at a remote site may differ in their processes.


# Example OU Experimental Factors, Continued

- Project dispersion:
  - Projects that, within the context of the project, are executed at one location or multiple locations that are inconvenient for daily face-to-face contact may have different processes.
- Project parent organization:
  - The "home" or sponsoring OU for a project may influence how processes are implemented depending on the support of management for the processes.
- Project complexity:
  - Projects that have complex life cycles may have different processes than simpler projects (e.g., spiral versus waterfall life cycle).
- Project customer and users:
  - Projects performed for different customers or users may use different processes depending on the customer or user's requirements.


# How to Select a Representative Sample for an Appraisal (1)

1. Determine the objectives of the appraisal with respect to the OU scope and process areas to be considered.
2. List the factors that may influence process institutionalization in the OU.
   - Be generous in listing factors—a factor that has no real impact is easily discarded through the application of DOE techniques, but omitting a factor of real influence may skew the appraisal conclusions.
3. Determine if any of the factors are clearly dependent on other factors. If so, these factors may be collapsed into fewer combined factors.
4. Determine the level settings for the factors, such as project size equals one of "large" or "small". Any given factor may have multiple levels, although two levels are easiest from a design and analysis perspective.
5. List all of the instantiations in the OU that are supposed to be using processes corresponding to the process areas under consideration.


# How to Select a Representative Sample for an Appraisal (2)

6. For each instantiation in the list, determine the factor level settings that describe that instantiation.
   - For example, project X may have factor levels of size=large, location=central office, and duration=long, where as project Y may have factor levels of size=small, location=field site, and duration=short.
7. Given the list of factors and their level settings, choose an experimental design.
   - This design will be determined by how much confounding between factors is tolerable and the budget limits on how many different instantiations can be examined in the appraisal.
   - A design catalog or statistical software that supports DOE is indispensable here for exploring the options!
8. Fill in the experimental design from step 7 with actual instantiations using the information in step 6.


# An Example of Selecting A Representative Sample Using DOE

- Suppose we are examining an OU that has five factors to be accounted for in an appraisal:
  - Project size: large or small
  - Project age: new or existing
  - Project geographic location: domestic or international
  - Project customer: government or commercial
  - Project complexity: high or low
- We have 5 factors at 2 levels that might influence process institutionalization in the OU.


# Full Factorial Design

- The full factorial design (all factors at all levels), we would have to examine 32 (2x2x2x2x2) instantiations!
- The design is given on the next page for illustration purposes.
- No one is expected to ever construct such an appraisal.

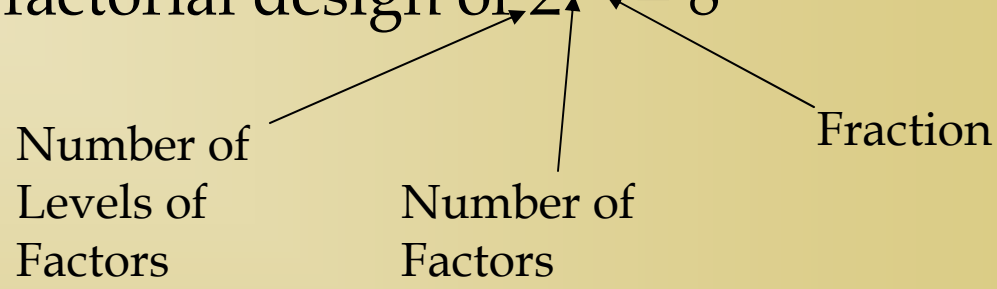

| | Size | Age | Location | Customer | Complexity |
|---|---|---|---|---|---|
| Instantiation = 1 | L = large | N = new | D = domestic | G = government | H = high |
| 2 | L | N | D | G | L = low |
| 3 | L | N | D | C = commercial | H |
| 4 | L | N | D | C | L |
| 5 | L | N | I = international | G | H |
| 6 | L | N | I | G | L |
| 7 | L | N | I | C | H |
| 8 | L | N | I | C | L |
| 9 | L | E = existing | D | G | H |
| 10 | L | E | D | G | L |
| 11 | L | E | D | C | H |
| 12 | L | E | D | C | L |
| 13 | L | E | I | G | H |
| 14 | L | E | I | G | L |
| 15 | L | E | I | C | H |
| 16 | L | E | I | C | L |
| 17 | S = small | N | D | G | H |
| 18 | S | N | D | G | L |
| 19 | S | N | D | C | H |
| 20 | S | N | D | C | L |
| 21 | S | N | I | G | H |
| 22 | S | N | I | G | L |
| 23 | S | N | I | C | H |
| 24 | S | N | I | C | L |
| 25 | S | E | D | G | H |
| 26 | S | E | D | G | L |
| 27 | S | E | D | C | H |
| 28 | S | E | D | C | L |
| 29 | S | E | I | G | H |
| 30 | S | E | I | G | L |
| 31 | S | E | I | C | H |
| 32 | S | E | I | C | L |


# Alternatives to Full Factorial

- Except in limited circumstances, a full factorial selection on instantiations is too expense and too time consuming.
  - Note: for this presentation, we are neglecting the idea that an appraisal might want to look at more than one instantiation for each setting of factors (replication). Looking at multiple instantiations for the same factors is a good idea—but the number of instantiations to be examined grows rapidly!
- Besides, who needs complicated math to try every combination of everything?
- DOE offers an alternative: fractional factorial designs.

# A ¼ Fractional Factorial Design

- In the example above, instead of using a full factorial design, we could also have conducted our appraisal using a fractional factorial design of $2^{5-2} = 8$ instantiations.




- A ¼ design in this case is a Resolution III experiment.
- The choice of a fractional factorial design will depend on the number of factors to be considered and the acceptable experimental resolution.


# The ¼ Fractional Factorial Design

| Instantiation | Size | Age | Location | Customer | Complexity |
|---|---|---|---|---|---|
| 1 | S | E | I | G | H |
| 2 | L | E | I | C | L |
| 3 | S | N | I | C | H |
| 4 | L | N | I | G | L |
| 5 | S | E | D | G | L |
| 6 | L | E | D | C | H |
| 7 | S | N | D | C | L |
| 8 | L | N | D | G | H |


# Still Too Many Instantiations!

- From the viewpoint of a SCAMPI A appraisal, using 8 instantiations across multiple process areas still seems like a lot!
  - Note: we're still doing better than a traditional representative sample selection method—we at least clearly understand the impact of different factors on our appraisal.
- Going to a 1/8 fractional factorial would give 4 instantiations to be appraised—but drops our resolution down to Resolution II.
- What to do . . . ?


# Using DOE with SCAMPI B and C

- DOE works best not as a single experiment, but as a sequence.
- This is ideally suited to SCAMPI:
  - Conduct early appraisals that examine many factors and instantiations as SCAMPI Cs.
  - Based on the results, eliminate factors (and the need for instantiations).
  - Conduct later appraisals that examine fewer *critical* factors and instantiations as SCAMPI As or Bs.
  - Note: changing lead appraisers from one appraisal to another allows you to block your design according to lead appraiser—if lead appraisers are unbiased!


# Example SCAMPI C

- Consider the setup in the example above: 5 factors that may influence the OU's process institutionalization.
- We would like to determine which factors really influence the process and which are not important.
- Eventually, we want to benchmark the OU using a SCAMPI A.
- We will start with a SCAMPI C.
- For illustration purposes, we will only look at the appraisal covering two process areas (PP and REQM) at Capability Level 2.
  - The example could be expanded to as many PAs as we like, but the calculations are lengthy in a presentation.


# Assigning Numerical Values to Response Variables

- As defined in the SCAMPI C method, we would usually assign a color (green, yellow, red) as the characterization of each instantiation's specific and generic practices.
- To aid in our analysis, we will assign numerical values against these characterizations:
  - Red = 0
  - Yellow = 0.5
  - Green = 1.0
- The assigned values may be changed, if desired.
- Similar values may be assigned for characterizations using in other SCAMPIs. For example:
  - NI = 0
  - PI = 0.5
  - LI = 0.75
  - FI = 1


# Aggregation at the Goal Level

- To aid in our analysis, we are going to take the arithmetic mean of the specific practices and generic practices at the goal level for each instantiation.
- For REQM (similarly for PP),

$$\text{Score(SG 1)} = \frac{Score(SP1.1-1) + Score(SP1.2-2) + Score(SP1.3-1) + Score(SP1.4-2) + Score(SP1.5-1)}{5}$$

$$\text{Score(GG 2)} = \frac{Score(GP2.1) + Score(GP2.2) + \cdots + Score(GP2.10)}{10}$$


# Characterization Data from the SCAMPI C

- We perform the SCAMPI C appraisal using the instantiations given above.
- The results:

| | S | A | L | C | C | PP SG 1 | PP SG 2 | PP SG 3 | PP GG 2 | REQM SG 1 | REQM GG 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Instantiation = 1 | S | E | I | G | H | 0.38 | 0.71 | 0.83 | 0.75 | 0.60 | 0.75 |
| 2 | L | E | I | C | L | 0.25 | 0.86 | 0.83 | 0.75 | 0.60 | 0.75 |
| 3 | S | N | I | C | H | 0.88 | 0.93 | 1.00 | 0.95 | 0.80 | 0.95 |
| 4 | L | N | I | G | L | 0.88 | 0.86 | 1.00 | 0.95 | 0.90 | 0.95 |
| 5 | S | E | D | G | L | 0.13 | 0.14 | 0.17 | 0.20 | 0.10 | 0.20 |
| 6 | L | E | D | C | H | 0.25 | 0.14 | 0.17 | 0.20 | 0.10 | 0.25 |
| 7 | S | N | D | C | L | 0.50 | 0.43 | 0.50 | 0.40 | 0.40 | 0.45 |
| 8 | L | N | D | G | H | 0.50 | 0.43 | 0.50 | 0.50 | 0.30 | 0.45 |


# Analysis of Data

- In a simple analysis, we account for the impact of any particular factor (e.g., instantiation size or age) by:
  - Adding the responses for the goal when a given factor is set "high";
  - Subtracting the responses for the goal when the same factor is set "low"; and,
  - Dividing the result by the number of high (or low) settings (i.e., 4 in this case.)
- Let R(x) equal the response value for instantiation x.
- For example, the impact of age on PP, SG 2 (across all instantiations) is:

¼ *[Sum(Responses when Age = "Existing") – Sum(Responses when Age = "New")] =

¼ * [R(2)+R(4)+R(6)+R(8)-R(1)-R(3)-R(5)-R(7)] = 0.017857


# Full Data Results

| | PP SG 1 | PP SG 2 | PP SG 3 | PP GG 2 | REQM SG 1 | REQM GG 2 |
|---|---|---|---|---|---|---|
| Effect of size | 0.00 | 0.02 | 0.00 | 0.03 | 0.00 | 0.01 |
| Effect of age | 0.20 | 0.20 | 0.25 | 0.23 | 0.25 | 0.21 |
| Effect of location | -0.58 | -0.55 | -0.58 | -0.53 | -0.50 | -0.51 |
| Effect of customer | 0.00 | -0.05 | 0.00 | 0.03 | 0.00 | -0.01 |
| Effect of complexity | 0.01 | -0.02 | 0.00 | 0.03 | -0.05 | 0.01 |

- Our conclusion is that instantiation location and age have an impact on process institutionalization.
- All other factors appear to have negligible impact.


# Next Steps

- The analysis given here is very elementary.
  - More sophisticated analysis techniques may be found at http://www.itl.nist.gov/div898/handbook/ and its references.
  - Additional designs, appropriate for many more situations, may be found at the same location.
- Given the analysis, our next appraisal might be a SCAMPI B that examines only two factors: location and age.
  - A design using only two factors is full factorial with $2^2 = 4$ instantiations.
  - We may wish to conduct the next appraisal with replication against some of the design elements, to provide more insight into institutionalization.


# What Have We Learned?

- DOE provides a technique to help us choose appraisal representative samples in a more rigorous manner.
- DOE fits with conducting a sequence of SCAMPI appraisals, leading to a benchmark SCAMPI A.
- DOE techniques may be applied in a SCAMPI context with similar schedule duration to traditional SCAMPIs.
- DOE can be a complex subject, but there are many software packages and online and print references to make applying it easier.


# What Haven't We Discussed

- DOE techniques actually work better for planning SCAMPIs for large OUs because there are more instantiations available for any given design.
- Instantiations that reflect some factor settings may not be available in all OUs—we haven't covered how to handle this situation.


# DOE and Interview Questions


# The Interview Dilemma

- Conducting interviews in an appraisal gives much the same challenge as choosing a representative sample.
  - There are many questions to ask and many people to whom we wish to ask them.
- How do we choose?
- **Note**: If we have information needs, then we will want to ask particular people specific questions!
- DOE is useful for general questions intended to fulfill face-to-face affirmation coverage requirements.


# Example: Designs for Interviews

- We can categorize personnel as “managers”, “engineers”, and various kinds of “support”.
- Each person will also have an instantiation (possibly more than one) associated with them.
- In this case, the personnel categories and the personnel’s binned instantiation provide the settings for the factors.
- We choose the questions to be asked of each person based an experimental design guiding us to sample certain combinations of personnel categories and instantiations.
- Due to SCAMPI coverage requirements, particularly for SCAMPI As, we will need a fractional factorial constrained design.
  - Unlike the regular fractional factorial, constrained designs are not conveniently available for reference.
  - Our only choice in this case is to use software that supports DOE.


# Example: Interview Design

- Note: this is an example to demonstrate how the technique might be applied, not a real design.
  - The ideas are the same as applied in choosing a representative sample, so we will not repeat the details!
- Suppose we have personnel categories "manager" and "engineer".
- Suppose we have two instantiations to consider: project 1 and project 2.


# The Design

| | Question 1 | Question 2 | Question 3 | Question 4 |
|---|---|---|---|---|
| Manager, Project 1 | Ask | No | No | Ask |
| Manager, Project 2 | No | Ask | Ask | No |
| Engineer, Project 1 | No | Ask | Ask | No |
| Engineer, Project 2 | Ask | No | No | Ask |


# Summary

- DOE provides a powerful method for designing reasonable representative samples.
  - DOE is of greatest benefit in dealing with large OUs with many factors and instantiations.
  - DOE works well in screening out instantiations that do not provide much "new" information through SCAMPI Cs and Bs.
- DOE provides a means during an appraisal for determining the general questions to ask various personnel types on different projects.
  - Specific questions to answer information are still directed as usual.


# Contact Information


- Bob Moore, Business Transformation Institute, Inc.
  - rlmoore@biztransform.net
- Will Hayes, Software Engineering Insitute
  - wh@sei.cmu.edu