THE JOURNAL OF AACE® INTERNATIONAL -
THE AUTHORITY FOR TOTAL COST MANAGEMENT®

COST ENGINEERING

July/August 2016

www.aacei.org

UNMANNED AERIAL VEHICLES FOR CONSTRUCTION MANAGEMENT

WHAT TO DO WITH UNKNOWNS IN SCHEDULE RISK ANALYSIS?

RECOMMENDED PRACTICES PROVIDE IMPLEMENTATION DETAILS FOR CMAA STANDARDS OF PRACTICE

aace International
The Authority for Total Cost Management®

TECHNICAL ARTICLE

What to Do With Unknowns in Schedule Risk Analysis?

Dr. David T. Hulett

Abstract: Schedule risk analysis explores how unknowns applied to the project schedules may derive a distribution of possible completion dates. Unknowns include known unknowns (we know the cause but do not know whether the risk will occur; and/or, if it occurs, its impact on activity durations), unknown unknowns (those risks that are not known today, whether they could be known or not with further inquiry) and "unknown knowns" (we know but do not want to discuss them in a public forum). This article describes the types of unknowns, and the methods used to incorporate the unknowns to drive the Monte Carlo simulation of the schedule. Methods include:

- Using the three-point estimate to represent inherent variability and estimating error and bias.
- Using the risk interview to uncover unknown knowns.
- Expanding the three-point range for "far future" (in the context of the project) activities.

A simple case study shows, through use of Monte Carlo simulation, examples of the methods and of their impact on final answers. This article was first presented at the 2015 AACE International Annual Meeting as RISK.1993.

Schedule risk analysis explores how unknowns applied to the project schedules may derive a distribution of possible completion dates. Unknowns include:

- Known unknowns for which the cause is known, but it is not known whether the risk will occur; and/or, if it occurs, its impact on activity durations. Interestingly, risks with 100 percent probability of occurring, though usually called "issues," are included if their impact is unknown.
- Unknown unknowns are those risks that are not known today but it may reliably be expected that they will occur in the future. It is not clear that these risks are "unknowable," because they have been experienced in many projects over time. People are often myopic about the risks they want to discuss, so near-term risks are often the focus of attention during interviews or workshops. Whether they could be known or not with further inquiry, or whether benchmarking to historical data can help is an area of interest.
- Unknown knowns may be a new class of unknowns to some people. It is known these risks exist and often their parameters (probability and impact) are known. The management does not want to discuss them in a public forum such as a risk workshop since they are sensitive or pessimistic, causing harm and even cancellation of the project. Confidential interviews always reveal risks not in the risk register that are agreed to by subsequent interviewees and, on inspection, turn out to be the most important risks of all.

This article describes the types of unknowns and the methods used to incorporate these unknowns to drive a Monte Carlo simulation of the schedule. Methods include:

- Using a three-point estimate to represent inherent variability, estimating error, and estimating bias.
- Expanding the three-point range for "far future" (in the context of the project) activities for unknown unknowns.
- Using the risk interview to uncover unknown knowns.

A simple case study shows, through use of a Monte Carlo simulation, examples of these methods and of their impact on the final answers.

The Problem

This article presents a discussion of the representation of different uncertainty concepts in quantitative analysis of a project schedule risk. The issue is whether quantitative schedule risk analysis covers adequately and explicitly, the various kinds of unknowns, typically classified as known unknowns and unknown unknowns. There are also known knowns that are understood and probably incorporated in the project schedule explicitly. Recently, another classification, unknown knowns, which means risks that are known but not talked about, has been added to the list. These unknowns must each be represented if the analysis results are to be useful to project managers.

The most troubling of these are the unknown unknowns, because how can one represent an uncertainty that one knows nothing about, particularly its cause, its likelihood of occurring, and its impact? It is argued here that unknown unknowns are not necessarily unknowable, but are not known at the moment. However, experience shows that people's attention during risk interviews or workshops are focused on the near-term uncertainties. This myopia may cause risks in the far-future of the project to be poorly, if at all, discussed. Also, the work of Bent Flyvbjerg, E.M. Merrow and John Hollmann has shown that benchmarking may reveal at least the existence of these risks that are unknown. This leads to the opportunity to discover them, or at least to represent them with wider ranges on the durations of activities in this far future (many months or years from now), depending on the length of the project.

Categories of Unknowns

Donald Rumsfeld, formerly the Secretary of Defense of the United States. famously said in February 2002:

"Reports that say that something hasn't happened are always interesting to me, because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we don't know we don't know.

And if one looks throughout the history of the US and other free countries, it is the (unknowns) in the latter category that tend to be the difficult ones" [11].

According to this description, there are three categories of knowns and unknowns that need to be considered and modeled, in a schedule risk analysis using Monte Carlo simulation of a project schedule.

Known Knowns

Known knowns are events or conditions that are known to exist and to have an impact known with certainty. Known knowns are items that should be included in the plan, schedule, and budget. There are still assumptions behind these known knowns, and those assumptions may be unknown as to existence, or impact, or both. Some plans do not include all the known knowns.

Variation Caused by Known Unknowns

Fundamentally, "There are no facts about the future" [5]. Known unknowns are events or conditions that can be described well enough to know the cause. However, one does not know whether this event or condition will occur; or, if it does, what its impact will be. These known unknowns include:

- Pure uncertainty for which there is no specific cause, but it is known with certainty to exist with an uncertain impact on the schedule.
- Risk events that can be described and for which the specific cause is known. These, in turn, may or may not occur, or may be certain to occur, but in either case their impacts are not known with certainty.

Included in pure uncertainty are those conditions that are known with certainty to exist, but they must be included in the risk analysis since their impact is still uncertain. Some organizations might call these "issues," and therefore exclude them from consideration in a risk analysis. This is incorrect, since the impacts are still in the future and are still uncertain. It does not matter what these 100 percent likely conditions are called, they are still included in the risk analysis.

- Uncertainty—the inherent variability in project activities that arise because people and organizations cannot do things reliably on plan.
- Estimating Error—attaches to all types of estimates.
- Estimating Bias—estimates may be slanted, usually toward shorter durations, to make desired project results.

Inherent Variability

Inherent variability is similar to "common cause" variation described by Walter A. Shewhart and championed by W. Edwards Deming. Common cause variability is a source of variation caused by unknown factors that results in a steady, but random, distribution of output around the average of the data. Common cause variation is a measure of the process's potential, or how well the process can perform when special cause variation is removed. Common cause variation is also called random variation, noise, non-controllable variation, within-group variation or inherent variation. A condition leading to common cause variability would be many X's with a small impact [9].

In other words, inherent variability is random "noise" in the schedule, attaching to durations indicating that individuals and organizations cannot be relied upon to perform to plan. A memorandum may take two hours or two days to write. This can be as a result of the unpredictability of the author's productivity, attention span, or the amount of interruptions. There are many of these influences that can be relied upon to occur, but they cannot be catalogued or addressed in the span of one project. Corralling this multitude of influences and their consequences, in the space of a single project with a specific team, is not possible. Advances to narrow the inherent variation would take repetition of the project many times with the same teams, contractors, and conditions. Such progress was made in World War II, as thousands of airframes were produced, and teams and their organizations learned how to be more routine and efficient. A project is not likely to experience such learning.

By being 100 percent likely to happen, inherent variability can be represented by applying a three-point estimate, with a probability distribution, directly to the activity durations that will experience this variability. The uncertainty of impact is thought of as being a fairly tight probability distribution around the estimated duration, which becomes the most likely value (mode) of that distribution. The pure randomness of inherent variability indicates that the impact or degree of variation is symmetrical, that is plus or minus the same fraction, multiplied by the activity duration. For instance, the multiplier used may be 90 percent, 100 percent or 110 percent; representing the low, most likely, and high multiplier of the activity's durations. The multiplying fraction is chosen at random from a symmetrical distribution (e.g., a symmetrical triangle, a Gaussian, or normal distribution) during iteration in a Monte Carlo simulation.

It should be noted that different types of activities may have different levels of inherent variability. So, for instance, detailed engineering and construction may have different inherent variability. In the risk analysis of the schedule, these categories of activities can have different "reference ranges" of uncertainty. The possibility that some categories of activities have narrower inherent variability than other categories does not make this uncertainty reducible. It is "inherent" after all.

Estimating Error

Estimating error can be attributed to a lack of information concerning specific issues needed to make up an activity's duration, though it may also be a result of insufficient or inexpert resources applied to the project scheduling. The contractor may not have specific vendor information until the vendors bid, but vendor information is required for completed engineering. Ultimately, one does not necessarily have contractor bids for overall schedule development.

Each of these sources of information can be helpful eventually to narrow the estimating error. Still, the estimates are uncertain, even after contractor bids.

The duration estimating range is often related to the "class" of estimate, determined by the level of knowledge and the method of estimating. With less knowledge, the "plus and minus" range would be large; but as more information is known, it may become smaller. However, at the time of the risk analysis, the schedule is populated with estimates made with some stage of information-gathering. Like inherent variability, estimating error is 100 percent likely.

Research shows that the range of uncertainty around estimates is larger than recommended by professional associations (including AACE) [3]. These plus and minus estimates should be symmetrical around the most likely duration; which, ideally, is included in the schedule. Perhaps, adding estimating error to the inherent variability could increase the three-point estimate to 80 percent, 100 percent or 120 percent.

Estimating Bias

Estimating bias is quite common, whether talking about costs or schedule durations. The symmetry of uncertainty ranges around the estimates of duration in the schedule should always be challenged, since only in optimum situations will the scheduler and the team leads who are responsible for building the schedule, be free to make their professional judgment paramount. Optimum conditions mostly occur in text books and training courses—real scheduling is messy, and fraught with compromises and pressures. Two such pressures, one for schedule and the other applying to cost estimates, are described by Edward M. Merrow [4]:

- "I want it NOW!" Merrow says that, "Schedule pressure dooms more megaprojects than any other single factor." These pressures may come from ambitious managers who see early completion as a way for promotions. Customers usually want the project to finish early, maybe because another project is scheduled to be finished at that time. But, the other project is also late and much anguish and many late nights are dedicated to achieving finish dates that may not be realistic or even necessary. Merrow says that every megaproject has an appropriate pace that becomes known early. Pronouncements about early, and unattainable completion dates, do not change this pace.
- "We need to shave 20 percent off that cost number!" Setting to work a construction cost-reduction task force is a counterproductive exercise that may just reduce estimates, in Merrow's opinion. Reducing estimates this way is foolish, in part because, usually unsustainable assumptions are needed. The task force may actually identify scope to come out in order to cut costs, but the scope needs to be added back in

later to complete the project, so the reduction in cost is only temporary.

In looking at the reasonableness of the duration estimates, the question typically asked is, "Was there pressure put to bear on the scheduler by management or the customer, by statements or directives, or was pressure for early finish implicit in the competitive process?" If the scheduler and the team leads are aware of management's commitment to an earlier date, the schedule will have many activities estimated "on the short side." (Activities will also be scheduled in parallel, rather than in sequence). When talking with project participants (management, team leaders, SMEs) the author often finds that they do not believe the duration estimates in the schedule.

With a range represented by optimistic, most likely, and pessimistic values, these people say the "most likely" duration or cost, is not the value in the schedule for activities or the estimate for cost elements. Often, the "most likely" multiplier is 1.05, or 1.1, or more, indicating that the estimates are viewed as being five percent, 10 percent, or more above those in the project documents. Sometimes, the values in the schedule or estimate are viewed as the optimistic value, or even worse, are not even deemed possible, optimistically.

One way to combat this pressure for unrealistically-short activity scheduling might be to ask the basic question: "How long would this scope of work take if no pressure for an earlier date were brought to bear?" In other words: "How long would this scope of work take if the estimates were purely professional, without prior expectations?" Contractors generally admit that the schedule would take longer without time pressure having been put upon them, and say that a nominally 24-month schedule should be anywhere from 26 to 30 months (8 percent to 25 percent) longer. But, and even in the same breath they may say: "We can do it!"

Summary of Sources of Uncertainty

In summary, since these sources of uncertainty (e.g., inherent variability and estimating error always occur, if estimating bias exists it has already occurred) are 100 percent likely, they can be represented by a three-point estimate of multiplicative factors applied to activities' durations.

The three-point estimates representing these factors are often represented as multiplicative impacts. For the balance of this article, assume some fairly standard values:

- Inherent variability might exhibit low, most likely, and high multipliers of 0.9, 1.0, and 1.1.
- Adding estimating error, these multipliers may become 0.8, 1.0, and 1.2.
- Adding estimating bias, if it exists, these multipliers may become 0.8, 1.05, and 1.3. Notice that the most likely multiplier does not necessarily equal 1.0. That means that the durations in the schedule are not necessarily the most likely values, as seen by impartial, professional schedulers.

Two Issues About Implementing Three-point Estimates for Uncertainty

Two issues facing the risk analysis practitioner implanting arise from:

- Under-estimating the range of the distribution, whether from motivational bias or cognitive bias.
- Narrowing of the total schedule uncertainty range because simulating the activity duration uncertainty ranges exhibits cancelling out of high and low values in the absence of correlation.

Under Estimating the Duration Impact Range—Motivational and Cognitive Bias

Under estimation of impact ranges is very common. Under estimating the probability of a risk occurring is common too. That fact is not relevant in a discussion of uncertainty, since uncertainty is 100 percent likely. Under estimation needs to be combated by the risk analyst' working with the interviewees; and challenging, with data or examples from other projects, their assumption of low probability.

One problem is the overall motivation of individuals to support, or not to challenge, the organization's assumption of project success. Individuals who participate in risk discussions are often subject to pressures that punish wide distributions for motivational or political reasons. Motivational risks are often caused by people's unwillingness to show uncertainty in the estimates, or expose their inability to make good estimates. Most motivational bias involves the conflict between professional, realistic estimating, and knowing that wide estimates will cause the overall project's finish to be later than desired by management.

Even with the best intentions to produce good estimates of duration variability, there is also cognitive bias. Individuals discussing uncertainty of estimates often refer to heuristics (rules of thumb) and exhibit naturally-occurring biases that have been studied and experienced [7].

One of the most common examples of cognitive bias is described as anchoring and adjusting, where some value, perhaps the duration estimate in the schedule, forms an anchor and the interviewee cannot imagine that low and high ranges vary from that anchor. Individuals commonly understand the agreed-upon duration in the schedule and may have actually made the estimate. They find it difficult to imagine even extreme cases of lower or higher

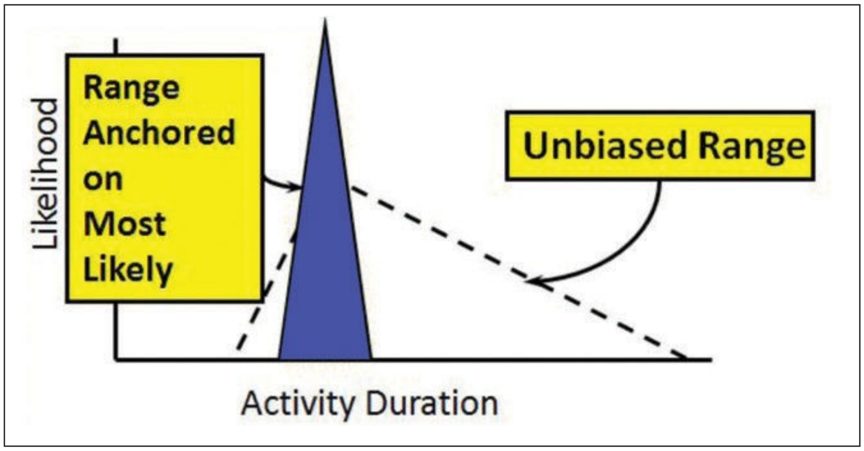


Figure 1 – Under-Estimating Durations Because of the Anchoring and Adjusting Bias

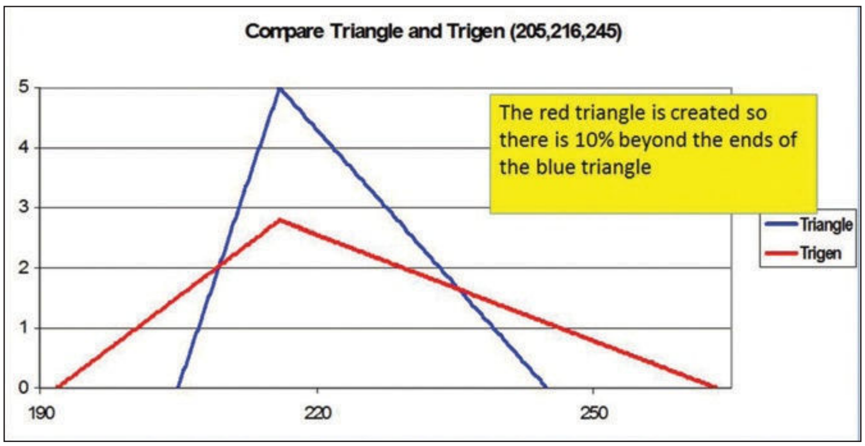


Figure 2 – TriGen Function Correcting for Underestimation of Impact Ranges

durations. A common result is found in Figure 1.

These estimates may be narrow, requiring a distribution-widening distribution. A common adjustment is to correct for an incorrect narrow distribution, by applying a TriGen (for "triangle generation") function to the three-point estimate provided by the interviewee. The risk analyst estimates that there is some percentage of the area under the unbiased triangular distribution that is below the low estimate and some percentage above the high estimate. The analyst makes this judgment, either with the help of the interviewee or upon reflection after the interview. A picture of a typical TriGen function compared to the function provided by the interviewee is shown in Figure 2.

Another example is illustrated, using the Monte Carlo method, as shown in Figure 3.

Correcting the input data for well-known biases is often required before performing a Monte Carlo simulation.

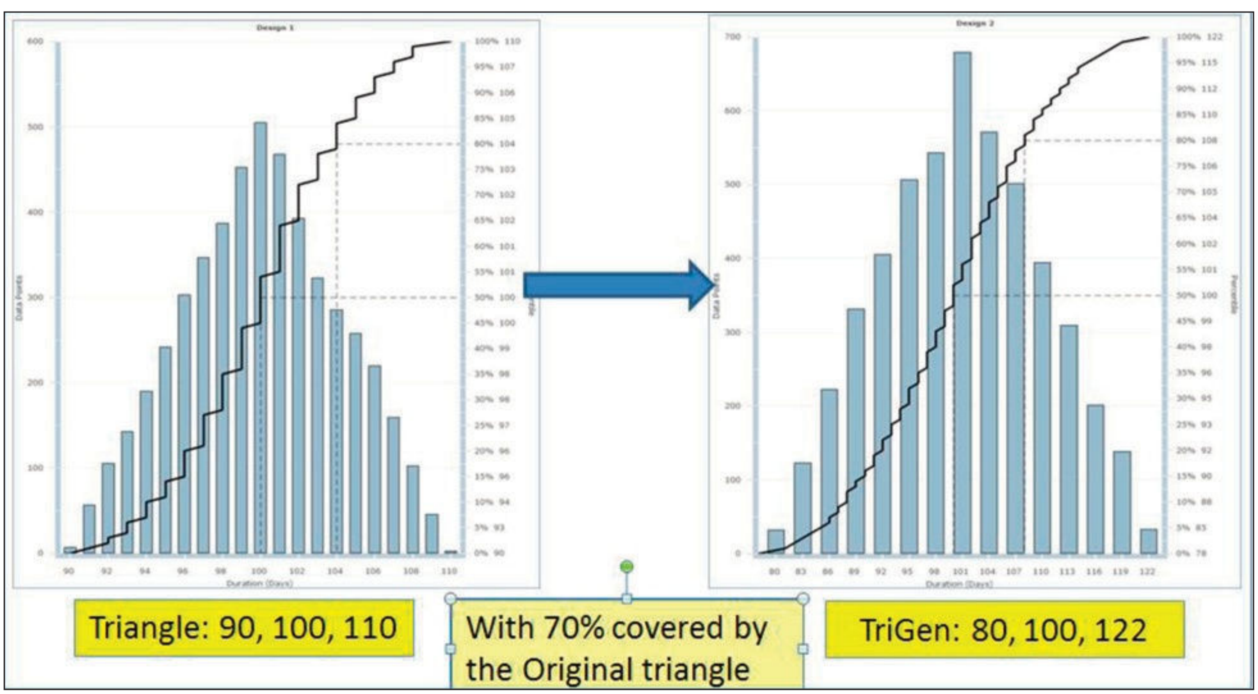
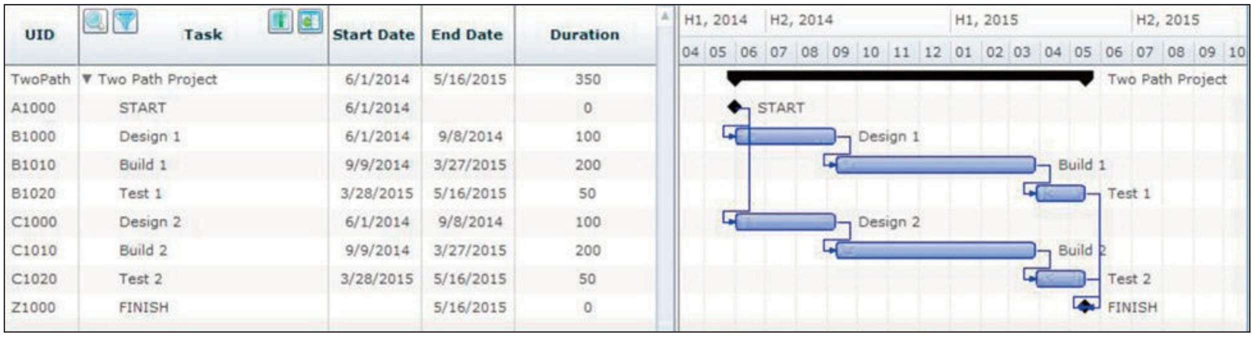


Figure 3 – Use of TriGen Function in a Monte Carlo Simulation



UID	Task	Start Date	End Date	Duration
TwoPath	▼ Two Path Project	6/1/2014	5/16/2015	350
A1000	START	6/1/2014		0
B1000	Design 1	6/1/2014	9/8/2014	100
B1010	Build 1	9/9/2014	3/27/2015	200
B1020	Test 1	3/28/2015	5/16/2015	50
C1000	Design 2	6/1/2014	9/8/2014	100
C1010	Build 2	9/9/2014	3/27/2015	200
C1020	Test 2	3/28/2015	5/16/2015	50
Z1000	FINISH		5/16/2015	0

Figure 4 – A Simple Schedule With Nominally 350 Days Duration

Many of these figures are screen shots from Polaris™ v. 1.8, an integrated cost-schedule Monte Carlo simulation program developed by Booz Allen Hamilton.

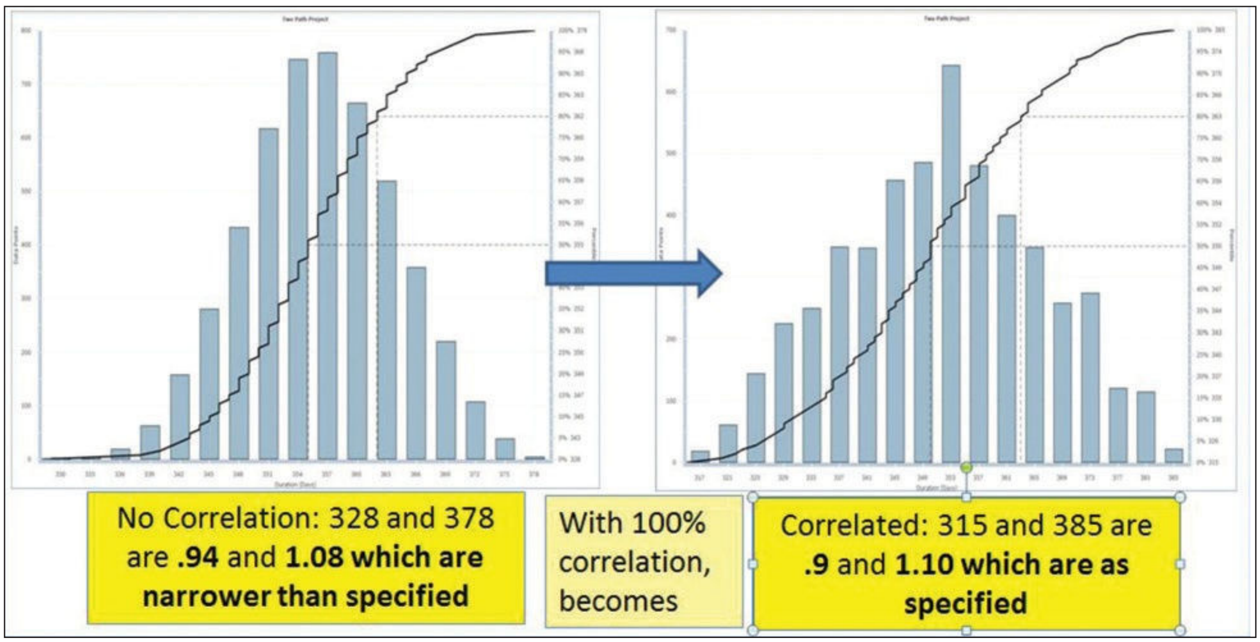


Figure 5 – Simulation Results Without Correlation and With Correlation at 100 Percent

Need to Correlate Duration Impacts

Often the interviewee or workshop will discuss their estimate of schedule uncertainty as a plus or minus percentage range around the entire schedule. Using a simple example, the source of the risk data may say: "The schedule uncertainty is plus 10 percent and minus 10 minus of the overall schedule duration." This means, in a simple example, a 20-month project could be as long as 22 months and as short as 18 months. The interviewee did not put forward an opinion about the uncertainty that is attached to the durations of individual activity durations, but the range needs to be put on individual activities at the most detailed level of the schedule.

In analysis, it is tempting to place the plus 10 percent and minus 10 percent values to the durations of the schedule's individual activities before simulating the schedule. However, the individual activity ranges will cancel each other out to some extent in any iteration because some durations will be high in their three-point estimate range, while others may be low or middling. The simulation will generate a finish date with a mixture of long, short, and middling durations. With cancelling out, the range around the total project duration becomes less in percentage terms than the interviewee indicated.

In order to use the plus or minus 10 percent ranges, for individual activities' ranges, and provide total project duration results that accurately reflect the input from the interviewee of plus or minus 10 percent, you need to correlate the individual activities' durations at 100 percent. Imposing a perfect positive correlation will cause iterations in which high durations or low durations will occur for all activities, together. It is the word "together" that indicates perfect positive correlation. As the simulation calculates finish dates, these high or low values for the detailed activity durations reinforce each other with 100 percent correlation and the total schedule results will mimic the interviewee's inputs.

The six-activity schedule, shown in Figure 4, is used to illustrate the need for correlation of uncertain durations during simulation.

In this case, with two paths of three activities each, the results without and with correlation clearly indicate that applying correlation to the activity durations causes the assumed percentage range on each activity to be carried over to the results for the entire schedule's duration:

- Applying 0.9, 1.0, 1.1 multipliers and no correlation yields overall schedule results of:
  - Optimistic is 328 days = **0.94** * 350
  - Pessimistic is 378 days = **1.08** * 350

Applying 0.9, 1.0, 1.1 multipliers with correlation set at 100 percent:
  - Optimistic is 315 days = **0.90** * 350
  - Pessimistic is 385 days = **1.10** * 350

Variation Caused by Known Unknowns – Risk Events

Unlike inherent uncertainty, estimating error and estimating bias, risk events are:

- Describable as "root causes" of variation in durations.
- Characterized by probability of occurring that is usually less than 100 percent.
- May be reducible by risk mitigation actions.

These are specified during risk interviews and implemented with estimated probabilities, impact ranges (three-point estimates of multiplicative factors) if they occur, and the activities or cost elements they influence. By modeling specific risks, the analyst can prioritize these risks for effective mitigation actions.

Identifiable or discrete risks, however they are implemented in the risk model, are similar to the category of special cause variability described by Walter Shewhart and W. Edwards Deming. Unlike common cause variability, special cause variation is caused by known factors that result in a non-random distribution of output. Special cause variability is also referred to as "exceptional" or "assignable" variation. Special cause variation is a shift in output caused by a specific factor, such as environmental conditions or process input parameters. It can be accounted for directly and potentially removed and is a measure of process control [10].

Variation determined by risk events, specifically identified as root causes of variability if they occur on the project (and during a Monte Carlo simulation), are described with a standard risk structure; including their cause, the risk (generally a statement that includes "may") and the effect if they occur. Each identifiable risk event is described with its:

- Probability of occurring with some noticeable effect on the duration of some activities.
- Impact ranges stated in days or multiplicative factors to be applied to the activities' durations if the risk occurs.
- The activities affected by the risk, if it occurs.

Case Study Showing the Application of Uncertainty

A simple illustrative case study of a large capital project, the construction of an offshore natural gas production platform, will show the elements of:

- Inherent variability
- Estimating error
- Estimating bias
- Discrete risk events

The schedule is presented in Figure 6.

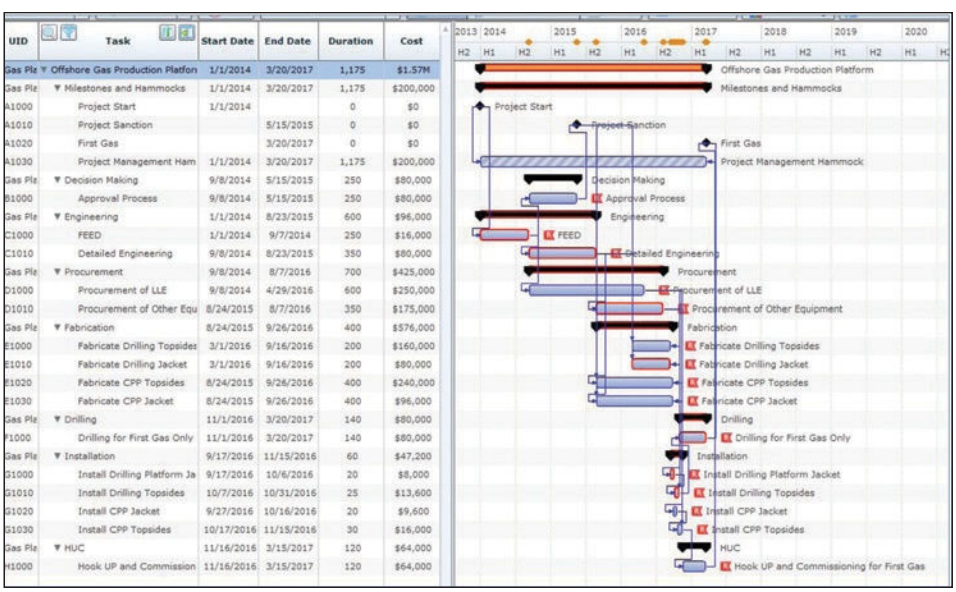


Figure 6 – Offshore Gas Production Platform Construction Project Schedule

Figure 7 – Specifying Inherent Variability to the Uncertainty Range

Figure 8 – Adding Estimating Error to Uncertainty

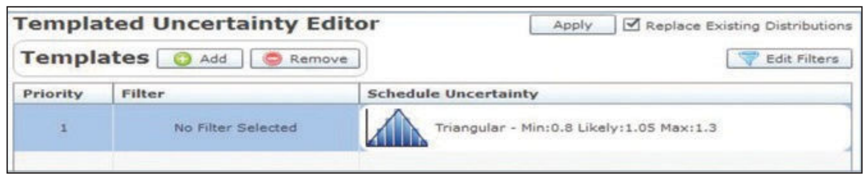


Figure 9 – Adding Estimating Bias to Uncertainty

Inherent variability is added by specifying ranges for uncertainty, in this case applying to all activities equally, as shown in Figure 7.

Next, estimating error is added, expanding the ranges of uncertainty as shown in Figure 8.

Finally, complete the modeling of uncertainty by changing the uncertainty factor to reflect inherent variability, estimating error and estimating bias. In some cases there are many risk events that are each small in impact and will not be modeled individually. These risks may be swept into the final uncertainty range, widening the range and making it more pessimistic since most of these risks are more threat than opportunity.

The results for each of these steps are shown in Figure 10. Obviously the parameters will differ for each project, but these are representative of data used on real projects.

In addition to specifying the range (low, most likely, and high) or (min, likely, and max) one has to also consider and implement a 100 percent correlation, so the results for the entire schedule will better approximate the input data from various interviewees.

The results for the simulations up to this point, representing various steps of uncertainty, are shown in Figure 10.

Notice in Figure 10 that the P-80 date is highlighted along with the increase in schedule contingency reserve needed from the scheduled finish date of 20 March 2017. Also highlighted is the calendar day spread between the P-5 and the P-95 dates. This range of days compared to the total duration of the project, (1,175 days), provides some measure of the adequacy of the input data and realism of the results. Finally, the probability of completing on or before the scheduled date is shown. The effect of perfect correlation is to widen both the opportunity and threat tails of the distribution so the probability of finishing on time rises to 32 percent.

Adding Risk Events

Discrete risk events need to be added to the risk model that contains

Adding Inherent Variability, Estimating Error and Estimating Bias						
		Scheduled Finish Date		20-Mar-17		
P-5	P-50	P-80	P-95	Delay in Months @ P-80	Cal. Days P-5 to P-95	Probability of Scheduled Date
Inherent Variability 0.9 - 1,0 - 1.1						
19-Feb-17	28-Mar-17	15-Apr-17	2-May-17	0.9	72	37%
Adding Estimating Error 0.8 - 1.0 - 1.2						
4-Feb-17	12-Apr-17	17-May-17	20-Jun-17	1.9	136	29%
Adding Estimating Bias 0.8 - 1.05 - 1.3						
25-Mar-17	17-Jun-17	2-Aug-17	13-Sep-17	4.4	172	4%
Inherent Variability, Estimating Error and Estimating Bias 100% Correlated						
1-Nov-16	17-May-17	30-Aug-17	29-Nov-17	5.4	393	32%

Figure 10 – Results for Each Step Representing Uncertainty

uncertainty. Risk events have a probability of occurring that is usually less than 100 percent, an impact range if they do occur, and an identifiable set of activities that they influence if they occur.

There are generally two ways to represent risk events:

- Risk register (term coined by many including in Oracle®Primavera Risk Analysis) by which the risks' impacts are characterized by days. A risk cannot have both opportunity and threat characteristics.
- Risk drivers (Risk factors in Oracle® Primavera Risk Analysis) by which the risks' impacts are described as distributions of multiplicative factors, not unlike those used above for uncertainties. Because the impact multiplier can be less than or greater than unity, a risk driver can have both opportunity and threat "tails," a smoother application than a risk register.

To this model, along with uncertainty, add seven specific risk drivers and one organizational risk driver assigned to all tasks, as shown in Figure 11. Each risk driver is specified with a TriGen function to offset anchoring and adjusting bias.

The results of simulation with uncertainty and risk events represented by risk drivers are shown in Figure 12. Notice that over seven months is added to the P-80 date and the spread from P-5 to P-95 almost doubles. These are illustrative data only, but are not illogical. The probability of meeting the scheduled finish date drops to 17 percent:

Modeling Unknown Unknowns – A Proposal

In this example, it is not know what these unknowns are or how significant they may be. Therefore, the author concludes from historical results that the knowledge of risk events and the size of uncertainty is limited and that more risks will be revealed as the project proceeds through its lifecycle. This results in a recommendation to repeat quantitative risk analysis periodically to identify the new risks, as well as retire existing risks.

In other words, experience tells that there exists unknown unknowns that will potentially affect the project. They will become known in the future as events unfold. Therefore, there is a need to reflect them in the analysis.

Since unknown unknowns can reliably be expected to occur on this project, a relatively conservative certainty target, e.g., the 80th percentile ("P-80") is recommended when discussing the need for schedule or cost contingency based on what is known about uncertainty and risk events that is revealed in the Monte Carlo simulation results. In other words, the author sees a need to be "ahead of the game" with respect to the risks that are known so that he is not "behind the eight-ball" already when unknown unknowns arise.

Interviews or workshops on risk often appear to focus on risks and uncertainties that are close-in or actually happening now. This myopia leads to insufficient consideration of risks that occur most likely in the "far future" of a project. In practical terms, this might mean that risks more than two years out from the interviews will not be brought to mind and included in the quantitative database for the risk analysis. Something needs to be done to represent these far future risks.

It is also arguable that unknown unknowns are not truly unknowable, but have not been thought of yet. This may be a result of a lack of attention to risks that may occur "down the road.:

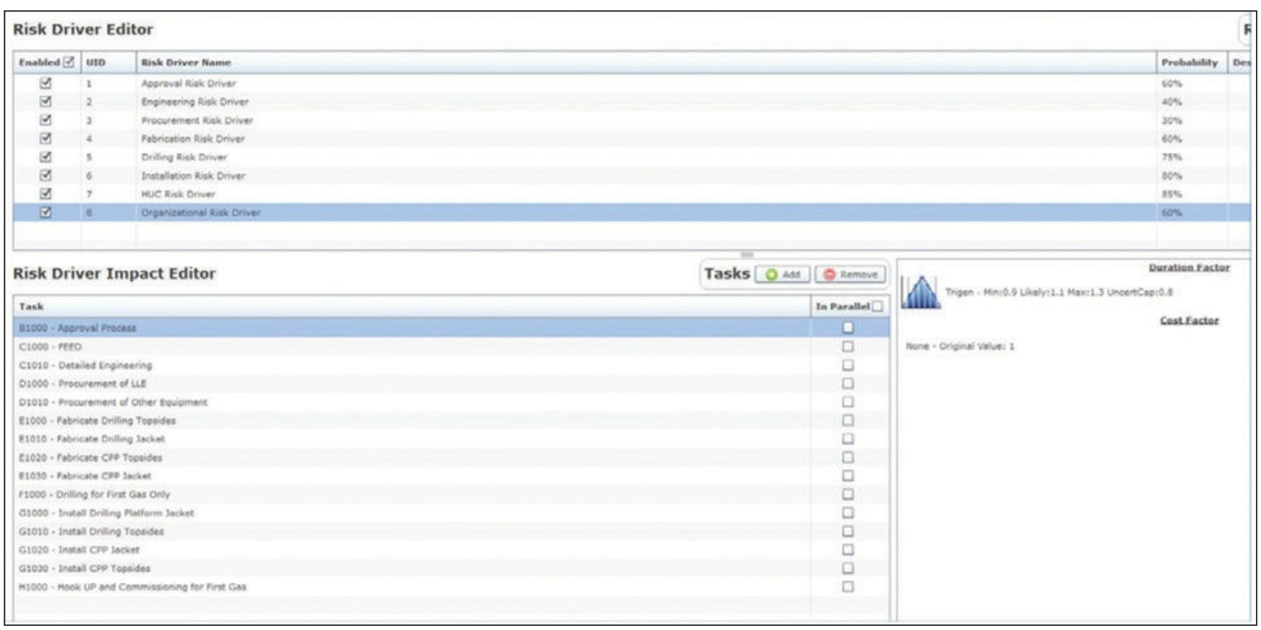


Figure 11 – Adding Risk Drivers to the Model

Adding Inherent Variability, Estimating Error and Estimating Bias						
		Scheduled Finish Date		20-Mar-17		
P-5	P-50	P-80	P-95	Delay in Months @ P-80	Cal. Days P-5 to P-95	Probability of Scheduled Date
Inherent Variability, Estimating Error and Estimating Bias 100% Correlated						
1-Nov-16	17-May-17	30-Aug-17	29-Nov-17	5.4	393	32%
Inherent Variability, Estimating Error and Estimating Bias 100% Correlated Plus 8 Risk Drivers						
26-Nov-16	22-Sep-17	19-Apr-18	19-Dec-18	13.0	753	17%

Figure 12 – Results from Adding Risk Drivers to the Model

If these risks are "unknown" right now, extra effort to focus on down-stream risk could improve the understanding of risks, including making known some "unknowns." Some of the risks can be identified with practice at thinking forward.

While unknown unknowns have no known specific cause, one can confidently anticipate that they will be revealed as the project execution continues. One can widen, beyond the ranges reported for the near-term uncertainty, the range of uncertainty that is applied to future activities to represent these unknown risks before one knows what they are. The degree of "widening" and the definition of "future" would be judgment calls for the project participants and the risk analyst. This idea of widening the uncertainty range for far future (within the project context) can be implemented as shown in Figure 13.

The degree of widening the ranges in the far future for the project is obviously a judgment based on inputs from many and experience with benchmarking. Notice that the new range for future uncertainty is quite wide, since it needs to encompass both uncertainty and risk events. The results in Figure 14 show adding about four months to the P-80 and other expected results.

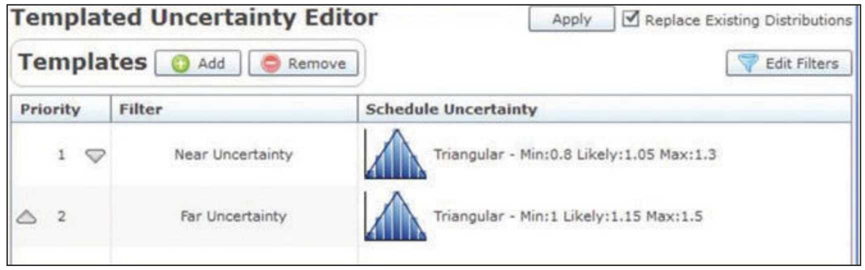


Figure 13 – Applying Wider Ranges Representing Unknown on Far Future Activities

Considering Unknown Knowns

Considering the discussion initiated by Secretary of Defense Rumsfeld, he missed one of the most important unknowns, the so-called "Unknown Knowns." However, people have really known for years that there are risks that cannot be discussed in public. These risks usually are detrimental to the project, contradict management's position, or are embarrassing. In some organizations these cannot be discussed in the halls, over coffee, or in risk workshops.

"Psychoanalytic philosopher Slavoj Zizek says that beyond these three categories there is a fourth, the unknown known, that which one intentionally refuses to acknowledge that he/she knows. German sociologists Christopher Daase and Oliver Kessler (2007) agree with a basic point of Rumsfeld in stating that the cognitive frame for political practice may be determined by the relationship between what we know, what we do not know, what we cannot know, but Rumsfeld having left out what we do not like to know" [11].

Many risk events are known by the project participants but they may be unable to discuss these in a public setting. This is confirmed because the risk register is always incomplete. The evidence of this fact is that in confidential interviews with project team members a number of new risks, risks that are not included already in the risk register are introduced with their parameters. Subsequent interviewees often contribute their

Adding Inherent Variability, Estimating Error and Estimating Bias						
		Scheduled Finish Date		20-Mar-17		
P-5	P-50	P-80	P-95	Delay in Months @ P-80	Cal. Days P-5 to P-95	Probability of Scheduled Date
Inherent Variability, Estimating Error and Estimating Bias 100% Correlated						
1-Nov-16	17-May-17	30-Aug-17	29-Nov-17	5.4	393	32%
Inherent Variability, Estimating Error and Estimating Bias 100% Correlated Plus 8 Risk Drivers						
26-Nov-16	22-Sep-17	19-Apr-18	19-Dec-18	13.0	753	17%

Figure 14 – Results With Wider Uncertainty Ranges on Far Future Activities

information on these new risks without objection. Many of these newly-introduced risk events are later found to be most important in determining the schedule risk results. Among these risk are the "Unknown Knowns" that which we intentionally refuse to acknowledge that we know or do not want to know.

Why are some of the most important risks not included in the risk register? The risk register is usually a formal document and is clearly described in the organization's handbook. Often there is someone assigned to manage the risk register. Sometimes expensive risk register software is purchased and maintained. Periodic risk workshops are conducted to amend, update, and correct the risk register. Yet, none of this matters if key risks never make it to the risk register.

The risks that are omitted are risks we know but are unwilling or unable to talk about. Ignoring these risks, or causing the corporate environment wherein these risks cannot be discussed, exhibits an ostrich approach to risk management.

The author views the incompleteness of the risk register to be the result of social pressures or group dynamics that limit debate in open workshops, including:

- Groupthink – people in groups often prefer unanimity, discourage dissent and make it difficult for people to raise new issues or voice an opinion different from that of the group.
- Moses Factor – some people will suppress their own ideas, adopting influential person's ideas instead, in a group.
- Cultural Conformity – The decisions match the group norms [2].

Risk workshops usually involve many people getting together for several hours and often include project managers. Workshops can be serious wastes of time with people not willing to be candid in discussing risk, or they are not able to talk much at all in a room full of other people, including the boss.

A more successful environment to gain the best data possible by Subject Matter Experts (SMEs) is the confidential interview approach. Confidential interviews bring people a degree of comfort not found in open meetings. In these interviews, new risks are discussed, whether they are hurtful to the project or embarrassing. People more openly talk about "the good, the bad, and the ugly" of a project in these interviews.

In the confidential one-on-one (or one-to-a-few) the interviewee(s) can talk in depth and with focus to someone and indicate a desire to know his/her opinion. Most interviewees open up and find it easy to discuss risks they cannot discuss in the team meetings, even the special purpose risk meetings.

In addition, interviewees appreciate the opportunity to talk about risk and that someone is listening, taking their viewpoints into account.

Some Evidence on Project Overruns

Quantifying unknown unknowns by widening uncertainty ranges on far future activities' durations is a "best guess" strategy, simply because we do not know what these unknowns are. As mentioned above, challenging the interviewees to think about future risks may reveal some of these unknown unknowns. Also, looking at the experience with other projects may help to calibrate the representation of unknown unknowns. At the least we should be somewhat modest in claiming any specific insight into unknown unknowns – that is one reason they are included in an expanded uncertainty range rather than as risk events.

Evidence from well-documented projects tells a startling story, at least as to cost results. Bent Flyvbjerg and associates have amassed some interesting findings: [1]

- Chunnel project was overrun by 80 percent (this was a commercial project but with heavy oversight and involvement of politics, regulations – e.g., safety).
- Great Belt Link – a bridge tunnel between east Denmark and Europe, was a 54 percent overrun.
- Oresund Link bridge between Sweden and Denmark was a 68 percent overrun.
- The Big Dig in Boston was a 196 percent overrun.

In addition, Flyvbjerg and associates found:

- Studies in Sweden on road and rail projects
  - o Road projects average 86 percent overrun, range +two percent to +182 percent
  - o Rail projects average 17 percent overrun, range -14 percent to + four percent
- US Department of Transportation rail projects
  - o Average overrun 61 percent, range -10 percent to +106 percent
- Aalborg University Denmark, nine of 10 projects underestimated
  - o Rail average overruns of +45 percent
  - o Tunnels/Bridges + 34 percent
  - o Roads + 20 percent

Flyvbjerg and colleagues concluded that cost underestimating is common. Coupled with overestimating the benefits, which are often "non-measurable, insignificant or even negative. This means that some of these projects should not have been approved. These are large-scale infrastructure projects, but projects in other industries appear to be similar or worse. Interestingly, there appears to be no learning about overruns – the activity of constructing large governmental projects has established equilibrium over time at substantial overruns"[1]. They believe that these results reflect both powerful incentives to underestimate to get projects started and weak disincentives to overrun.

A dissenting voice is that of Edward Merrow who has stated that: "Large projects have a dismal track record because we have not adjusted our practices to fit the difficulty that the projects present." Merrow was quoted in a 2012 AACE International Transactions article [3].

Finally, John Hollmann finds that research has shown that "the approximate ranges for accuracy or uncertainty around the reference amounts are as follows:

- P10: -32 percent to +8 percent (average about -9 percent).
- P50 or mean: 0 percent to +88 percent (average about 21 percent).
- P90: +34 percent to 190 percent (average about 70 percent)" [3].

The point of benchmarking against some real results is to put a check on the results of a risk model based on the CPM schedule, using data gathered at least in part by talking with SMEs and employing Monte Carlo simulation techniques. We can learn more by comparing results from two quite different methodologies and data collection approaches than if we just stick to one method.

Conclusion

This article presents four different concepts of knowns and unknowns and shows how they are, or could be, represented in a schedule risk model using uncertainty, risk events and confidential interviews.

Known knowns are those facts of which we are sure. They should be reflected in the baseline and current updated schedule. This is not new or controversial.

Known unknowns include several concepts of uncertainty and risk events. The uncertainties include:

- Inherent variability, which is known to exist even if its origins are non-specific. A probability distribution range of plus and minus multipliers helps to implement this concept, since inherent variability is 100 percent likely and thought to be irreducible.
- Estimating error is also known to exist and to depend on the quality and maturity of the information at hand at the time of creating the schedule. Again, at any one time this is 100 percent likely to exist so it can be incorporated in the probability distribution of impact multipliers applied directly to the durations of the activities.
- Estimating bias may or may not exist, but if it exists it has already occurred and forms the basis of the project CPM schedule that is used as a platform for the schedule risk analysis. Estimating bias may cause the pessimistic or high value of the three-point estimate to be asymmetrically further away from the most likely value than the optimistic or low value. In addition, the most likely value may not be the duration in the project schedule since that may have been biased, usually downward to accommodate management's requirements.

Having described setting the parameters of a three-point estimate-based distribution representing uncertainty, this article discussed the two issues of: using the TriGen function to adjust for a common underestimation of the ranges, and correlating the values of the uncertain activity durations. The interviewees who describe the plus and minus ranges have usually described their opinions about overall schedule variation rather than variation of individual activity durations.

The category of known unknowns that contains risk events is neatly addressed and handled by the use of risk drivers. In this article, the common failing of underreporting the risk impact, because of anchoring and adjusting bias, is addressed by creating TriGen functions that widen the ranges of the triangular distribution that are too narrow even in confidential risk interviews.

The new category of unknown knowns is introduced as a class of risks that is known about, but not spoken of in the context of public meetings such as risk workshops. This issue is handled by conducting data gathering in confidential interviews, and evidence that these interviews work to address unknown knowns is that many risks that do not appear in the risk register are discussed in those interviews.

It is not surprising that the method of handling unknown unknowns, applying wider ranges on far future activities, is least satisfying of all. done does expect there to be risks in the project that are revealed as the progress occurs, so the existence of unknown unknowns is well established, at least early in the lifecycle. They must be handled as uncertainty rather than as risk drivers since their nature is unknown at the point of analysis. It is suspected from experience that the interviewees talk at length and in much detail about the risks that exist today or are in the near future. It is suspected that if one works at expanding the time dimension of the interview discussion, that some of the unknown unknowns will become known. Alternatively, the author of this article proposes to implement an expanded range of uncertainty to take account of unknowns that are discovered in the future. This begs the questions of how much to expand the range and when the far future begins.

Finally, some sobering data from recent studies of Flyvbjerg and Hollmann about project cost overruns needs to be factored into the mix when explaining and calibrating unknowns of whatever type. Coming at project schedule results using quite different methods, in this case database examination of comparable projects, may improve the ability to understand the degree to which project performance does not match project plans.

REFERENCES

1. Flyvbjerg, Bent, Nils Bruzelius and Werner Rothengatter, 2003, **Megaprojects and Risk: an**

Anatomy of Ambition, Cambridge University Press, Cambridge England.
2. Hillson, David and Ruth Murray-Webster, 2005, **Understanding and Managing Risk Attitude**, Gower Publishing, Aldershot, Hants, England.
3. Hollmann, J., 2012, RISK.1027: Estimate Accuracy: Dealing with Reality, **AACE INTERNATIONAL TRANSACTIONS**, AACE International, Morgantown, WV.
4. Merrow, E. M., 2011, **Industrial Megaprojects**, John Wiley & Sons, Hoboken, New Jersey.
5. Moses, L., 1977, **Energy Information Administration Annual Report to Congress**, Volume 3, Administrator's Message, US Government Printing Office, Washington, DC.
6. Taleb, N. N., 2007, **The Black Swan: The Impact of the Highly Improbable**, Random House, New York.
7. Tversky, A. and D. Kahneman, 1974, *Judgment Under Uncertainty: Heuristics and Biases*, **Science**, Sept. 26, 1974, Science/AAAS, Washington DC.
8. http://jxb.oxfordjournals.org/content /60/3/712.full/ February 5, 2015.
9. http://www.isixsigma.com/dictionary/common-cause-variation/ February 6, 2015.
10. http://www.isixsigma.com/dictionary/variation-special-cause/ February 6, 2015.
11. http://en.wikipedia.org/wiki/There_are_known_knowns/ January 24, 2015.

EDITOR's NOTE

Figures 4,5,6,8,9,11,12,and 13 are screen shots from Polaris™ v.1.8, an integrated cost-schedule Monte Carlo simulation program developed by Booz Allen Hamilton.

ABOUT THE AUTHOR

Dr. David T. Hulett, is a principal with Hulett & Associates, LLC. He can be contacted by sending email to: david.hulett@projectrisk.com

FOR OTHER RESOURCES

For additional resources go to: www.aacei.org/resources/vl/

Do an "advanced search" by "author name" for an abstract listing of all other technical articles this author has published with AACE. Or, search by any total cost management subject area and retrieve a listing of all available AACE articles on your area of interest. AACE also offers pre-recorded webinars, an Online Learning Center and other educational resources. Check out all of the available AACE resources.