New Leaf
PROJECT MANAGEMENT
500 Thoreau Street
Concord, MA 01742 USA
+978.369.9009
info@newleafpm.com
www.newleafpm.com

(1 PDU for $9.95)

When you are ready to earn your PDUS:
Please return to www.NewLeafPM.com
- Click on the "White Papers" tab.
- Select "White Paper Power Quizzes for PDU Credit."

Good Estimates, Bad Biases, and QPM™ Games

How can we create a good work estimate despite the human biases in our judgment?

John M. Nevison, PMP
President, New Leaf Project Management

"The fault, dear Brutus, is not in our stars, but in ourselves…"
- Cassius, in William Shakespeare's *Julius Caesar*

Introduction

When we do a poor job of planning a project that involves a new technology or unfamiliar work, we can blame our bad estimates on the "difficulties of unfamiliar work." However, when our work estimates remain inaccurate for projects involving even familiar work, we must ask ourselves if the source of our difficulties lies closer to home. Experienced project managers know that skillfully estimating project work, both effort and duration, is always a challenging task.

It turns out we have two major difficulties with our problem of estimation. One because we are human and one because we are, as a group of practitioners, somewhat limited in our understanding of how to correctly compensate for our human tendencies.

In early 2011, New Leaf set out to produce a set of interesting games on the web that would teach some of the quantitative fundamentals involved in great project management. QPM™ (Quantitative Project Manager) was the result [1]. The first and biggest educational challenge in QPM was dealing with the subtleties of estimating project work.

Quantitative Games

QPM became a sequence of 15 lessons, each with a game or a quiz, each set up to award the successful player with a "stripe." The lessons explore estimation and other rarely understood foundations of familiar project practices, including:

- How to combine range estimates in useful ways
- Improvements to Agile's planning poker

© 2012-2104 New Leaf Project Management You may not post the PDF to a web page. You may, however, link to its source: www.newleafpm.com

- Simple ways to apply the critical path algorithm
- How to calculate convergence in a network logic diagram (without resorting to simulation), and, finally
- How best to use mid-project performance to anticipate where the project is heading

As we worked on the challenge of estimating project work, we discovered that range-based estimates, which had been dismissed by some critics of PERT, have been thoroughly reinstated by recent work on biases in human behavior spearheaded by the Nobel-Prize-winning psychologist, Daniel Kahneman. These discoveries about biases have rekindled an interest in how we can improve our project work estimates [2].

High or Low?

When we estimate project work, do we estimate high or low? Are we by nature cautious or bold?

We could estimate *high* because we are *cautious:*

- To give ourselves some freedom, some room to maneuver
- To protect ourselves against some uncertainty
- To make a commitment that we are sure we can keep
- Because we are under-confident in our knowledge about the work
- Because we are under-confident in our ability to do the work
- Because we are pessimistic in our general outlook

We could estimate *low* because we are *bold*:

- To give ourselves a challenging goal
- To expose ourselves to some uncertainty
- To make a commitment that will look good when we keep it
- Because we are overconfident in our knowledge about the work
- Because we are overconfident in our ability to do the work
- Because we are optimistic in our general outlook

Looking over the two lists, we might hope that the tendencies would cancel each other out. Unfortunately, research shows that, left to our unconscious instincts, we will estimate *low*.

The following discussion will review five biases that contribute to bad estimates, show how they affect our ability to plan project work, suggest ways to deal with them, and show where they appear in particular QPM games the reader can play to learn more.

The theme of this discussion will be, "Forewarned is forearmed." The hope is that as we become aware of these biases, we can engage in compensating behavior that will lead to a less-biased, improved estimate. Better estimates of cost and duration mean a better project plan, better project control, less re-planning, and greater project success.

© 2012-2104 New Leaf Project Management You may not post the PDF to a web page. You may, however, link to its source: www.newleafpm.com

Project Planning and Human Biases
Among the many biases that affect our judgment are five traps that can seriously affect project work estimation:

1. Being too sure of ourselves: The overconfidence trap
2. Seeing what we want to see: The confirming-evidence trap
3. Neglecting distributional information: The neglecting base-rate trap
4. The meaning of the mean: The regression trap
5. Slanting probabilities and estimates: The prudence trap [3]

The first four traps tend to generate work estimates that are *unrealistically low*; the fifth trap, *unrealistically high*. Our goal will be to end up somewhere in the *realistic middle*. We will examine each trap in turn, what to do about it in our day-to-day practice, and how it has been incorporated into the games in QPM [3].

The Overconfidence Trap: Being too sure of ourselves
When individuals are given tests and asked how confident they are in their answers, their score for "100% certain" answers is about 80%; their score for "80% confident" answers is 50%. Repeated experiments over different ranges confirm that, in general, we human beings are bold estimators because we are *overconfident* [4].

Overconfidence can also lead to *too narrow* an estimate of an interval that brackets some unknown number, such as the birth date of a famous person, or the distance between two given cities. When given 10 unknowns to bracket with a "90% confidence" interval, people succeed only 60% of the time. QPM Stripe 1 gives a player a chance to become aware of the overconfidence bias and, through repeated plays of a bracketing game, develop the ability to create a true 90% confidence interval.

So as a species we are afflicted with the bias of *overconfidence*. In our projects, overconfidence can mean overconfidence in our ability to do the work, overconfidence about our understanding of the work to be done, and even overconfidence in our ability to estimate the work.

When we decide to combat this overconfidence by declaring not a single value, but a high-low interval that we are 90% sure brackets the correct value, that high-low interval, before we get self-conscious about it, will also be too narrow! Our overconfidence bias is a true lesson in humility.

When estimating a project activity, the remedy for this trap is a true humility about our ability and our knowledge, followed closely by an active curiosity about a *reference class of similar work examples*. Such examples are often found in old project plans and in our experience on prior projects. By gathering relevant historical evidence, we temper our initial overconfidence with data about what has actually happened in similar circumstances.

QPM Stripe 2 asks the participant to think about data from past everyday activities, such as commuting to work. Stripe 3 considers some examples where there are enough past cases to construct a histogram of possible values for the estimate.

The Confirming Evidence Trap: Seeing what we want to see
When proponents and opponents of capital punishment are presented with two scientific studies, one with supporting data and one with opposing data, everyone, on both sides, emerges from their reading assignments even *more certain of their original position.* We human beings select evidence to confirm our beliefs and we ignore evidence that conflicts with our beliefs.

The confirming-evidence trap takes an estimator's natural pride of work (overconfidence) and adds the tendency to notice only reinforcing examples of how well the estimated work can be completed.

A good antidote to this trap is to remind ourselves, as we build our reference class of similar past work, *to pay attention to data that do not agree with our beliefs* about the work. In particular, as we examine our reference class of other project work, we try to assess, not just a high-low range of values, but *at least three* points of information:

1. What's the fastest this has ever been done? (Easy, because it agrees with our flattering self-image.)
2. What's the slowest this has ever been done? (Difficult, because it contradicts our flattering self-image and probably requires a little extra pessimistic push on our recollection.)
3. What, in between the extremes, is the most likely single value? (Likely to drift towards the flattering end unless supported by a larger number of examples in the reference class.)

QPM Stripes 4 through 8 offer the player several games to experience what it feels like to deal with three-point (and other kinds of) range-based estimates and then dynamically manage project performance using these estimates. Stripe 5 shows the user how to simulate a real project network with a pair of dice! Stripe 7 gives the user a spreadsheet tool that is a "flexible estimator" where the user can chose to apply one of four estimates (single point, 90% interval, 3-point triangular, or histogram) to an activity. All of these stripes encourage the player to look at data that might not comfortably fit with his or her beliefs.

The Neglecting Base-Rate Trap: Neglecting distributional information
When a publisher is asked about the future sales of a new novel on her list, she tends to extrapolate more from information about the particular book, its author, plot, and style than from the general statistics about the sales of similar novels. Repeated experiments confirm that we tend to focus on the *singular information* about the particular case under consideration and pay little or no attention to the base-rate data about the *distributional information* of outcomes in similar situations. In the words of the previous bias, one of the things we don't want to take time to see is statistics.

Here's an experiment that illustrates this tendency. A person is asked to identify "Steven's profession" as either a librarian or a sales rep. Steven is described as "Very shy and withdrawn, invariably helpful, but with little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure and a passion for detail." Most people, given this description, decide that Steven's profession is "librarian."

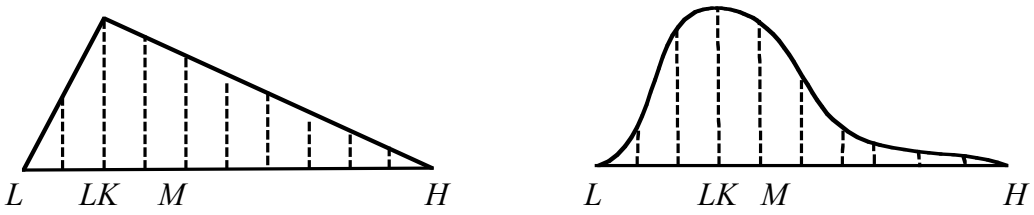
Even after people are told that in our society there are 100 sales reps for every male librarian, they will still choose librarian. We ignore the overwhelming odds that a small percentage of shy sales reps will significantly outnumber the infinitesimal number of male librarians and we continue to choose the wrong answer.

Our behavior illustrates the neglecting base-rate trap where we focus on *singular* data richly detailing one individual and neglect *distributional* data about the individual's group.

For our project work estimates, the way to avoid the neglecting base-rate trap is to ask two additional questions of the reference class (after we have good values for low, likely, and high). The questions are "What is the shape of the whole distribution and what is its weighted center (also called the mean)?"

For example, if the reference class is a histogram of past cases, we can calculate the mean and the variance of those points to find the weighted center. If we make a simplifying assumption that the reference class is shaped like a triangle continuously distributed over the range with a low of 2, a likely of 4, and high of 8, the distribution has mean a equal to (2 + 4 + 8) / 3 or 4.7. If we assume the distribution is shaped like a "beta" distribution, the mean is equal to (2 + 4 x 4 + 8) / 6 or 4.3.

Figure 1. Triangular and Beta Distributions.



Choosing the value of the mean automatically incorporates an assumption about the distribution of the estimates and is the antidote to the neglecting base-rate bias. QPM begins to consider the mean value in Stripe 2 and continues to develop the mean's importance in the games from Stripe 3 through Stripe 8.

The two distributions in Figure 1 are often associated with three-point range estimates in projects and they exhibit another noteworthy characteristic—they are asymmetric in shape. In both distributions [5] the mean, M, is *higher* than (to the right of) the "most likely," LK. QPM Stripe 3 deals with the roots of this asymmetry, while QPM Stripe 4 illustrates how distributions "shape-shift" from skewed individual activity distributions to converge on symmetric, normal project distributions.

QPM also carefully explains the mechanics of combining activity estimates into overall project estimates, beginning with the construction of a "98% reliable" target in Stripe 3.

The Regression Trap: The meaning of the mean

Over one hundred years ago Francis Galton discovered that tall parents have shorter children and short parents have taller children. Any group with an average that differs from the overall population will have offspring that, on average, are closer to the overall population's average. The offspring's average will "regress towards the mean" of the population (while a new and different group will be identified as the tall group in the offspring's generation).

Yet when we are informed that Group 1 consists of everyone who scored in the top 10% of History Test A and we are asked where Group 1's average score will fall when presented with History Test B (of equal difficulty), we predict "in the same top 10%." We forget that the old Group 1's average will regress towards the mean and probably be lower than the new Group 2 composed of those who scored in the top 10% on Test B.

The likelihood is that any group whose average was above (different from) the mean the first time will be closer to the mean the second time. (The Group 1 score may still be above the mean, but it will probably be below the high first score, closer to the mean, and, by definition, lower than the new Group 2's score.)

When estimating a project cost or schedule, in order to include an initial, intuitive guess in our final estimate, we need to counteract the regression bias by taking some final steps. After we have established a reference class, a range of values, an assumption about the distribution, and a mean value, the final steps begin with a look back at our initial guess and at our level of confidence in that guess. If our initial guess equals the mean, we are done. If the guess is off the mean, we should move it back towards the mean. (If we are very sure of our initial guess, we should only move it back a little bit; if we were wildly guessing, we should move it back most (or all) of the way to the mean.) In short, our final steps should *regress our initial intuitive estimate* towards the mean [6].

Because the games in QPM always choose the mean of the reference class data and *completely ignore any initial intuitive guess*, the game estimates are free of regression bias.

The Prudence Trap: Slanting probabilities and estimates
If our past estimates have been inaccurately low and we wish to avoid being wrong again, we may be tempted to "pad" our estimate a little. Our boss might do the same thing. Somewhere up the decision chain, someone decides there is too much padding in the figures and arbitrarily cuts the estimate. Prudence, "worst-case" anxiety, and then "fear of padding" lead to an estimate that has now become completely *unhinged* from reality.

The prudence bias leads us to *alter an estimate by an amount that is not supported by the evidence* of the reference class we have assembled. To compensate, we must keep our estimates *honest*, stay with what we know, and be sure that others know that our estimate is *not* padded.

To avoid padding in its games, QPM teaches the estimator how to use the variance of the distributions of project cost and schedule to develop a "98% reliable" target. These targets provide handy ways to communicate with stakeholders as a project unfolds.

Conclusion: Work in Process
In summary, in order to combat our inherent biases to make an overconfident, prejudicial, single-case-focused, regression-ignoring estimate we should follow a careful process in our day-to-day work:

1. Select a reference class (historical results of similar activities to establish base rates for this work).
2. Assess the distribution of outcomes by asking:
    a. What's the best this has ever been?
    b. What's the worst this has ever been?
    c. What's the most likely single value for this?
3. Find the shape of the distribution and the *mean* (the weighted center) of the distribution.
4. If you wish to include an initial, intuitive estimate:
    a. Position your intuitive, original, (optimistic?) estimate in the distribution.
    b. Assess the *reliability* of your original estimate. (It's very reliable, 0.9 [tomorrow's weather]; it's very unreliable, 0.1 [total runs in tomorrow's baseball game]).
    c. If your original estimate is not the mean of the distribution, adjust your estimate by regressing towards the mean.
        i. When your reliability is 0.3, correct, by going 70% of the way back to the mean; when 0.4, 60% and so on...
        ii. If in any doubt about your initial estimate, use the mean. [7]

The QPM games give players an entertaining and convenient way to gain experience by applying the above process to sample projects. The process itself lets us all counteract our innate human biases and produce less-biased work estimates. If we treat these less-biased estimates honestly (without padding), our project plan has

a much higher likelihood of being true, our project work can be better managed against the plan, our project results can arrive on a better schedule, and our stakeholders will be better served!

Notes

1. QPM™ is a registered trademark of New Leaf Project Management and is used with permission. QPM is available on the web at newleafpm.com/games/on-line-games/.
2. PERT critics include, in 2012, public comments on the contents of the *PMBOK Guide* 5th Edition and private communications to the author ("When people can't make one good estimate of an activity; why would you expect them to do better with three?"). Supportive remarks come from Lovallo and Fleming below.
3. The language for all the traps but "regression" is taken from Hammond, et al., below. Hammond's "base-rate" language was modified in this article. Regression is found in Kahneman, 1982.
4. The data on the experiments for all of these traps can be found in Kahneman, 1982. Several anecdotes are also recounted in Hammond, et al., with additional references.
5. Details on the Beta distribution, the Triangular distribution, and general work estimation are in Nevison.
6. This regressing towards the mean is described in Kahneman, 1982, and how to deal with regression is described in Lovallo.
7. A process similar to this is presented in Lovallo.

References

1. Hammond, John S., Ralph L. Keeney, and Howard Raiffa, (1999) *Smart Choices*, Boston, MA: Harvard Business School Press.
2. Fleming, Quentin W. and Joel M. Koppelman, (2010). *Earned Value Project Management, Fourth Edition*, Newtown Square, PA: Project Management Institute (PMI)
3. Kahneman, Daniel, (2011) *Thinking Fast and Slow*, New York, NY: Farrar, Straus and Giroux.
4. Kahneman, Daniel, Paul Slovic, and Amos Tversky (editors), (1982) *Judgment under uncertainty: Heuristics and biases*, New York, NY: Cambridge University Press.
5. Lovallo, Dan, and Daniel Kahneman, (July, 2003) "Delusions of Success: How Optimism Undermines Executives' Decisions," *Harvard Business Review*.
6. Nevison, John M., (2010) *Embracing the Dragon's Tail: Range Estimation of Project Cost and Schedule*, White Paper, Concord, MA: New Leaf Project Management. newleafpm.com/wp-content/uploads/2010/06/aWP%20EmbracingTheDragonsTail.pdf
7. Tversky, Amos and Daniel Kahneman, (October, 1992) "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, Volume 5, Number 4.

About the Author

John M. (Jack) Nevison, PMP, is the author of six books and numerous articles on computing and management. In his business career, Nevison has built and sold two businesses, managed projects, managed project managers, and served as both an internal and external consultant to Fortune 100 companies. He is a past president of the Mass Bay Chapter of the Project Management Institute (PMI®) and a past president of the Greater Boston Chapter of the Association for Computing Machinery (ACM). He was an author of the original *PMBOK® Guide* and led the team that revised Chapter 7 (Cost) for the *PMBOK® Guide, 5th Edition*. Nevison is a certified Project Management Professional (PMP), and a Phi Beta Kappa graduate of Dartmouth College.

You may not post the PDF to a web page. You may, however, link to its source: www.newleafpm.com