

Cost of Delay



The Goal – what to start next

Given a set of things we could do next, is one more advantageous economically to start first.

The Challenge – Doing one thing delays others

Every item has a different cost by being delayed. The cost will be a balance of lost value of other items, and how long they are deferred.

“The problem with any prioritization decision is [it is a] decision to service one job and delay another.” Don Reinertsen

The basic concept – balance \$ & time

$\$ = \text{Value lost due to delay} \times \text{delay duration}$

Same duration,
do > lost \$ first



Same lost \$,
do < duration first



Harder to decide
when both value
and duration
change



There are many ways that economically optimal work ordering can be derived. These rank from no attempt to find a more optimal order to computing the impact of delay on profitability over the products useful lifespan. The goal is to use the technique that is appropriate for the level of impact being wrong. For major decision with huge financial impact, use the techniques more to the right.

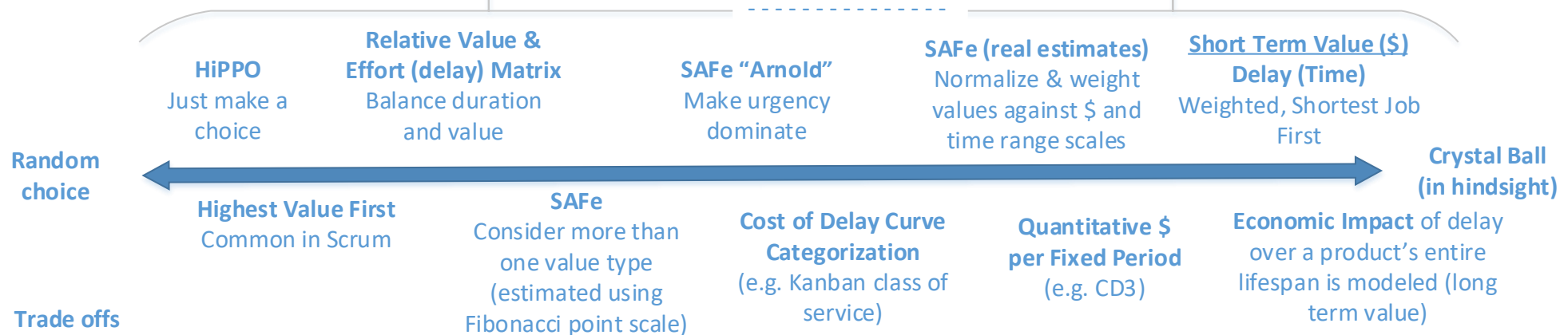
Two major approach categories: Qualitative, where proxy measures are used for potential value lost and duration to streamline analysis; and Quantitative where analysis is performed in currency and time. The more quantitative the analysis the more effort and research required. Balance the tradeoff between change of making a sub-optimal decision with the analysis time needed to be “sure.”

More Qualitative

(effort and value estimated using proxy measures)

More Quantitative

(effort in time, lost value is in currency of choice)



Chance of making a sub-optimal decision

Effort required to get optimal decision

Highest Value First: Common in Scrum

Scrum proposes starting the highest customer value work first. Some teams use a qualitative low, medium and high. Some attempt to estimate it in dollars. This is better than random ordering, but often leads to “Biggest liar wins.” It also doesn’t consider how long each item will take, meaning more value might be delivered in a number of smaller items that sum to greater value.

SAFe and SAFe “Arnold Mod”

SAFe (Scaled Agile Framework) uses a weighted shortest job first balancing technique. It uses subjective measures and approximates optimal starting order using the following formula (highest first):

$$\frac{\text{Value} + \text{Criticality} + \text{Risk Reduction or Opportunity Enablement}}{\text{Job Size}}$$

Joshua Arnold offers the following modification to make time criticality more dominant:

$$\frac{\text{Criticality} \times (\text{Value} + \text{Risk Reduction or Opportunity Enablement})}{\text{Job Size}}$$

Economic Models and WSJF

Donald Reinertsen in his book “Principles of Product Development Flow” offers a variety of scheduling techniques. The most popular is Weighted Shortest Job First where optimal starting order is calculated using economically modeled inputs (dollars and size) starting the highest first using the formula:

$$\frac{\text{Cost of Delay}}{\text{Duration}}$$

SAFe Weighted Shortest Job First

The Scaled Agile Framework proposes an ordering system based on Don Reinertsen's Weighted Shortest Job First (WSJF) principles.

More info: <http://www.scaledagileframework.com/wsjf/>

Proposed features are assessed on multiple value and size axis. It is proposed to use a Fibonacci scale of values for each input parameter (1,2,3,5,8,13,20). The full calculation process is described as -

1. Rate each parameter against the other features using the scale: 1,2,3,5,8,13,20. Do one column at a time, and calibrate the lowest value to be a "1" - each column MUST have one "1"
2. Calculate the WSJF value for each column using the formula shown below
3. Do the feature that has the HIGHEST WSJF value first if possible

SAFe's Weighted Shortest Job First formula (upper), and a typical data capture table (lower)

$$\text{WSJF} = \frac{\text{User-Business Value} + \text{Time Criticality} + \text{Risk Reduction} + \text{Opportunity Enablement Value}}{\text{Job Size}}$$

Feature	User-Business Value	Time Criticality	RR OE Value	Job Size	WSJF

SAFe Weighted Shortest Job First Variations (un-sanctioned)

Some variations of the basic SAFe formula and technique have evolved to make the computation more likely to match ideal.

1. "Arnold Mod"

In an email thread conversation between Martin Burns and Joshua Arnold, the suggestion of making Time Criticality more dominant was suggested. This solves the theoretical problem that something "Critical" might score a lower WSJF due to a high business value or risk reduction or opportunity enablement or a low size. Martin noted that these rarely occur due to earlier decision processes, but this suggestion would solve these even if they slipped through.

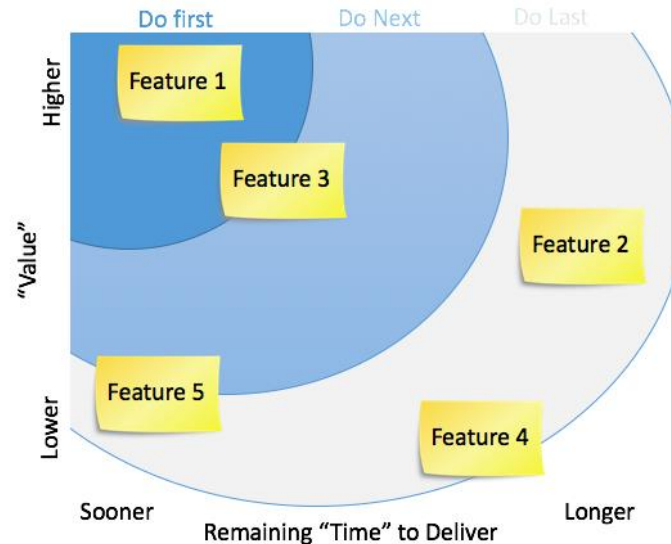
$$\text{WSJF} = \frac{\text{Time Criticality} \times (\text{Value} + \text{Risk Reduction} + \text{Opportunity Enablement})}{\text{Job Size}}$$

2. Scale and Weight the Arguments (solve the mathematical issues of different argument units)

The use of Fibonacci numbers for the input arguments is an attempt to make the estimates relative to each other for the same input argument, but there is a chance that the magnitude is different for each value. For example, a "5" in value might be \$100,000, but a "5" in risk reduction might be \$500,000. If we just added them as the original SAFe formula says, the result makes little intuitive sense. To correct, either scale the values to normalize across arguments, or multiple each Fibonacci value by a weighting multiplier to correct the magnitude mismatches.

Matrix Techniques (Arnold's and Eric's techniques)

Have the teams place a feature on a matrix of delay time and value whatever definition is appropriate for your company's context. Do higher value, shortest delivery time items first. Erik Willeke uses a variation where ONLY product owners can move items up or down (indicating higher or lower value) and the development team left and right (shorter or longer to deliver).



DOs

- Encourage better economic decisions
- Use the lightest analysis method to get a decision
- Use these methods to help have conversation about what value means to each feature or product
- Find better ways to measure and estimate
- Involve a diversity of viewpoints on both value and delay.

DON'Ts

- Use complex analysis on small items. Ideally only for features and larger
- Ignore delivery time or its proxy job size; this leads to sub-optimal ordering
- Create an arms race for "value" by prioritizing on it alone (biggest liar wins syndrome)
- Use the highest paid person's opinion if at all possible, offer alternatives!

REFERENCES

Donald Reinertsen:
Books: Principles of Product Development Flow has great Cost of Delay ideas and concepts.
Video: Cost of Delay: Theory & Practice with Donald Reinertsen <https://www.youtube.com/watch?v=OmU5ylv7vRw>

SAFe:
<http://scaledagileframework.com/wsjf/>

Joshua Arnold's blog:
<http://blackswanfarming.com/category/cost-of-delay/>

Chris Matts Blog:
<https://theriskmanager.wordpress.com>

Troy Magennis:
Spreadsheets for Cost of Delay <http://bit.ly/SimResources>
Blog: <http://focusedobjective.com/blog/>

How value is lost due to a delay – Urgency Profiles

Value erodes differently for different products and markets. The most commonly calculated is just the loss of revenue on the front end because of being late. But, the lost value can be much more than that if the delay causes a permanent erosion of market share, or if the market window is short. It can be difficult to calculate the longer term value erosion, but it may be significant.

