Chapter 3
Benefit Points for the Project

> Vision without action is a dream. Action without vision is simply passing the time. Action with vision is making a positive difference.
>
> JOEL BARKER

Abstract We start by looking at projects and their most abstract product elements, the epics, and show how to estimate their benefit using benefit points. Then we show how to sort epics according to a benefit-cost index to help decide the order in which to put epics into releases. Instantiating points with a monetary value provides added means of prioritizing and determining when to stop sending epics into construction. We show two modes of estimating benefit: one where the purpose is to fulfil a given goal (confirmatory mode), and the other where the purpose is to explore where to set the goal (exploratory mode).

3.1 Overview

In a development project, *product elements* are represented by requirements specifications in some form, such as user stories. The *benefit criteria* (Fig. 2.1) are then given by *project objectives*. Project objectives express the organization's reasons for initiating the development project in the first place. The purpose of the project is to fulfil these objectives.

Figure 3.1 shows a project as part of the agile fractal (Fig. 2.6), with its high-level requirements specifications in the form of epics. At this stage, the epics are to be assessed on project objectives and have not yet been distributed to project releases. The epics' benefits are estimated according to their assessed contribution to the objectives. This is the *effect* relation in Fig. 3.1. The system under development is expected to have an impact on business processes. This impact is effected through the system's functionality, designed with the intent to enable users and other systems to perform tasks in an overall better or more efficient manner.

The project objectives are, in turn, assessed on planned returns. This is the *worth* relation in Fig. 3.1. The worth relation has nothing to do with the system's functionality. Rather, the relation expresses the expected gain in value the objectives imply, once they are fulfilled.

J. E. Hannay, *Benefit/Cost-Driven Agile Software Development*,
Simula SpringerBriefs on Computing 8,
https://doi.org/10.1007/978-3-030-74218-8_3

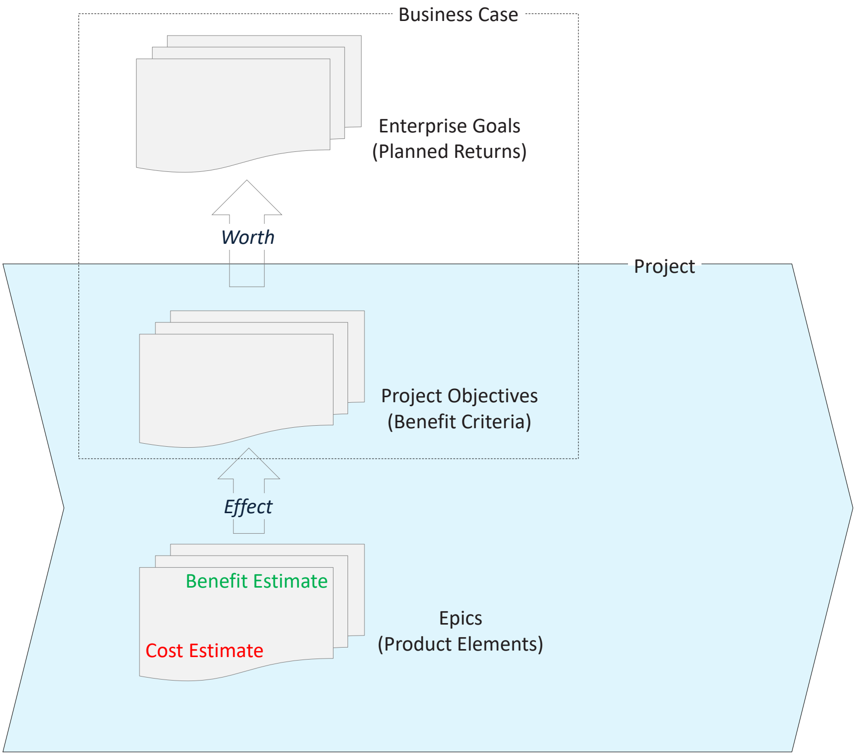

Fig. 3.1 A project has specific objectives. An epic's *effect* is assessed in terms of its contribution to the objectives. Objectives have different *worth* in terms of their contributions to planned returns. We state that *benefit* = *effect* × *worth*.

Then, *benefit* = *effect* × *worth*; that is, the benefit of an epic is its effect on project objectives times the objectives' worth in returns. Assessing *effect* and assessing *worth* are fundamentally different tasks, and we make a point of assessing these two relations one at a time. Indeed, the two assessments can be made by different stakeholder groups, and carried out in any sequence or in parallel. One should not attempt to combine the two assessments into one. Assessing the effects of epics on objectives, while simultaneously adjusting for the various objectives' worth exceeds most people's cognitive capacity. Because projects usually lack conceptual clarity when it comes to benefits management, projects often end up assessing benefit in a way that effectively combines and collapses these two steps.

Figure 3.2 (bottom portion) shows examples of epics in a development project for a public service organization.

Product elements can also be expressed in terms of related notions such as *minimum viable change* and *minimum viable transformation*, which emphasize the change in business processes induced by product elements.

Case 4. An internal revenue administration recently implemented changes to the way salary information is registered, processed, and reported. Since this involved substantial alterations to end-user processes and internal data processing, the product owner decided to deploy rather quickly the simplest possible version of the new web-based functionality to users in a limited region with uncomplicated life and work situations, in other words, a minimal viable change. This piece of meaningful functionality providing immediate benefit allowed the project to learn early from a low-risk real-life deployment. After that, further functionality was rolled out to successively larger portions of the population.

Returns:
Ret1: Reduced number of man-hours
Ret2: Reduced number of compensations
Ret3: Improved public image of the organization

Objectives:
Obj1: Reduce average case processing time by 30%
Obj2: Reduce number of wrong case decisions by 90%
Obj3: Reduce the average interaction time between the applicant and the application processor by 70%

Epics:
E1: *As* an applicant *I can* secure my identity in the application process *by using* MyID module *to* authenticate myself
E2: *As* an applicant *I can* increase speed & accuracy of the application process *by using* MyID module *to* autofill personal data
E3: *As* a case processor *I can* find all relevant information for a case *by using* the Cross Search module *to* retrieve applicant information from all relevant and permissible data sources in a single search
E4: *As* a case processor *I can* receive alerts when deadlines are approaching *by using* the Reports module *to* finish cases on time and avoid complaints
E5: *As* a case processor *I can* view graphical trends over cases per status *by using* the Reports module *to* increase planning and motivation
E6: *As* a division manager *I can* manage my division's productivity *by using* the Reports module *to* view statistics to monitor the time and quality of case processing
E7: *As* a returning applicant *I can* obtain an overview of earlier applications *by using* the Reports module *to* obtain an overview of my history with the public sector
E8: ...
...

Fig. 3.2 Example of epics, objectives, and returns (public sector).

3.2 Project Objectives

A project should have designated *objectives* that express the project's intended effects on the organization's business processes. Figure 3.2 (middle portion) shows examples of objectives in a development project for a public service organization.

We will present two modes of benefit estimation. One mode is where the objectives are set that the project must fulfil. We call this the *confirmatory mode*. The other mode is where stakeholders try to determine what the project will be able to deliver on the given objectives. We call this the *exploratory mode*. As with all top-down and bottom-up tactics, it is sensible to combine the two modes in an interleaving manner, especially in an environment geared to project learning and adaptation.

We will explain the principles of benefit estimation for the confirmatory mode, because things are simpler in that mode. Then, we will explain how to estimate benefit in the exploratory mode.

3.3 Effect Points: Benefit Points for the Effect Relation

For the effect relation, benefit estimates are assigned to epics according to how much each epic is perceived to contribute to objectives, in terms of relative benefit points. For the effect relation, benefit points are called *effect benefit points*, or *effect points* for short.

Since there are several objectives, the assignment of effect points is more complex than assigning story points for cost. Figure 3.3 shows a table with effect points for eight epics assessed on three objectives. As a rule, all epics should be assessed on one objective before moving to the next, as indicated by the vertical lines in Fig. 3.3. This is because objectives can have different metrics (time, money, quality, etc.), and special attention is required to perform relative assessments across metrics.

	Obj1	*Obj2*	*Obj3*	Total
E1	16	8	12	36
E2	25	35	8	68
E3	25	4	7	36
E4	10	13	3	26
E5	1	5	31	37
E6	6	9	8	23
E7	15	13	12	40
E8	2	13	19	34
Total	100	100	100	300

Fig. 3.3 Effect benefit points (effect points) assigned by distributing 100 points per objective. Benefit points provided by the stakeholder group are shown on a white background. The totals in the shaded area are computed automatically by your tool.

In this example, stakeholders have used a technique in which they distribute 100 linear points for an objective over the epics. This *parts of the whole* technique is suitable in the confirmatory mode: 100 points represent an objective's complete fulfilment, and they can be distributed among epics according to their relative contributions. You can also use open-ended scales, such as the Fibonacci sequence familiar from planning poker, in the confirmatory mode, but the calculations are slightly more complicated. Consult Section 3.13 on this topic later.

It is essential that you only concentrate on the effect relation at this stage: do not be concerned with the fact that objectives can represent different levels of worth! That consideration belongs to a different exercise, which we will address shortly. For more on the technique of assigning benefit points, read about *benefit poker* in Section 3.15, and find out more on the issue of *multiple objectives* in Section 3.16.

3.4 Planned Returns

Having explicit, preferably measurable objectives for one's project is one of many signs of organizational maturity. The assignment of benefit points to product elements in terms of those objectives is a first step to handling a project's generation of business value.

However, the project objectives represent the project's estimated effects, and therefore coexist for the duration of the project. To link the project to the organization's long-term goals, one must link project objectives to the business return, as planned in strategic goals. For example, a planned return for the public service organization example above could involve the goals in Fig. 3.2 (top portion).

3.5 Worth Points: Benefit Points for the Worth Relation

A project's objectives, once fulfilled, are expected to yield various degrees of return for the enterprise. This is the worth relation. *Worth benefit points* (or *worth points*) are used to express estimates of worth. The benefit criteria (Fig. 2.1) are then the planned returns above. Figure 3.4 exemplifies the technique of distributing 100 points per return: reaching *Obj3* is assessed to yield on *Ret1* as much as the two other objectives combined (see the *Ret1* column in Fig. 3.4).

For the public service example, this means that stakeholders assess that a 70% reduction in the average interaction time between the applicant and the application processor will reduce man-hours to the same extent as reaching the other two objectives together.

Returns are outside a project's domain of argument, and the project assumes the goals expressed in returns as given. For the project, the worth relation is therefore confirmatory, by definition. When we discuss portfolios in Chapter 4, we will see the worth relation in an exploratory mode as well.

	Ret1	*Ret2*	*Ret3*
Obj1	25	29	33
Obj2	25	43	40
Obj3	50	29	27
Total	100	100	100

Fig. 3.4 Worth benefit points (worth points) produced from distributing 100 points per return. Benefit points provided by the stakeholder group are shown against a white background. The totals and weights in the shaded area can be computed automatically by your tool.

Again, techniques such as the distribution of 100 points are suitable in the confirmatory mode. For the worth relations, this implies that one plans for the project's objectives to fulfil the returns entirely; in other words, the returns represent exactly the expected business value of the project. A planned return, say, *Reduced number of man-hours*, could be a strategic goal spanning several projects, initiatives, and programmes in an enterprise, but, here, only the part of the return that the project is expected to fulfil is considered.

3.6 Monetary Returns

Effect points represent estimates of the system's effect on business processes, and worth points represent the return in terms of strategic goals from changing those business processes. Both effect points and worth points are relative assessments. In particular, worth points express the relative degree to which project objectives contribute to returns. If one now estimates a project's returns in monetary terms, one can determine the project's estimated monetary benefit. Project returns can also pose as a strategic management goal, further emphasizing the confirmatory mode.

Suppose, then, that project stakeholders and strategic management assess that the project objectives, once fulfilled, will yield monetary returns as follows: *Ret1*, 40 million; *Ret2*, 14 million; and *Ret3*, 22.5 million. The objectives' worth points then imply that the project's objectives *Obj1*, *Obj2*, and *Obj3* are estimated to contribute

	Ret1	*Ret2*	*Ret3*		Weight
million:	40	14	22.5	Total	Project
Obj1	25	29	33	21.50	0.28
Obj2	25	43	40	25.00	0.33
Obj3	50	29	27	30.00	0.39
Total	100	100	100	76.50	1.00

Fig. 3.5 Returns are given monetary value (in millions of one's favourite currency). For each objective, one calculates the column denoted 'Total', by multiplying the monetary values by the objective's expected proportions of contribution and summing the results. For example, for *Obj1*, we obtain (40 * 0.25) + (14 * 0.29) + (22.5 * 0.33) = 21.50.

21.5 million, 25 million, and 30 million, respectively, to the total of 76.5 million (the 'Total' column in Fig. 3.5). Thus, the project's objectives, once fulfilled, contribute unevenly to the project's return. This is due to objectives contributing differently to returns, as expressed in their worth points, together with the different estimated worth of the project returns.

3.7 Balanced Effect Points

The fact that some project objectives are worth more than others must be reflected in the way the project prioritizes the backlog.

The 'Weight' column in Fig. 3.5 shows the weights of the objectives according to their contribution to returns. When objectives contribute unevenly to returns, a benefit point with respect to one objective will represent a different unit of benefit than a benefit point given with respect to another objective. To keep things manageable, we balance the number of benefit points so that a benefit point always represents the same amount of benefit, regardless of the objective.

Quite simply, multiply the effect points for an epic by the relevant objective's weight; for epic *E1* on *Obj1*, 16*0.28=4.48. We can then define a *balance* function as

$$\mathsf{balance}(BP_p, w_c) = BP_p * w_c \tag{3.1}$$

where BP_p is the number of benefit points for product element p, and weight w_c is the weight of criterion c. So if BP_{ij} is the number of benefit points for epic i on objective j and w_j is the weight of objective j, the general formula for balancing effect points is

$$\mathsf{balance}(BP_{ij}, w_j) = BP_{ij} * w_j \tag{3.2}$$

Figure 3.6 shows the resulting balanced benefit points for our example.

	Obj1	*Obj2*	*Obj3*	
Weights:	0.28	0.33	0.39	Total
E1	4.48	2.64	4.68	11.80
E2	7.00	11.55	3.12	21.67
E3	7.00	1.32	2.73	11.05
E4	2.80	4.29	1.17	8.26
E5	0.28	1.65	12.09	14.02
E6	1.68	2.97	3.12	7.77
E7	4.20	4.29	4.68	13.17
E8	0.56	4.29	7.41	12.26
Total	28.00	33.00	39.00	100

Fig. 3.6 Effect points, balanced according to the worth of objectives.

	Obj1	Obj2	Obj3	
Weights:	0.28	0.33	0.39	Total
E1	13.44	7.92	14.04	35.40
E2	21.00	34.65	9.36	65.01
E3	21.00	3.96	8.19	33.15
E4	8.40	12.87	3.51	24.78
E5	0.84	4.95	36.27	42.06
E6	5.04	8.91	9.36	23.31
E7	12.60	12.87	14.04	39.51
E8	1.68	12.87	22.23	36.78
Total	84.00	99.00	117.00	300

Fig. 3.7 Effect points balanced according to the worth of objectives and normalized to 300 points in total.

If you want to keep the total number of effect points in the project (300 in this example) constant in your tables (for cosmetic reasons), you can multiply by the ratio of the desired total number of benefit points (300 here) by the current total number of benefit points (100 here); for epic *E1* on *Obj1*, 16*0.28 * 300/100 = 13.44. We can define a *normalize* function as follows:

$$\mathsf{normalize}(BP_p, BP_{\mathsf{desired\ total}}, BP_{\mathsf{total}}) = BP_p * (BP_{\mathsf{desired\ total}}/BP_{\mathsf{total}}) \tag{3.3}$$

where BP_p is the number of benefit points for product element p, $BP_{\mathsf{desired\ total}}$ is the desired total amount of benefit points, and BP_{total} is the current total amount of benefit points. Thus, the formula for normalizing the amount of balanced effect points $b_{i,j} = \mathsf{balance}(BP_{ij}, w_j)$ for epic i on objective j is

$$\mathsf{normalize}(b_{ij}, BP_{\mathsf{desired\ total}}/BP_{\mathsf{balanced\ total}}) \tag{3.4}$$

where $BP_{\mathsf{balanced\ total}}$ is the total number of effect points after balancing.

Balancing and normalizing should be carried out automatically in your spreadsheet or project management tool. Figure 3.7 presents the resulting normalized balanced benefit points for our example, where keeping the total amount of benefit points (300) illustrates how the original points in Fig. 3.3 are redistributed according to the objectives' worth.

3.8 Cost Estimates: Size Points

Story points are routinely assigned for estimating cost in projects, and we assume procedures for doing this, such as planning poker, are known. However, we will make a few remarks in the context of benefit/cost management.

Benefit manifests itself after deployment; therefore, to obtain a sensible benefit-cost measure, the cost estimates should include post-deployment costs in addition to

	SP
E1	8
E2	8
E3	3
E4	5
E5	13
E6	13
E7	5
E8	8
Total	63

Fig. 3.8 Size points (SP).

development costs. We will use *size points* for this. Traditionally, story points reflect development cost only. However, life cycle cost is often assumed to be proportional to, or linearly dependent on, development cost (for more details, see Section 3.17 and e.g. [27]). Under that assumption, size points can be assigned if they are story points, since the relative proportions between story points remain the same for development and life cycle costs. Our methods apply, regardless of that assumption. However, under that assumption (and when it is warranted), some of the methods can take on a simpler form. In any event, for our running example, we assume the size points presented in Fig. 3.8.

3.9 Benefit-Cost Index

One can now immediately calculate the benefit point-to-size point ratio in Fig. 3.9 (left) to obtain a relative benefit-cost measure. The effect points are obtained from Fig. 3.7, and the size points are from Fig. 3.8. Size points can be divided by benefit

	BP	SP	BP/SP
E1	35.40	8	4.43
E2	65.01	8	8.13
E3	33.15	3	11.05
E4	24.78	5	4.96
E5	42.06	13	3.24
E6	23.31	13	1.79
E7	39.51	5	7.90
E8	36.78	8	4.60
Total	300	63	4.76

	BP	SP	BP/SP
E3	33.15	3	11.05
E2	65.01	8	8.13
E7	39.51	5	7.90
E4	24.78	5	4.96
E8	36.78	8	4.60
E1	35.40	8	4.43
E5	42.06	13	3.24
E6	23.31	13	1.79
Total	300	63	4.76

Fig. 3.9 Benefit-cost index. The ratios (BP/SP) of effect benefit points (BP) to size points (SP) are presented in the left panel, and sorted in descending order in the right panel.

points because both types of points are on a so-called *ratio scale*[1]. It is common to use nominal schemes – such as MoSCoW [25], which produces four categories of importance (textitMust have, *Should have*, *Could have*, and *Won't have*) – to assess benefit. In that case, benefit estimates cannot be divided by cost. To obtain a benefit-to-cost measure from MoSCoW, one could order the product elements by increasing cost within each category and then order the backlog by selecting the ordered elements in the Must have, then Should have, Could have, and Won't have categories. However, it is entirely possible for an element in a less important category to have a higher benefit-cost ratio than a given element in a more important category, due to low cost. Without a sound measure of benefit-cost provided by ratio scales, one would not become aware of such incidents.

There are several useful things that can be done with a benefit-cost index. If one wanted to realize maximum benefit relative to cost early, one would consider putting epic *E3* into construction first. Figure. 3.9 (right panel) shows the sorted epics, with those with the highest benefit-cost index at the top. We will investigate this and other ways to use benefit/cost estimates in later chapters.

3.10 Instantiating Points with Money

Points-based estimates are relative estimates, where monetary value is abstracted away. Practice and research suggest that it is easier to perform comparative judgements (one is larger than the other), rather than judgements on spot values.

An additional, powerful aspect of using relative sizes, such as benefit points and size points, is that one can assign actual monetary values to points, according to

	Benefit	Cost	Benefit/Cost
E3	8.45	1.80	4.70
E2	16.58	4.80	3.45
E7	10.08	3.00	3.36
E4	6.32	3.00	2.11
E8	9.38	4.80	1.95
E1	9.03	4.80	1.88
E5	10.73	7.80	1.38
E6	5.94	7.80	0.76
Total	76.50	37.80	2.02

Fig. 3.10 Benefit/cost. The same as Figure. 3.9 (right panel), but with effect points instantiated at 1 BP = 0.255 million and size points instantiated at 1 SP = 0.6 million.

[1] Scales come in several flavours. A *nominal* scale categorizes items by name, with no ordering. An *ordinal* scale puts on ordering on items, without stating distances between them. An *interval* scale orders items with fixed distances; two items classified as a '1' and a '2' have the same difference in magnitude as two items classified as a '2' and a '3', but a '4' is not double that of a '2' (examples are the Celsius and Fahrenheit temperature scales. A *ratio* scale has a defined zero point, which enables multiplication and division.

current knowledge. Figure 3.10 shows the results of Figure 3.9 with effect points instantiated at 1 BP = 0.225 million and 1 SP = 0.6 million. The monetary value (0.225 million) of an effect point is set by dividing the total benefit budget (76.50 million) by the number of effect points assigned to the project (300 BP). [2] The monetary value representing the life cycle cost of a size point is set at 0.6 million, say, based on structured stakeholder meetings and past experience from earlier projects with similar characteristics. For example, structured discussions could have established the development cost of a size point at 0.3 million. Then a linear model of post-deployment costs might suggest that the life cycle cost is twice the development cost. Thus, the life cycle cost estimate is 37.8 million, and the life cycle benefit estimate is 76.5 million. For this example, these values could be the initial estimates for the business case, prior to project learning. However, one can instantiate points with alternative values that reflect an initiative's current understanding or different scenarios. We will see this practice in action later.

With monetary values, benefit and cost have the same denomination. With the values set as above, it is evident that, according to initial estimates, epic *E6* has a benefit-cost ratio below one, which means that this epic, as a whole, should not be put into construction, since it will return less benefit than it will cost. Depending on a project's expectation levels and the level of risk the project is willing to take on, one might want to look out for *E5* as well. Its expected benefit is only about 40% more than its cost. Thus, benefit-cost deliberations can help one decide not only the order in which to construct product elements, but also when to stop construction.

3.11 Soft Returns

Return *Ret3* in the example in Fig. 3.2 is a typical qualitative, or 'soft', return. It does not directly refer to quantifiable measures. Since qualitative returns could be an essential part of business value, it is important to be able to include them in our scheme. Another example could be *Ret4*: *increased information infrastructure capability in society*. Such expected returns could be more important than quantitative financial ones, for example, in terms of political justification for initiating a development project or in terms of environmental and ethical sustainability goals.

The problem is that such returns can be very hard to quantify. Sometimes explicit quantification in terms of the monetary value of qualitative returns is required by law, such as in government-funded development projects, where there are obligations to follow socioeconomic models for the analysis of societal benefit. However, insisting on the hard quantification of qualitative values could be perceived as practically impossible and lead to the omission of such returns. In line with satisficing rather than optimization [44] and simplicity, we propose a method for implicitly quantifying soft returns, the *model for integrating soft and hard returns on invest-*

---

[2] In Section 4.3, we will discuss how one might set the total benefit budget.

ment, or *MISHRI*. The idea is the same as that presented by [5] for a slightly different context . So, how did we assess that return to be worth 22.5 million?

The entire methodology in this book is based on small steps that can be overcome by human cognitive resources. The required expert estimations are based on relative comparison, which is also what we recommend to quantify qualitative returns. For this, one needs to fix the value of at least one return, say, *Ret1*. One can now ask how important *Ret3* is relative to *Ret1*. If it is equally important, its monetary value should be set at 40 million; if it is less important, the same question can be asked relative to *Ret2*. One could also determine that *Ret3* is more important than *Ret2*, but closer to *Ret2* than to *Ret1* by, say, 10%, which implies a monetary value of 22.5 million. In other words, the quantitative returns can be used as markers for comparing qualitative returns.

Relevant stakeholders should be involved in this process, and one can use similar techniques as for the other expert estimates in our approach.

This inclusion of soft returns means that one can take into account their influence on project objectives. Soft returns will thus influence the backlog order. Later, one can choose whether to include soft returns in the actual return calculations. This might not always be appropriate, because there will not necessarily be any actual cash flow from soft returns. We leave that discussion for later. It is easy to include and exclude soft returns (and compare their effects). For *Ret3*, we simply set its value to zero and determine how the automatic calculations in your tool change. Figure 3.11 shows how not considering *Ret3* produces a different distribution of effect points on the epics, and therefore different ordering in terms of the benefit-cost index.

The relativistic approach to integrating soft returns above is designed to be non-intrusive in daily work as a simple, good-enough approach to an inherently difficult problem. In contrast, there are comprehensive approaches to quantifying planned returns that are far more elaborate, such as that of [14], but they will require a great deal of effort. One should be aware of both approaches.

	BP	SP	BP/SP
E3	32.34	3	10.78
E7	39.24	5	7.85
E2	61.56	8	7.70
E8	38.34	8	4.79
E4	23.46	5	4.69
E1	35.52	8	4.44
E5	46.20	13	3.55
E6	23.34	13	1.80
Total	300	63	4.76

Fig. 3.11 The benefit-cost index when *Ret3* is not taken into account. Compared to Fig. 3.9, *E7* and *E2* have changed places, as have *E4* and *E8*.

	Obj1	*Obj2*	*Obj3*	Total
E1	13	5	8	26
E2	21	21	5	47
E3	21	2	5	28
E4	8	8	2	18
E5	1	3	21	25
E6	5	5	5	15
E7	13	8	8	29
E8	2	8	13	23
Total	84	60	67	211

Fig. 3.12 Effect points obtained by assigning numbers from the Fibonacci sequence.

3.12 Effect Points in the Exploratory Mode

We have illustrated the main idea of benefit points using the technique of distributing a fixed number of points (100 in the examples). An alternative is to use open-ended scales, where one assigns points without assuming that the sum must be a certain number.

If you are familiar with planning poker, chances are that you will have used Fibonacci numbers to assign story points in an open-ended fashion. You can use the Fibonacci sequence for benefit points as well, and in this section, we will quickly go through the same steps as above to assign effect points, but now using the Fibonacci sequence in an open-ended fashion. It is possible to use the Fibonacci sequence as a fixed scale as well, and there will be examples of that later. Figure 3.12 shows our example with eight epics and three objectives, where numbers from the Fibonacci sequence have been used to assign the effect points. Again, epics are assessed against one objective at a time.

Whereas the technique of distributing 100 points prompts one to conduct an assessment relative to the total (100), using open-ended scales puts more emphasis on the direct relative assessment between items. The reason is that there is no global target (the upper bound of 100) to relate to. So, in the example, epic *E1* has been estimated to contribute substantially less than *E2* and *E3* to objective *Obj1*, but substantially more than *E4* and equally to *E7*. For objective *Obj2*, epic *E4* is assessed to contribute as much as *E5* and *E6* combined.

Now, since the scale is open ended, the effect point totals for the objectives may very well differ, as it does in Figure 3.12. This result can be interpreted in two ways: (A) the differences are an artefact of the estimation method or (B) the differences reflect a perception that the objectives are fulfilled to different degrees.

In the exploratory mode, (B) is the relevant interpretation.[3] There are various ways in which the project could be in an exploratory mode. We will mention a few possibilities.

[3] For interpretation (A), see Section 3.13.

3.12.1 Exploring the Effect of Epics

The project may want to determine what it can realistically achieve in terms of given objectives. For example, if the total number of effect points for an objective is substantially lower than for the other objectives, this could indicate that the stakeholder group thinks the epics do not have the potential to fulfil that objective to the full extent stated. This should prompt the project to re-evaluate the epics and perhaps redesign them so that they do fulfil the objectives.

3.12.2 Exploring the Feasibility of Objectives

The project, in the exploratory mode, could also start questioning the objectives themselves. This would initiate a discussion with the stakeholders responsible for defining project objectives.

Case 5. A new web-based customer solution was to be developed in an organization that provides services for handling intellectual property rights. Epics were specified and benefit estimated using planning poker cards. The resulting effect points under each objective (prior to normalization) differed to such a degree that the project leader questioned whether the benefit estimation group thought the system under development would be able to fulfil the planned objectives. The objectives were therefore revised, and benefit estimation reinitiated.

A more deliberate use of the exploratory mode than the one above would be to assess the epics' benefit with the explicit intent to determine the effects of the planned system. This would amount to exploring and determining what the objectives should realistically be, rather than setting objectives at the outset, as in the confirmatory mode. In such a case, one could start with rudimentary objective formulations, such as 'Reduce wrongful payments', without specifying by how much. The group of stakeholders can use effect points in benefit poker sessions as a means to discuss the effects of epics and eventually arrive at concrete objectives, such as 'Reduce wrongful payments by 70%'.

3.12.3 Working in the Exploratory Mode

In the exploratory mode, brainstorming-type discussions can be useful. To inform the discussion, effect point assessments can be used informally to reveal stakeholders' perceptions of the effect of epics or the viability of the objectives.

	Obj1	*Obj2*	*Obj3*	
Weights:	0.28	0.33	0.39	Total
E1	11.09	4.96	9.53	25.58
E2	17.92	20.84	5.95	44.71
E3	17.92	1.98	5.95	25.86
E4	6.83	7.94	2.38	17.15
E5	0.85	2.98	25.01	28.84
E6	4.27	4.96	5.95	15.18
E7	11.09	7.94	9.53	28.56
E8	1.71	7.94	15.48	25.12
Total	71.68	59.54	79.78	211

Fig. 3.13 Effect points obtained by assigning numbers from the Fibonacci sequence.

However, it is also possible to prioritize backlogs and, indeed, use all the other techniques in this book in the exploratory mode. Figure 3.13 shows the effect points from Table 3.12 balanced according to the objectives' worth, using Equation (3.2), and normalized to a total of 211 points, using Equation (3.4). The different objective totals indicate (under interpretation B) that the objectives are fulfilled to different degrees. Note that, here, the objectives' worth and returns are given in a confirmatory mode. In other words, the objectives are assumed to fulfil the returns in full, and the monetary value of worth points is also given, but the degree of the epics' fulfilment of objectives is under exploration and the monetary value of effect points is also unknown.[4]

3.12.4 *Partial Fulfilment of Objectives*

It could also be the case that one's project is not presented with project-specific objectives, but, instead, more general objectives. Then, the intention is not the project's total fulfilment of objectives. In this case, Figs. 3.12 and 3.13 express the epics' partial fulfilment of objectives.

However, when open-ended scales, such as the Fibonacci sequence, are used, the semantics for partial fulfilment are not obvious. For example, in Fig. 3.13, one would have to fix the fulfilment degree for at least one of the totals. For example, if we manage to determine that *Obj2* is $x\%$ fulfilled by its approximately 60 effect points, we can calculate the remaining degree of fulfilment for the other objectives; if Obj_i has a total of BP_i effect points, its degree of fulfilment is approximately $BP_i/60 * x\%$. Alternatively, the absolute value of an effect point could be given. This can be the case if, after some time, organizations settle on conventions analogous to those in planning poker for cost estimation: through extended experience [7], stakeholders can recognize a product element as, for example, a typical five or a two. In other words, benefit point amounts become universal quantities applicable across projects

[4] We will look at *worth* points in the exploratory mode in the next chapter.

in or, perhaps, even across organizations. In that case, the degree of fulfilment is given directly by the effect point total of an objective relative to the objective's total worth points, which expresses its worth when totally fulfilled.

3.12.5 Closed Scales in the Exploratory Mode

If the intention is, indeed, the partial fulfilment of objectives, then using a *parts of the whole* assessment (distributing 100 points, say) could be easier than using open-ended scales, such as the Fibonacci scale. In that case, one would have to use an absolute assessment rather than a relative assessment: one would still distribute points from a pool of, say, 100 points among the epics, but one would have to evaluate each epic's absolute contribution to the objective. If an objective then receives a total of, for example, 44 points of 100, this would presumably indicate a fulfilment of 44% of that objective by the project's epics.

3.12.6 Ending Up in the Confirmatory Mode

In practice, a smooth combination of the exploratory and confirmatory modes is likely the most sensible. At the end of such a process, project objectives should arise that are to be met in full by the system under development. In other words, the exploratory mode should result in epics and objectives that the project addresses in the confirmatory mode.

We also promote the use of project-specific objectives and the semantics of total fulfilment. Even when provided objectives that are not project specific, one can derive project-specific objectives by determining what part of the general objectives the project will actually fulfil.

A common argument in favour of objectives not being specific to a project is that benefit occurs through synergy between multiple projects. This is, of course, true, and it is a good thing to acknowledge potential dependencies and the importance of a holistic perspective. However, it is also important to be able to express the benefit of each part. This is crucial for and at the heart of thinking in terms of minimum viable products (MVPs). MVPs are supposed to yield integral benefit, and that integral benefit should be asserted. The project is an organizational unit, and it is necessary to assert what that unit is capable of in terms of the benefit of its MVPs alone. Synergies with other projects' MVPs are the business of portfolio management, which is the topic of the next chapter.

	Obj1	*Obj2*	*Obj3*	Total
E1	10.88	5.86	8.40	25.14
E2	17.58	24.62	5.25	47.45
E3	17.58	2.34	5.25	25.18
E4	6.70	9.38	2.10	18.18
E5	0.84	3.52	22.04	26.40
E6	4.19	5.86	5.25	15.30
E7	10.88	9.38	8.40	28.66
E8	1.67	9.38	13.65	24.70
Total	70.33	70.33	70.33	211

Fig. 3.14 Effect points equalized for the total fulfilment of objectives (and normalized to 211 points in total). The points are automatically computed by your tool.

3.13 The Confirmatory Mode with Open-Ended Effect Points

Finally, it is, of course, possible to use open-ended scales in the confirmatory mode. This is highly relevant to projects whose participants are already accustomed to using the Fibonacci numbers for cost estimation. If the intention is to estimate in the confirmatory mode, then differences in effect point totals per objective must be viewed as an artefact of the estimation process (interpretation A above): when not distributing parts of the whole, it is generally hard and distracting for stakeholders to ensure equal benefit totals in the end.

To neutralize this unintended difference, we simply equalize the benefit points so that the objective totals are equal, that is, we divide by the total number of benefit points for the objective. For example, the equalized benefit points for *E1* on *Obj1* in Fig. 3.12 are 13/84 = 0.15. We define the following *equalize* function:

$$\mathsf{equalize}(BP_{pc}, BP_c) = BP_{pc}/BP_c \tag{3.5}$$

where BP_{pc} is the number of benefit points for product element p on benefit criterion c, and BP_c is the total number of benefit points on criterion c. So, if BP_{ij} is the number of effect points for epic i on objective j and BP_j is the total amount of effect points on objective j, then the formula for equalizing effect points is

$$\mathsf{equalize}(BP_{ij}, BP_j) = BP_{ij}/BP_j \tag{3.6}$$

Figure 3.14 shows the effect points of our example equalized. The points are also normalized to 211 points in total by $\mathsf{normalize}(e_{ij}, 211, BP_{\mathsf{equalized\ total}}$ for $e_{ij} = \mathsf{equalize}(BP_{ij}, BP_j)$, where $BP_{\mathsf{equalized\ total}})$ is the total number of effect points after equalizing.

Then, Fig. 3.15 shows the balanced and normalized effect points for our example, using Equation (3.4).

Therefore, Fig. 3.16 (left panel) shows the epics sorted as in Fig. 3.9, but now with Fibonacci-based effect points. Figure 3.16 (right panel) has effect points instantiated at 1 BP = 0.36 million and 1 SP = 0.6 million. With 211 total effect points, the amount per effect point is different than in Section 3.10, where the total number of effect points was 300.

We can also compare Figs. 3.16 and 3.10. The values are comparable, but not equal. The differences are an artefact of using two different scales.

3.14 To Sum Up...

We have introduced benefit points for epics, called effect points. We have also introduced benefit points for objectives, called worth points. Using simple methods, you can assign benefit points based on a project's business case, using stakeholder knowledge and project expertise. This comprises a *core practice* alongside story point, or size point, estimation. Now, since you can assign both cost and benefit estimates to your product elements, you have the basics to monitor and learn from your project, to work towards generating as much benefit as possible, in addition to controlling cost.

A key feature to this core practice is a loosely coupled approach that allows a focus on one relation at a time. You are to focus on the relation between epics and objectives, disregarding the relation between objectives and returns, and to focus on the relation between objectives and returns, without having to think about epics. The combination of your assessments of the two relations can, and should, be generated automatically by the project management tool the project is using.

In contrast, assessing an epic's contribution to an objective while taking into account that objective's contribution to various returns and reflecting all this in the number of benefit points for the epic is hard. Trying to do all of that for several objectives is close to impossible. Yet, in practice, this is precisely what projects set out to do; not deliberately, but because they lack clear concept of benefit. For similar reasons, it is important to clearly delineate cost and benefit as separate concerns

	Obj1	*Obj2*	*Obj3*	
Weights:	0.28	0.33	0.39	Total
E1	9.18	5.75	9.88	24.80
E2	14.83	24.13	6.18	45.13
E3	14.83	2.30	6.18	23.30
E4	5.65	9.19	2.47	17.31
E5	0.71	3.45	25.94	30.09
E6	3.53	5.75	6.18	15.45
E7	9.18	9.19	9.88	28.25
E8	1.41	9.19	16.06	26.66
Total	59.30	68.95	82.75	211

Fig. 3.15 Effect points equalized for the total fulfilment of objectives and balanced against the objectives' worth (and normalized to 211 points in total). The points are automatically computed by your tool.

	BP	SP	BP/SP
E3	23.30	3	7.77
E7	28.25	5	5.65
E2	45.13	8	5.64
E4	17.31	5	3.46
E8	26.66	8	3.33
E1	24.80	8	3.10
E5	30.09	13	2.31
E6	15.45	13	1.19
Total	211	63	3.35

	Benefit	Cost	Benefit/Cost
E3	8.45	1.80	4.69
E7	10.24	3.00	3.41
E2	16.36	4.80	3.41
E4	6.28	3.00	2.09
E8	9.67	4.80	2.01
E1	8.99	4.80	1.87
E5	10.91	7.80	1.40
E6	5.60	7.80	0.72
Total	76.50	37.80	2.02

Fig. 3.16 Left: Benefit-cost index (BP/SP), with effect benefit points (BP) divided by size points (SP), sorted in descending order. Right: Benefit/cost, where the points are instantiated at 1 BP = 0.36 million and 1 SP = 0.6 million, sorted in descending order.

when providing estimates, by using, for example, size points and benefit points. Several existing methods do not explicitly support separating these concerns. The result is, again, a confounding of concepts, with ensuing confusion as to how to proceed with benefit-cost deliberations.

We introduced our approach to over 500 IT professionals in a triannual industry workshop on agile management. When participants first inadvertently attempted to estimate benefit without a clear picture of the objectives and returns, they expressed frustration over having to keep track of large numbers of factors at the same time. After being encouraged to concentrate on one relation at a time, they found that complexity disappeared and perceived the task as easy.

Jumbled concepts and a lack of clarity regarding the estimation task one is currently undertaking are not unusual. We regularly witness, in projects and larger development programmes, how notions akin to objectives, returns, and various metrics are confounded. This seems to create a dull confusion, halting effective benefits management. Although you might want to use other notions for goals than the objectives and returns in this book, we encourage you to adopt a disciplined approach to those notions and to be deliberate about exactly what you are estimating at a given time.

Although relatively new, the concepts presented in this paper have started to emerge in various organizations. Several projects in the public and private sectors have used benefit points to estimate the contribution of epics to business objectives, and subsequently used this for backlog prioritization. MISHRI has turned out to be a particularly popular technique, since it has made it possible for project leaders to include soft returns when presenting business cases for senior management and prioritizing backlogs. The general feedback from project members so far is that the benefit estimation process yields improvements over earlier practice, particularly in terms of a better understanding of project objectives and a clearer perception of the expected value of project deliverables. Benefit estimation also contributes to aligning project and business resources with respect to the impacts to expect from project deliverables.

Throughout this discussion, those portions of tables that you are required to provide estimates for have white backgrounds. Portions that are automatically calculated by your tool (e.g. Excel), have shaded backgrounds. Note that a modest number of expert estimates need to be provided and that they are not complicated measures, but are intended to capture the project's knowledge that is currently available.

The remainder of this chapter contains optional sections, which can be skipped in a first reading and consulted if needed. They contain material that addresses more frequent questions people have asked us when teaching or using these techniques. For example, see Section 3.19 for comparisons with other, related techniques, and Section 3.20 for more on the underlying principles of our techniques.

3.15* Benefit Poker

The key to assigning benefit points is to assess how much you think each epic contributes to the project's objectives. Here, we describe how to adapt the familiar practice of planning poker to a game of benefit poker. We illustrate this with effect points.

A benefit poker session could proceed as follows. A group of stakeholders estimates the relative contributions to an objective, one epic at a time. Each stakeholder bids a number face down, after which everyone reveals their bid simultaneously. The stakeholders with the highest and lowest bids are prompted to express their grounds for their bids. In this way, different assumptions and perspectives on the product element (and the objective) are highlighted. Nuances in understanding, knowledge, experiences, and ambitions contribute to useful clarifications and refinements. A host of group processes will likely be ongoing in such sessions, perhaps not all of them beneficial, but the rationale is that the positive effects of such poker sessions still outweigh the negative.

Bid rounds continue until the bids converge towards common agreement. In our experience, three rounds often suffices. If the bids still deviate substantially, the product owner can choose the average bid or the majority (if six of eight have identical bids, say). The resulting number represents the benefit points for the epic on the given objective. The group then turns to the next epic in the backlog and estimates its relative contribution to the same objective as before by repeating the bidding procedure.

It is common in planning poker to use a standard card deck with a slightly revised Fibonacci sequence, namely, 0, 1/2, 1, 2, 3, 5, 8, 13, 20, 40, 100 [4]. Grenning's [15] original paper on planning poker used a set of cards with the sequence 1, 2, 3, 5, 7, 10, ∞. The author also stated that the participants should feel free to use sums by showing two cards at once. This is also a practice we have used successfully. The important thing is not the Fibonacci numbers as such, but that the values are on a ratio scale and that the scale enables good differentiation between estimates.

Benefit poker can be used for both *parts of the whole* assessment (percentages, distributing 100 points, etc.; see Section 3.3) and open-ended scales (Section 3.12). The Fibonacci sequence can be used in both cases. To distribute 100 points, say, use the deck of cards above, with 100 as the highest number. To distribute 100 linear points, use cards from zero to 100, perhaps in intervals of one, five, or 10.

The argument for using the Fibonacci sequence would be to adapt the state of practice in cost estimation to benefit estimation. There is, however, no evidence yet to determine which scale provides better accuracy and reliability in assessments. It has been argued that the Fibonacci sequence is favourable due to what is known as the *Weber-Fechner* law [8]: as magnitudes increase, it becomes harder to distinguish between them. In fact, differences between magnitudes must increase exponentially for our senses to be able to detect the differences. Use of the Fibonacci sequence would then facilitate differentiation.

On the other hand, one study [46] suggests that the use of the Fibonacci sequence leads to lower estimates, on average, than the use of a uniformly distributed scale (e.g. 1, 2, 3, 4, 5, 6, ...), possibly due to the *central tendency of judgement* effect [24], where assessors tend towards the middle value of the perceived pool of possible values. In a standard deck of Fibonacci-style planning poker cards, as above, the middle value is five; that is, substantially lower than the middle value of a zero–100 linear scale. In cost estimation, where the general tendency is to give estimates that are too low, this can exacerbate the situation. In benefit estimation, it is not yet known if, or in what direction, estimates tend to deviate from actual values. If the general trait is overoptimism, one would expect benefit estimates to tend to be too high. However, even if the Fibonacci sequence can counter this tendency, its use for that purpose would ostensibly be for the wrong reason if due to the effect above.

3.16* One Combined Objective

Assessing epics against a number of objectives can seem quite complex. In practice, effect point estimation can be quite rapid. The stakeholder team will likely need a few moments to get calibrated to the scale it is using (perhaps starting with a reference epic), but once at cruising altitude, our experience is that it takes only a couple of hours to assess 10 to 20 epics on four to six objectives.

However, an alternative to considering several objectives is to estimate the effects of epics on some notion of a single objective. This practice is common today, where, for example, issuing one or several pluses (+) or minuses (-) to pieces of functionality against some potentially unspecified notion of benefit is a prevalent mode of operation.

Ostensibly, there are pros and cons to both approaches. The consideration of specific objectives allows one to think in more detail, but substantially increases the complexity of the benefit point estimation process. On the other hand, considering all objectives as one single, perhaps fuzzy entity can mean that you, as a stakeholder, are not really able to use your expertise and knowledge of the domain properly, even though the estimation process is substantially less complex.

There are theoretical grounds for choosing the first, more complex approach. Judgement and decision making theories predict that people will be affected by a host of unconscious biases that are likely to affect judgement in ways that can neither be predicted nor controlled [30, 17]. These biases add considerable noise to judgements. However, if one is able to use *task-specific* knowledge at key points in the judgement process, one should be able to boost the conscious elements in the judgement process, so that the decisions are the results of knowledge to a greater extent [38]. This is a case for strengthening the signal of a conscious knowledge-based process over the noise of unconscious biases. Consideration of each objective in turn stimulates that conscious signal.

Empirically, a controlled experiment indicates that the first approach generates estimates with less inter-rater variance than the second approach. This phenomenon could be a manifestation of lower noise, as theorized above. Additionally, less variance between job performers is an indication that a task has been defined such that expertise both is applicable and can be built [31, 7].

3.17* Life Cycle Cost Estimation

In Section 2.4, we stated that, to obtain a sensible expression of benefit/cost, the estimates of cost must take into account the life cycle of the product element, not only its construction. Size points (Section 3.8) reflect life cycle costs, and a simplifying assumption is that life cycle costs can be computed as proportions of the estimated *build cost*. In this section, we provide simple heuristics for doing so.

Here, we define the estimated build cost b as the estimated cost of development and unit testing of a specific product element. Thus, b is typically a value that can be verified after sprints and the basis of sprint burndown charts. Then, to arrive at an estimated *construction cost* for the product element, we need to add hours necessary for design, integration testing, documentation, and ceremonies, which will depend on the organization's development methods and standards. Experience from several large public sector organizations suggests that this sum ends up in the neighbourhood of $2b$.

To arrive at a *release cost* estimate for the product element, work related to product ownership, architecture, management, and operations must also be accounted for. This will again depend on how development is organized in the enterprise. Experience tells us that this release cost can amount to somewhere around $6b$. This is also called the *investment cost* of the product element, which can be verified at the end of a release.

To arrive at a *life cycle cost* estimate, we must now take into account two more cost drivers:

- Work related to teaching and motivating the end users and other stakeholders and
- Work related to maintenance (bug fixes, simple changes, and software component upgrades) after the first deployment.

The cost of these drivers can be estimated as proportions of the investment cost. Again, these proportions depend on the methodology and the organization. Let i be the investment cost. In one public agency, the first driver above turns out to be, on average, approximately $0.15i$, while the second driver is stipulated to vary over four years in production, as follows: $0.15i$ for the first year and $0.12i$, $0.11i$, and $0.10i$ for the following years. Using these rules of thumb yields a life cycle cost over four years of $9.78b$. For other public agencies, the factors are similar to these, but not identical.

The exact factors by which to multiply the build cost will vary according to the type of project, and the above factors should be considered merely as examples. The take away is that the life cycle cost can be estimated as proportional to the build cost.

Of course, there will be uncertainty in all of this. However, although the life cycle cost estimate as a whole will have a slow evaluation cycle, one can adjust the factors involved after sprints and releases according to incremental experience, thus removing some of the uncertainty.

Hardware and other infrastructure cost drivers might have to be computed separately. For example, one might distribute these cost elements at either the project level or the enterprise portfolio level down to the product elements in question.

The main thing to bear in mind is that the life cycle cost estimates should be computed for the same time period as the benefit estimates. We return to periodization and the concept of present value in Chapter 7.

3.18* Negative Benefit

Product elements can have a negative effect on objectives. Consider again our example project in Fig. 3.2, where *Obj1* is 'Reduce average case processing time by 30%'. Suppose the stakeholders and product owner want to include the following new epic *E9* in the backlog: 'As a security officer in the agency, I want to perform a check of the applicant for a certificate of good conduct before granting the application'. The epic might not be that costly, but it will impact the average case processing time in a negative manner. The project can consider other epics to compensate for this negative impact, so that *Obj1* will still be met.

Objectives can also add negative worth to returns. Consider again the example project in Fig. 3.2, where we now introduce *Obj4*, 'Case processing should fully meet the new quality and security standards'. Objectives such as this can be costly and, moreover, conflict with other objectives and impact returns in a negative manner. For our specific example project, the stakeholders could decide that, to meet this objective, the agency will have to allocate resources to the relevant departments, and they could estimate that *Obj4* will have a negative impact on *Ret1* ('Reduced number of man-hours').

Since benefit points are purely relative estimates, negative benefit can be handled by assigning monetary benefit point values so that benefit points below a certain threshold represent a negative monetary value. For example, if using a 100-point scale, one could set the zero point at 50 benefit points and set 1 benefit point at 3.7 million for all amounts of benefit points above 50, and 1 benefit point at -3.7

million for all amounts of benefit points below 50. However, this can be awkward, especially if using the Fibonacci scale, since the distances between negative values will always be smaller than the distances between positive values.

A better alternative is to use explicit positive and negative benefit points. Consider the benefit points from the Fibonacci sequence in Fig. 3.12. The effect points for the new epic *E9* above could be estimated at -8 on *Obj1* (contrasting the positive impact of epic *E4* on the same objective), zero for *Obj2* (no impact on 'number of wrong case decisions'), and -5 for *Obj3* (contrasting the positive impact for *E2* on *Obj3*).

3.19* Other Approaches

Agile at scale frameworks, such as Large Scale Scrum (LeSS) and Scaled Agile Framework (SAFe) present alternative models for prioritizing product elements, or product backlog items, as they are referred to in these frameworks. In LeSS, one is prompted to, 'with *relative value points* (RVPs) as a lightweight proxy for "value", use *planning poker* to experiment with *relative value points* (RVPs) and their estimation' [34, p. 139]. This method is not described in detail. Instead, it is argued that value is not a simple attribute or number, and one is advised to move beyond the simplistic notion of value towards multiple weighted factors, such as stakeholder preferences, strategic alignment, relative points for value and effort, and risk.

In SAFe, the prioritization of product backlog items is based on several parameters. Building on the concept of the *cost of delay* [39], one should use an algorithm to compute the sequence to implement the product backlog items [35]. This approach is called the *weighted shortest job first* (WSJF) method:

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

where the indicated parameters are estimated with relative sizes using the Fibonacci sequence. The complexity of these measures contrasts with what we are advocating. Combining benefit, cost, risk, and duration parameters in one go is not easy and mixing different parameters can inhibit measuring, reporting, and project learning.

We designed the current framework to be intuitive and straightforward to maintain, and the key to this is the clear separation between the cost and benefit parameters. Our approach is minted towards supporting stakeholders' conscious processes.

3.20* Satisficing, Fast, and Frugal

Human-based benefit and cost estimation are judgement-based tasks. Such tasks are often inherently difficult and *inconsistent* (with different people developing different successful strategies) [1, 3, 2] and *ill structured* (where it is hard to even define successful strategies) [26, 43, 40, 48]. Research shows that practitioners of inconsistent and ill-structured tasks can apparently spend their careers *not* learning and *not* improving their performance beyond a very narrow subset of consistent tasks [28, 42].

In the field of judgement and decision making, cognition is often modelled as two sets of subprocesses: the *analytical* and the *intuitive*. The former is deliberate and strives to take into account all relevant cues. It is therefore slow. The latter relies on only a few cues, might not be fully conscious, and is regarded as rapid.

There are reasons to favour the analytical process; after all, rational thinking, taking into consideration all relevant factors with a tight focus on explicit deliberation [41], adds comprehensiveness [9] and is something most of us are trained to value (e.g. the so-called *worship of reason* [16]). Several studies show how humans ostensibly fail to make correct judgements when they do not follow analytical processes, due to biases and undue heuristics [29, 47].

However, human working memory and other cognitive functions limit humans' ability to process all relevant factors, let alone rapidly, when the number of factors becomes large and the relations complex [37, 13]. A large body of research has investigated how to take advantage of the quicker intuitive processes, including the fast and frugal heuristics approach to judgement and decision making [13, 19, 20, 12] and naturalistic decision making [32, 36, 18]. All of these approaches acknowledge the almost impossible task of supplying the sufficiently reliable information required to predict accurately how to proceed in complex situations. Both human decision makers and tools fail to deliver good results under uncertain circumstances when attempting to gather and analyse all relevant data correctly. Instead, it is argued, human cognitive judgement is geared towards processing unreliable partial information rapidly and with sufficient accuracy for the purpose at hand, in line with *satisficing*, rather than optimizing [44]. Following this argumentation, we determine that methods and tools should be designed to support this mode of decision making, rather than geared towards analysing the totality of a situation [13].

The underlying principles in our methods are in line with these ideas. We design our methods so that stakeholders

- Consider a limited number of cues at a time and a single relation at a time,
- Provide a modest number of assessments (the white portions of the tables in this book), from which additional measures are automatically calculated in a transparent manner (the shaded portions of the tables), and
- Perform relative points-based estimations by comparing product elements, rather than producing absolute monetary estimates on individual product elements.

Moreover, the methods are designed to facilitate project and community learning. In Hogarth's [22, 21] terms, intuition is expertise that is internalized [6], perhaps after extended experience and deliberate practice [7]. Intuition can therefore be trained. Klein [32] suggests aiming to learn like an expert, that is, provide meth-

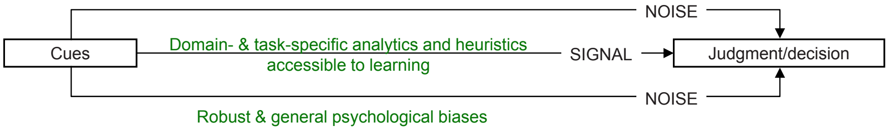

Fig. 3.17 Processes amenable to learning (signal) and processes robust to learning (noise).

ods and tools that support *learning to become* an expert, in addition to *acting* as an expert. Accordingly [32, 33], we design our methods so that stakeholders

- Engage in *deliberate practice* by assessing and reassessing product elements that are associated with goals (returns) and evaluation criteria (objectives),
- Obtain *feedback that is accurate and diagnostic and reasonably timely* by evaluating and re-evaluating assigned benefit points and size points and their monetary values, following stakeholders' evaluation of MVPs in increments, and
- Enrich *experiences by reviewing prior experiences to derive new insights and lessons from mistakes* by using points-based estimates to monitor project progress.

The facility to access relevant domain knowledge systematically is central to learning. For inconsistent tasks, it is important to stimulate processes that are sensitive to domain knowledge [38, 32] and learning [7]. These processes can be seen to increase the desired signal against the noise of competing processes that are driven by general psychological traits that are not domain specific and not amenable to learning [30, 11, 19, 20, 23, 12] (see Fig. 3.17). In the spirit of Stewart [45], one must increase the reliability of estimates, in the sense of decreasing undue inter- and intra-rater variance. We do want variance in a group of diverse stakeholders due to their respective domain- and task-relevant perspectives, but we do not want variance due to the misperceptions or inaccessibility of the knowledge in question and the host of undue biases. We advocate methods such as relative and pairwise comparisons that help stakeholders tap into domain knowledge and structure its use in assessments. Pairwise comparisons are a core element in the conscious cognitive processes of judgement[38]. To strengthen conscious comparisons further, methods that focus on differences are beneficial (e.g. the repertory grid technique [10]). We also advocate structured group methods that are intended to elicit and illuminate domain knowledge from various perspectives.

References

1. P.L. Ackerman, "Individual differences in skill learning: An integration of psychometric and information processing perspectives," *Psychological Bulletin*, vol. 102, pp. 3–27, 1987.
2. J.P. Campbell, "Modeling the performance prediction problem in industrial and organizational psychology," in *Handbook of Industrial and Organizational Psychology*, M.D. Dunnette and L.M. Hough, Eds. Consulting Psychologists Press, Inc., 1990, vol. 1, pp. 687–732.
3. J.P. Campbell, R.A. McCloy, S.H. Oppler, and C.E. Sager, "A theory of performance," in *Personnel Selection in Organizations*, N. Schmitt and W.C. Borman, Eds. Josey-Bass, 1993, pp. 35–70.
4. M. Cohn, *Agile Estimating and Planning*. Prentice Hall, 2005.
5. M. Denne and J. Cleland-Huang, *Software by Numbers: Low-Risk, High-Return Development*. Prentice Hall, 2003.

6. H.L. Dreyfus and S.E. Dreyfus, *Mind over Machine*. The Free Press, 1988.
7. K.A. Ericsson, "An introduction to Cambridge Handbook of Expertise and Expert Performance: Its development, organization, and content," in *The Cambridge Handbook of Expertise and Expert Performance*, K.A. Ericsson, N. Charness, P.J. Feltovich, and R.R. Hoffman, Eds. Cambridge Univ. Press, 2006, ch. 1, pp. 3–20.
8. G.T. Fechner, *Elemente der Psychophysik – Erster Teil*. Breitkopf und Härtel, 1860.
9. D.P. Forbes, "Reconsidering the strategic implications of decision comprehensiveness," *Academy of Management Review*, vol. 32, no. 2, pp. 361–376, 2007.
10. F. Fransella, R. Bell, and D. Bannister, *A Manual for Repertory Grid Technique*. John Wiley & Sons, Ltd., 2004.
11. G. Gigerenzer, *Gut Feelings. The Intelligence of the Unconscious*. Viking, Penguin, Ltd., 2007.
12. G. Gigerenzer and D.G. Goldstein, "Reasoning the fast and frugal way: Models of bounded rationality," *Psychological Review*, vol. 103, no. 4, pp. 650–669, 1996.
13. G. Gigerenzer and P.M. Todd, Eds., *Simple Heuristics that Make Us Smart*. Oxford University Press, 1999.
14. T. Gilb, *Competitive Engineering—A Handbook For Systems Engineering, Requirements Engineering, and Software Engineering Using Planguage*. Butterworth-Heinemann, 2005.
15. J.W. Grenning, "Planning poker or how to avoid analysis paralysis while release planning," 2002.
16. J. Haidt, "The emotional dog and its rational tail: A social intuitionist approach to moral judgment," *Psychological Review*, vol. 108, pp. 814–834, 2001.
17. T. Halkjelsvik and M. Jørgensen, "From origami to software development: A review of studies on judgment-based predictions of performance time," *accepted to Psychological Bulletin*, 2012.
18. K.R. Hammond, "Upon reflection," *Thinking and Reasoning*, vol. 2, no. 2/3, pp. 239–248, 1996.
19. R. Hertwig, U. Hoffrage, and L. Martignon, "Quick estimation: Letting the environment do the work," in *Simple Heuristics that Make Us Smart*, G. Gigerenzer and P.M. Todd, Eds. Oxford University Press, 1999, ch. 10, pp. 75–95.
20. U. Hoffrage, R. Hertwig, and G. Gigerenzer, "Hindsight bias: a by-product of knowledge updating?" *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol. 26, pp. 566–581, 2000.
21. R. Hogarth, "Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics," *Psychological Bulletin*, vol. 90, no. 2, pp. 197–217, 1981.
22. R.M. Hogarth, *Educating Intuition*. The University of Chicago Press, 2001.
23. R.M. Hogarth, "When simple is hard to accept," in *Ecological Rationality: Intelligence in the World*, P.M. Todd and G. Gigerenzer, Eds. Oxford University Press, 2006.
24. H.L. Hollingworth, "A survey of the characteristics of projects with success in delivering client benefits," *The Journal of Philosophy, Psychology and Scientific Methods*, vol. 7, no. 17, pp. 461–469, 1910.
25. International Institute of Business Analysis, "A guide to the business analysis body of knowledge," 2009.
26. E.J. Johnson, "Expertise and decision under uncertainty: Performance and process," in *The Nature of Expertise*, M.T.H. Chi, R. Glaser, and M.J. Farr, Eds. Lawrence Erlbaum Associates, Inc., 1988, pp. 209–228.
27. C. Jones, *Estimating Software Costs: Bringing Realism to Estimating*, 2nd ed. McGraw-Hill, 2007.
28. M. Jørgensen, "Forecasting of software development work effort: Evidence on expert judgement and formal models," *Int'l J. Forecasting*, vol. 23, pp. 449–462, 2007.
29. D. Kahneman and S. Frederick, "A model of heuristic judgment," in *The Cambridge Handbook of Thinking and Reasoning*, K.J. Holyoak and R.G. Morrison, Eds. Cambridge Univ. Press, 2004, pp. 267–294.
30. D. Kahneman and G. Klein, "Conditions for intuitive expertise—a failure to disagree," *American Psychologist*, vol. 64, no. 6, pp. 515–526, 2009.

31. C.T. Keil and J.M. Cortina, "Degradation of validity over time: A test and extension of ackerman's model," *Psychological Bulletin*, vol. 127, pp. 673–697, 1987.
32. G. Klein, "Developing expertise in decision making," *Thinking & Reasoning*, vol. 3, no. 4, pp. 337–352, 1997.
33. G. Klein, *Sources of Power: How People Make Decisions*. MIT Press, 1998.
34. C. Larman and B. Vodde, *Practices for Scaling Lean & Agile Development: Large, Multisite, and Offshore Product Development with Large-Scale Scrum*. Addison Wesley, 2010.
35. D. Leffingwell, *Agile Software Requirements: Lean Requirements Practices for Teams, Programs and the Enterprise*. Addison Wesley, 2011.
36. R. Lipshitz, G. Klein, J. Orasanu, and E. Salas, "Rejoinder: A welcome dialogue – and the need to continue," *J. Behavioral Decision Making*, pp. 385–389, 2001.
37. G.A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychological Review*, vol. 63, pp. 81–97, 1956.
38. T. Mussweiler, "Comparison processes in social judgment: Mechanisms and consequences," *Psych. Review*, vol. 110, no. 3, pp. 472–489, 2003.
39. D. Reinertsen, *Principles of Product Development Flow: Second Generation Lean Product Development*. Celeritas Publishing, 2009.
40. W. Reitman, *Cognition and Thought*. Wiley, 1965.
41. E. Salas, M.A. Rosen, and D. DiazGranados, "Expertise-based intuition and decision making in organizations," *Journal of Management*, vol. 36, no. 4, pp. 941–973, 2010.
42. J. Shanteau, "Competence in experts: The role of task characteristics," *Organizational Behavior and Human Decision Processes*, vol. 53, pp. 252–266, 1992.
43. H.A. Simon, "The structure of ill-structured problems," *Artificial Intelligence*, vol. 4, pp. 181–201, 1973.
44. H.A. Simon, *The Sciences of the Artificial*, 3rd ed. MIT Press, 1996.
45. T.R. Stewart, "Improving reliability of judgmental forecasts," in *Principles of Forecasting: A Handbook for Researchers and Practitioners*, J.S. Armstrong, Ed. Kluwer Academic Publishers, 2001, pp. 81–106.
46. R. Tamrakar and M. Jørgenseni, "Does the use of Fibonacci numbers in planning poker affect effort estimates," in *Proc. 16th International Conference on Evaluation & Assessment in Software Engineering (EASE 2012)*, 2012, pp. 228–232.
47. A. Tversky and D. Kahneman, "Judgement under uncertainty: Heuristics and biases," *Science*, vol. 185, no. 27, pp. 1124–1131, Sept. 1974.
48. J.F. Voss and T.A. Post, "On the solving of ill-structured problems," in *The Nature of Expertise*, M.T.H. Chi, R. Glaser, and M.J. Farr, Eds. Lawrence Erlbaum Associates, Inc., 1988, pp. 261–285.