Data complexity does not improve knowledge

BY DONALD KENNEDY AND SIMON P. PHILBIN

EXECUTIVE SUMMARY

When a new engineering project begins, there often are lengthy discussions and disagreements over how to set up the financial cost codes to be used. This situation can be compounded by a high staff turnover on engineering projects. We found that discussions about coding held 20 years ago are mostly unchanged from those today, though the needs of projects and their cost accounting are often separate from firms' long-term asset management financial requirements. Here we provide insights for the structuring of cost codes to better manage and understand projects so that engineering risks are properly mitigated and project resources can be deployed more effectively. These lessons can be applied in many other areas of the planning, monitoring and controlling aspects of management.


54.5
45.6
37.4
42.1
74.6
58.0
49.6
51.5
Q3 2013
Q4 2013
Q1 - 2014
Q2 - 2014
Q3 - 2014
Q4 - 2014
Q1 - 2015
Q2 - 2015

Considerable effort is devoted in organizations to the tracking and analysis of cash flows. A primary metric used to evaluate the effectiveness of a management team is the financial performance of the organization.

Following the prison sentences of senior management at Enron, it became easier to obtain the resources from senior management to manipulate accounting data in the manner desired by departments in control of financial reporting. The level of manipulation can become highly sophisticated, including processes such as producing algorithms to allocate specific fixed costs to the various activities directly involved in the value-adding performance of the business.

Before 2000, the accounting and cost systems available in most organizations were much simpler than today's. Many legacy accounting platforms had been adapted from the first computers from decades before and it was considered by most organizations to be too costly to overhaul their data systems. These costs typically ran up to $300 million for multinational corporations.

Concerns over the "Y2K bug" provided the narrative to convince executives to take the leap to convert to integrated enterprise resource planning (ERP) systems to control financial operations. At the time, it was realized that the high short-term costs of these upgrades to more sophisticated ERP systems would not have an immediate payback considering the loss in productivity as workers learned the new system. But it was speculated that the learning curve and adaption to the capabilities of the ERP systems would provide long-term net benefits. In our experience, these analyses did not include the ongoing upgrades and major efforts required by the organizations as both hardware and software evolved.

Given this history, around 2000 organizations discovered they had powerful tools available to integrate data collection for financial tracking, quality control and resource planning. In our experience, ERP systems have been the center of attention as they struggle to develop internal systems that utilize the power at their fingertips. Indeed, industrial companies in 2021 appeared to be facing the same questions organizations faced in 2000, especially in how to best set up their systems for cost and quality control.

These questions that were worked through in the 1980s and 1990s are more difficult now to answer because of two developments. One is the ability to handle data broken into very small details given the implementation of ERP systems. In earlier times, data that could be feasibly handled manually prevented developing too complex processes.

The second difference is much higher employee turnover compared to previous decades. Knowledge workers can be expected to stay at one job for about as little as three years. The high turnover of knowledge workers has created a common situation where an experienced manager comes into a new organization with a powerful ERP system yet with their peers disagreeing how to best utilize the tools to improve their ability to control and understand the organization's performance. Sometimes these tools have been acquired recently to replace abandoned systems deemed unsatisfactory. Other times, the processes previously used are blamed for the shortcomings and the focus is on revamping existing tools to more practical processes.

The facets of data granularity: Engineering cases

A number of consequences are derived from decisions on the amount of coding used in the cost accounting of engineering projects. In Figure 1, we present six charts to demonstrate a general impact for having various levels of detail in the coding structures used. We have referred to the fineness of detail desired on each chart as the granularity.

ERP systems have been the center of attention as they struggle to develop internal systems that utilize the power at their fingertips.

Engineering projects with few codes would have low granularity; projects with multiple levels of codes and subcodes would have high granularity. The magnitude of the impacts are illustrative but resulted from discussions with three senior project managers with in-depth involvement in the execution of engineering projects in the multimillion- to billion-dollar range.

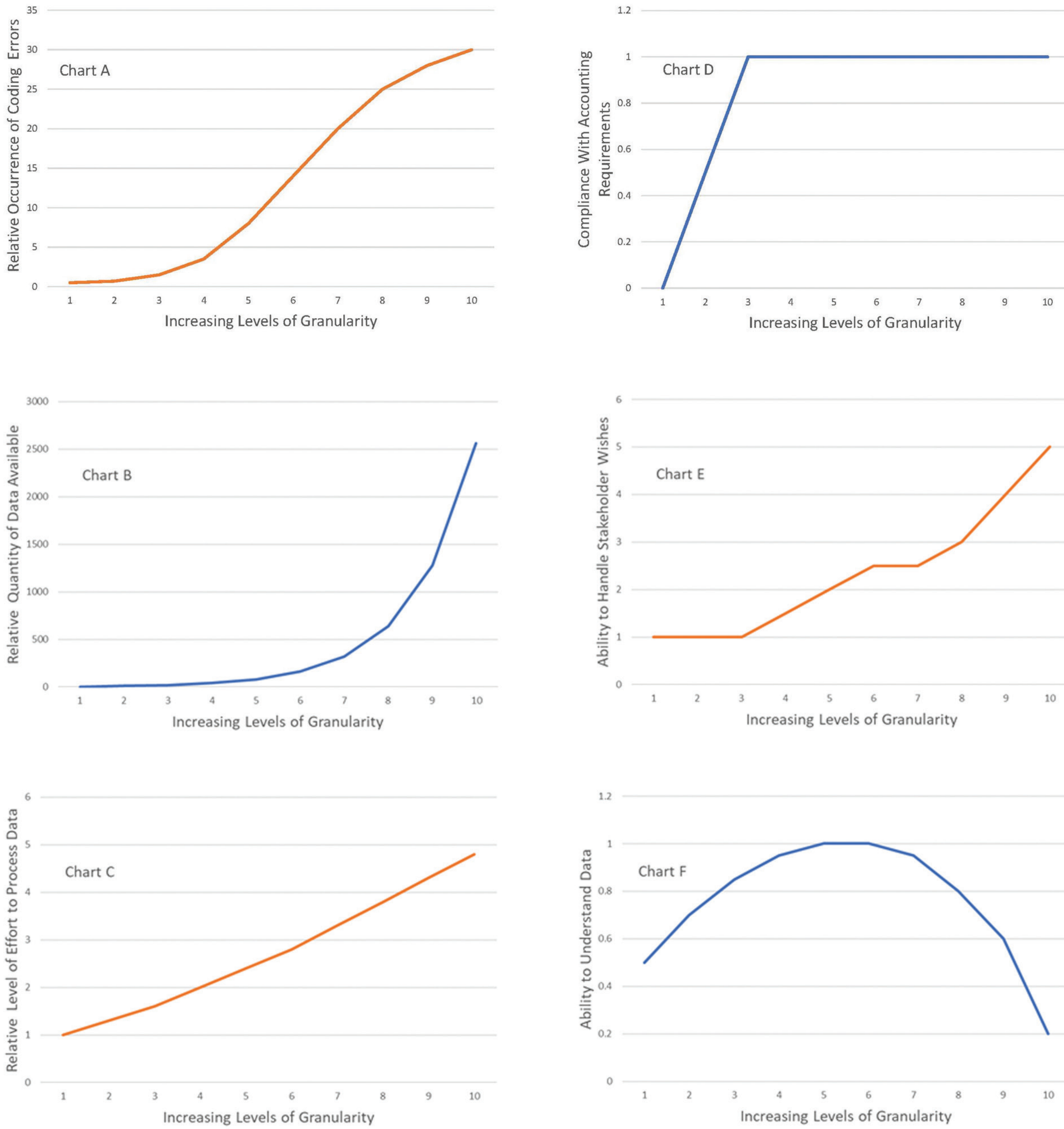
Chart A demonstrates how the increased number of codes increases the number of errors encountered in the project. We submit that a major negative to increased coding is the loss of validity in the amounts assigned to each code. We offer four general causes for error. The first is the familiar typographical error. The second is due to intentionally assigning a code the user understands is probably incorrect but does not believe there is value in taking time to figure out the correct assignment. The third cause is having codes with descriptions that are subject to interpretation by the user. Two users may not agree on the "correct" code for the given item; the final cause is creating different processes for different users.

To support the shape of the curve in Chart A and to demonstrate evidence for these causes, Figure 2 (Page 16) represents the actual costs, as recorded per facility on a $55 million engineering project. The scope required building almost identical process facilities at 11 different sites. There were some differences due to site specific variations, such as soil conditions, available electrical infrastructure and road access. The expectation was that all 11 sites would cost the budgeted $5 million. The costs for each site were tracked separately with a code that started with the first letter of the name of the location. Figure 2 shows the costs with the locations sorted alphabetically.

The general trend seen is that sites

GRANULARITY FACTORS

Figure 1. The general impact for having various levels of detail in the coding structures used on engineering projects.



lower in the alphabetical order are lower in cost. By interviewing team members, we discovered that when there was uncertainty on how to code a particular item, the easiest way to process the cost was to pick the first item in the drop-down code that would accept that item. If 11 pumps were required, it did not make sense to the buyer to charge three hours of time for processing the purchase order split 11 ways. It seemed more practical to charge the three hours to one site only, with the cost often assigned to the first one to appear in the alphabetically sorted drop-down menu.

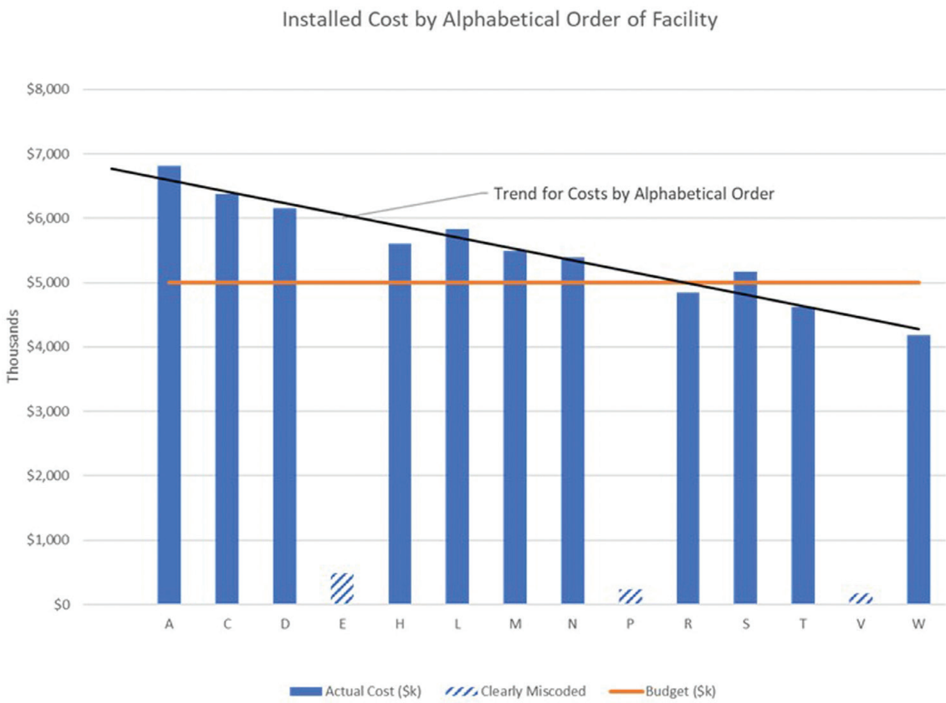
It is seen that locations E, P and V had costs allocated to them. These costs were clearly identifiable as errors in coding as these sites were not involved in this project. Although these incorrectly coded costs only represent 1.4% of the total final cost, one cannot assume these are the only mistakes. It is reasonable to assume that miscoding impacted all sites about equally and the lack of any legitimate costs for those sites only highlight these errors. One can assume that perhaps 10% of all the costs are coded incorrectly due to typographical errors.

Perhaps 10% of all the costs are coded incorrectly due to typographical errors.

Chart B in Figure 1 demonstrates the amount of data generated by the accounting systems as a function of the coding granularity. This should be almost self explanatory, as the number of coding options increases the number of accounting entries. As an example, for one multimillion-dollar engineering project, the company developed three independent coding

RECORDING COSTS

Figure 2. Costs per site for a multisite construction project.



structures. One accommodated the permanent fixed asset accounting rules for the organization. The second suited the desires of the project management team. The third split the costs according to the various functional departments' operating budget tracking systems. These were not relatable between departments. The number of possible codes therefore became a product of the number of codes in each structure.

On one $40 million project recently completed, there were 400,000 accounting entries spread over 640 cost codes. The project manager struggled to determine answers to questions such as "How much was paid to a certain subcontractor?" The costs were spread over 15 codes but these codes were also used for other items so it was very difficult to use the system to determine such answers.

Chart C represents the administrative effort to manage the cost system as a function of the granularity. The effort is not proportional to the amount of data generated. This is due to the acceptance of the existence of errors and lack of understanding needed to properly manage the more complex systems.

Here are two examples to illustrate this point: On one engineering project with a budget of $350 million, the cost engineer was allowed the freedom to set up the project codes. One person was able to administer the cost coding at the project level. On another project with a budget of $80 million, corporate managers dictated processes that established a much more granular coding structure. On that project, three people worked in the same function doing the same general work as the single person on the much larger but simpler coded project.

It is very difficult to track all the time spent by all stakeholders (accounts payable, accounts receivable, project management, executives and auditors) on the administration and analysis of the coding, but we are assuming the levels of effort would be similar to that shown in Chart C.

Chart D represents our findings on the actual granularity typically required for accounting departments to comply with their needs. This is different from meeting the desires of the people working in the accounting departments. As an example, on one engineering project the accounting department presented its coding structure the clerk stated was required for the project. The company was government-regulated and there were legislative requirements for compliance. One code was for tracking the cost of fuel filters on diesel drive engines and another for a category named "crossover" piping. The accounting clerk said the exclusion of the codes was not an option due to government regulations.

On one $40 million project recently completed, there were 400,000 accounting entries spread over 640 cost codes.

We traced back the origins of the coding structure and the specifics of the government regulations. For the latter, the regulations stated there were seven different categories of capital costs the company had to report. All other codes were at the company's discretion.

The company had developed its accounting structure in the 1950s when diesel engines were the main source of power for their process equipment. At the time of facility setup (around 1951), there was a major project to adjust some of the piping in the plants, an initiative called "crossover piping." This category accumulated significant costs that year, and, due to typographical errors and other mistakes, had a few costs assigned over the next decades that were basically meaningless data. Indeed, the Canadian Revenue Agency lists 22 different capital cost allowance categories but many of these would not be required to track on most projects, such as codes for fishing vessels and goodwill.

For these reasons, Chart D demonstrates that too few codes would not meet the accounting requirements for the organization but the amount of granularity needed is low to fully meet the base criteria.

Chart E represents the opposite of Chart D, namely expanding the number of cost codes to meet stakeholders' wishes. We touched on the details typically requested by the accounting personnel. For an example on the project management side, a director of engineering projects wanted sufficient coding to track costs associated with each of offloading, handling, fit-up, welding and hydrotesting of two different sizes of piping. The goal was to provide the estimating department sufficient cost details to better hone its estimates. The project manager wanted a single code to handle all piping-related labor costs. The director was able to override the project manager and 10 additional cost codes were added. At the end of the project, the costs assigned to each of the codes did not make logical sense.

In speaking with the tradespeople who did the work, this was apparent by determining the effort for each of the steps. In one case, a large load of materials arrived at the site including one spool of piping of each of the two sizes in question. A welder removed the items and installed them in a day.

When it was time to decide what part of the day was spent on any task, it would be a matter of guessing. Should the welder split the time carrying the two spools between the four codes for unloading and handling? And should such a split be equal for the two sizes or based on weight or length or some other dimension? There were no costs assigned to the hydrotesting because the spools were tested as part of the main test of the systems they tied into.

The shape of Chart E is based on how there will be a certain granularity to satisfy many stakeholders but it would be a challenge to have sufficient granularity to meet everyone's wishes.

Chart F represents the usefulness of the data collected as a function of the granularity of the codes. This is not intended to consider the level of effort required to extract knowledge from the data, or determine a point of diminishing returns for the effort expended. The essence of the graph is that the extra effort required to produce data to fit a large number of codes does not create better information than a more manageable system.

In the case presented in the discussion of Chart E, there was a desire to understand the costs associated with the installation of two sizes of piping. With a single code for all pipe installation tasks, as desired by the project manager, there is no way to differentiate between activities or pipe sizes. With the 10 codes as requested by the director, the data is there to interpret, whether the results are useful or not.

To find the total cost to install the piping requires knowing which codes to sum. Let's say the load arriving with the two spools from the case in Chart E was postponed by eight hours. The welder who performed all the tasks had to charge time to some code, so the time is assigned to handling the larger spool. Such incidences will result in certain codes having high costs compared to others that may not reflect the true effort for the task measured. As more codes increase the number of errors, the values from the data are more suspect than a simpler system.

Resource allocation on engineering projects and system granularity

Many engineering consulting companies require workers to only do work that is billable to a client. The clients' expectation is that they are billed for value-added work, and therefore hours charged to the engineering projects would be for work by the engineer that is related to them. As late as the mid-1990s, a review of the offices of engineers in our area revealed it was common for engineers to have computers in their offices for regular use. A common question asked then was whether computers increased the productivity of these knowledge workers. Would the client receive value for the money charged to them for the engineers' time that included using the computer?

In our experience, detailed records for time invariably poorly match the actual work performed. In discussions with dozens of individuals over several decades, and verifying the recorded outcomes with the three senior project managers mentioned above, certain common behaviors create a misalignment between the work performed and how this time is recorded.

As more codes increase the number of errors, the values from the data are more suspect than a simpler system.

We offer three types of sources of error. Firstly, it is common for organizations to be pressured into performing tasks relating to work that has a high certainty of being formally approved but is not yet officially in the ERP system. Considerable savings can be achieved by setting up and planning early, as well as effectively tuning the system parameters to the needs of the engineering projects. In such cases, the work is carried out but the ERP system does not allow recording the time where it belongs. Yet the workers will typically charge their time to approved projects that may be totally unrelated to their work.

In these cases, and for the observed outcomes of time tracking, there is a desire to produce a finely granular system to track resource use and better understand the nature of the business. But in reality, the data collected is mostly meaningless in terms of learning enough to improve. Therefore, based on our insights across the engineering projects and cases we have described, we recommend the granularity of financial accounting systems needs to be reconsidered.

Furthermore, we suggest that the level of granularity should be contingent on the circumstances of the project, such as the pace, level of technological complexity and interconnectedness across the project, resource base and stakeholders. This way, project financial codes can be implemented more effectively and meet the needs of all project stakeholders while ensuring effective decision-making is maintained and engineering projects realize successful outcomes. ❖