DECTECH'S BRIEFS
Covering What Matters

Finding Oil, People and Other Treasures:
The Anatomy of Distributed Business Decisions

Dectech has been working on many superficially unrelated problems - lending, recruiting, exploration, site research - and noted what is more than a family resemblance. They are all filtering processes with multiple decision stages distributed across organisations.

Making the right decisions is hard enough. It is even harder when lots of different people are involved at different stages. This brief describes the key characteristics of such processes and what actions managers can take to improve them.

In the case of graduate recruitment, interview judgements are inaccurate and misdirected and the whole process can be reconfigured to improve its effectiveness. Consequently, current practices are close to five times less predictive than they could be.

People know what they like (one person in our office is obsessed with sausage and marmalade sandwiches). They can also tell you what they thought they'd like but didn't (for example, Beowulf[a]). Then there are the things they avoided but actually would have enjoyed (err…we don't know which).

This brief concerns such judgement errors in business. For example, "how did we recruit that disastrous employee?" or "why did we turn down a retail outlet there?". In some shape or form, the performance of every business depends on filtering out a few good deals from many bad ones - the wheat from the chaff. Hence, these decision processes are pervasive and improving them is an opportunity in every industry.

Multistage Decisions

Filtering usually takes place across several stages located in different parts of the business. Consider the two main stages of oil exploration. At the start, geologists survey new acreage looking for promising underground rock formations. Having estimated the potential reserves in a trap, they then make a decision on drilling. Next, the engineers drill an exploratory well. They appraise the trap's commercial value and then make a decision on development.

Each decision risks two potential errors. Type I are the false negatives (i.e. not drilling a good trap) and Type II are the false positives (i.e. drilling a bad trap). Both errors are caused by mis-estimation. If the geologist overestimates the reserves, a small trap gets through (Type II) and vice versa (Type I). Decision-making effectiveness is therefore driven by three key process characteristics; firstly, the predictive power of the available diagnostics at each stage; secondly, the sequencing and interaction of these diagnostics; thirdly, the tuning of the process to the financial implications of errors. Let's look at these in turn.

Graduate Recruitment: A Cautionary Tale

It is clear that you will make better decisions if you make better predictions. Therefore it would seem obvious that you should track and improve your predictiveness for any commercially significant decision, particularly repetitive ones. Yet people don't always do this. To illustrate this point, we report some of our recent findings on graduate recruitment.

There are two main problems with recruiting. The first is illustrated by Figure 1. Each dot is a candidate and the co-ordinates show how they were rated by two groups containing about five interviewers each[b]. As can be seen, there is little agreement. For example, some of Set A's best candidates are Set B's worst and vice versa. So these interviewers can't have much predictive power because they hardly even agree with each other[c].

Figure 1. Interviews Are Close to Contentless
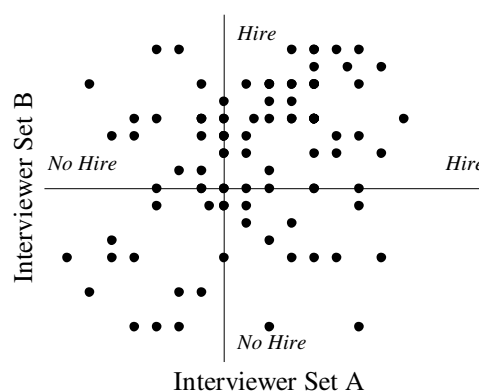
Then there is the second problem - even if interviewers agree on the "good" candidates, are they correct? To examine this question, we evaluated several hundred recently recruited graduates to compare their skills with how they subsequently performed at their jobs. Table 1 summarises the findings. The second column shows the relative importance of different traits for diagnosing good hires. Basically, it was the ambitious and diligent ones that flourished (captured by personality Traits III and IV), a conclusion supported by other research conducted in this field[d].

Unfortunately, to the extent that interviews filter for anything, they do not filter for these criteria. The first column shows how the diagnostic value of interviews tends to favour the more intelligent candidates and actively removes the diligent ones[e]. Likewise, this firm's psychometric testing was also misaligned. Meanwhile, their analysis of CVs was entirely qualitative and the CV information was effectively discarded once a candidate reached interview.

Table 1. What Makes People Successful...

	…In Interview	…In Work
Trait I	10.0	
Trait II	7.7	
Trait III	-6.8	9.7
Trait IV		10.0

Anyway, in this example the interviewing process has failed twice. Firstly, it has little or no content. Secondly, what content there is does not help select better candidates. In fact, it marginally works against them. We discuss how firms get into this situation and how they can get out of it presently. In the meantime, we simply note that graduate recruitment is an important decision-making process that isn't sufficiently researched.

Geological Faults

The second key characteristic of distributed decision processes is how the diagnostics are glued together. Again oil exploration helps to illustrate the issues. Clearly geologists and engineers should try to estimate reserves accurately and without bias[f]. But even if they make the best individual estimates possible, Figure 2 shows how their interaction has a further, and potentially equally important, effect.

Figure 2. Combining Two Estimates
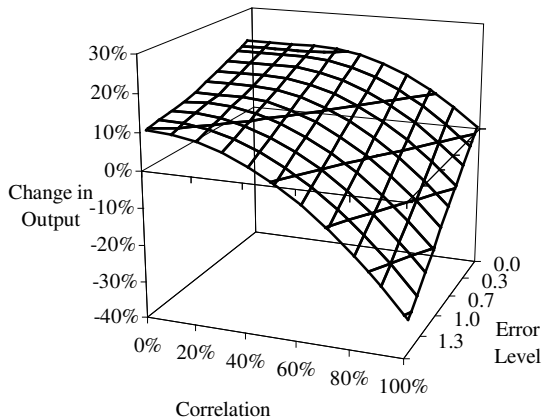
Specifically, the surface shows how the reserves of the average well selected by exploration changes with the error size and interaction. In the base case errors are 1.0x normal levels and 65% correlated (i.e. if the geologist underestimates the reserve, the engineer is more likely to make the same mistake). The figure shows how output improves when errors are either reduced or more independent. So even if the geologists and engineers continue to make similar sized mistakes, the average well reserve can be increased by 12% if those mistakes are no longer correlated. That's equivalent to an 80% error reduction.

Unfortunately some correlation is inevitable because the engineer uses the geologist's data. However, the engineer's estimate will also contain further, unnecessary contamination - for example, from being aware of, and therefore influenced by, the geologist's conclusions. So there is scope to improve such processes simply by ensuring that the different parties form their judgements independently[g].

Naturally, this is not the only process design issue. For example, perhaps the engineer stage should be expanded into an adversarial one, where two engineers make the case for and against. Likewise, there are questions on how to incentivise each decision stage in order to reduce errors and optimise interactions. For example, different bonus structures for geologists and engineers will lead to tensions that can be either good (i.e. productive cross-checking) or bad (i.e. unproductive cross purposes).

How High The Hurdles?

Improving prediction is all very well. But how can complete decision-making processes be made economically optimal? We return to recruiting to describe some of the issues.

Figure 3 shows the progress of two students through a three-stage recruiting process - CV screening followed by first and second round interviews. Initially, the red candidate is a likely reject, since he or she falls below the line. But as this view only uses CV data there is uncertainty, as shown by the error bars. Then, as more information is collected at interviews and added to the CV perspective, red's rating improves and the estimation errors decrease. By the final stage, red is a clear hire. Meanwhile, blue is a probable hire, but still carries some risk. We now discuss four issues associated with the economic performance of this process.

Figure 3. Evolving View on Two Candidates
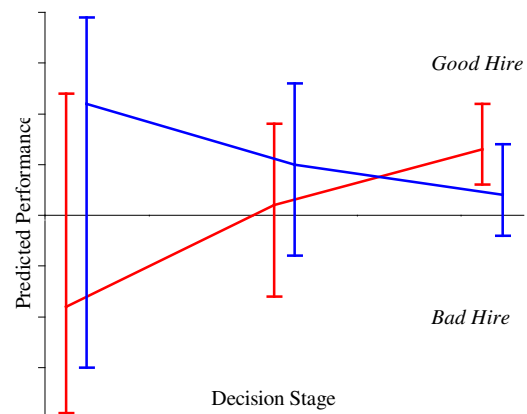
Firstly, the diagnostics have been placed in decreasing order of information gain-to-cost. CV screening is first because it reveals most about the candidate for least expense. Likewise there are fewer interviews at first round, sometimes augmented with psychometric testing. Overall the aim is to obtain as much information as possible as inexpensively as possible as early as possible. On that basis, the net can be thrown wider, on a lower budget, involve less interviewing time, and still generate better hires.

Secondly, there is the question of who gets through each stage. Because false positives only cost a first round interview and false negatives mean forgoing a good candidate, it will probably be better to have a relatively low CV threshold and admit borderline rejects, like red, to the next round. However, the thresholds would be different if the error costs were different. For example, the situation would be reversed if first stage false positives involved wasting $20mm drilling an exploratory well. In this case even blue might be rejected as too risky.

This then introduces the third issue - risk. Acceptance thresholds clearly need to be set based on accuracy and error costs. However, they also depend on risk appetite. For example, the organisation may want to pursue a portfolio of low probability but high yield prospects or it might prefer smaller, safer bets with lower expected value.

The fourth point is that in reality many decisions, like recruiting, are not "take it or leave it". This means thresholds are more flexible. For example, the attractiveness of a candidate at the final stage of Figure 3 might then be used to either adjust the offered salary or scale the candidate conversion effort.

Conclusion

There are many reasons why companies continue to operate poor decision processes. We discuss a few and then review the action steps outlined in this brief.

One major cause of broken decision processes is that no one knows they are broken. If we set the thresholds high enough we can curb the false positives and sleep soundly, entirely ignorant of all the rejected wunderkinds and undiscovered oil. Similarly, it is unsurprising that interviews have low content. Typically they are given by successful managers who built careers on their people skills. Yet decades of research tells us that humans, even experts, need rapid and plentiful feedback to make good judgements - conditions that are absent in recruiting[h].

Another reason why processes go unfixed is that there are vested interests in the status quo. By trying to apply scientific scrutiny to what was previously a "dark art" the business may gain a cheaper and more effective process, but the current operators may fear loss of power and livelihood. Hence, lending officers initially insisted that credit scorecards shouldn't be applied in corporate banking and interviewers, save an enlightened few, tell us that you can't apply statistics to recruitment.

Finally, it is hard to do. People say that to get a good salesman you have to hire ten and fire nine. But this amounts to surrender. Likewise, we are often told that such analysis would be interesting, but some part of the business has previously blocked reforms and will do so again. We suggest that roadblocks persists because no one bothered to do the quantitative work needed to persuade, or even force, change. Sure, this might involve some ambitious analysis, but that's what well-run companies do. Try telling the shareholders it's too hard.

So, in summary, what can be done? Firstly, invest in state-of-the-art diagnostics. You may think that you already have them, but even assuming you do there will always be scope for further improvements. You may think that the issues are unquantifiable, but this ignores scientific progress and abandons the process to something far worse--shamanism. Secondly, you need to understand how these different diagnostics fit together. What order should they go in? How independent are they? Should there be audits and devil's advocates? Thirdly, the process needs to be tuned to optimise value. This means setting the acceptance thresholds at each stage according to the financial implications of mistakes and the organisation's risk appetite. It also means having incentive structures that reward the desired outcomes.

So, if you're reading this, then you clearly already have an excellent filtering system for choosing which unsolicited articles to browse. Now, you just have to attend to all the other decision processes in your business. But first, we'd like to recommend the sausage and marmalade sandwich.

References and Footnotes

a. "If I had to live my life all over again, I'd do it all exactly the same - only I wouldn't read Beowulf." attributed to Woody Allen.

b. These interviewers included very senior members of the business, were highly motivated, and had many years of interviewing experience.

c. The R^2 is 10%, so each rating is 90% noise. Also note that these are aggregated ratings. Individual interviews are even noisier. Indeed, interviews are less predictive than a simple CV scorecard. For further interview bashing and corroboration that recruiting performance can typically be improved by a factor of five see Robertson, I. T., & Smith, M. (2001) Personnel selection. *Journal of Occupational and Organizational Psychology 74* pp: 441-472.

d. Bono, J. E., & Judge, T. A. (2004) Personality and transformational and transactional leadership: A meta-analysis. *Journal of Applied Psychology 89(5)* pp: 901-910.

e. For this work we developed some cool techniques for getting around selection bias. Essentially, the risk is that clever people don't seem to perform well because only clever people get recruited. The results presented here are corrected for this attenuation. These candidates were all clever enough once they got through the CV screen so that prioritising this in interviews was a wasted effort.

f. So the first section was about better prediction. This means reducing error variance (accuracy) and eliminating error mean (bias, e.g. in their enthusiasm, geologists might consistently overestimate small reserves). This section is about the effects of error correlations.

g. Note that this doesn't mean ignore the geologist's conclusions. The idea is first to get an independent assessment, use it to update the prior prediction and *then* make a decision based on both inputs.

h. This is a big topic. For an opener see Keren, G. B. (1987) Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes, 39* pp: 98-114. More generally, there are many techniques developed by experimental psychologists that can be used to radically improve interview content.