PERFORMANCE EVALUATION

Carolyn Stringer
Department of Accountancy and Business Law, University of Otago, P O Box 56, Dunedin, New Zealand 9054

Abstract

Performance evaluations systems are a critical part of overall performance management systems. This intensive case study provides insights into how the use of subjective performance evaluations in a complex organisational setting has led to perceptions of injustices (e.g., procedural, distributional, interactional), and unintended consequences. The key injustices were mixed practices, unclear criteria, financial focus, little differentiation between good and poor performers, stickiness in ratings, inequities in target setting, higher ratings at higher grades, and the predetermined theory. The consequences include the lack of trust, generous bonuses, gaming, inability to influence, resource allocation, bonuses are expected and not performance-related, and the system is costly to administrate. Multiple sources of evidence support the findings, and the patterns are consistent over a three year period. Future research needs to develop a deeper understanding of how these parts interrelate, where subjective performance evaluations work, and where they do not work, and to encourage breaking down the divide between functional specialisations (e.g., human resources, accounting).

Keywords: performance evaluations, incentives, performance management, forced distribution, balanced scorecard, EVA®.

Introduction*

Performance evaluations are an essential part of performance management in organisations. This paper responds to calls for more research from Ittner, Larcker and Meyer (2003) into the multifaceted and complex nature of subjective performance evaluations. This is a growing issue, especially when performance evaluations are linked to cash bonuses. A related issue is that subjective performance evaluation is an under-researched topic in the performance management and control literatures (Otley, 1999), and there is a pressing need to know more about the way performance evaluation systems operate, and their impact on the broader performance management framework.

This 2½ year intensive case study paper examines the role performance evaluation processes play as part of the overall performance management and control system (PMC) in Sensol, a large and complex Australasian organisation.[1] Sensol's incentive system for their top 700

---

* I wish to express thanks to the academic community whose comments have contributed to the re-write of this paper including Dr Fiona Edgar. I am also grateful for the comments from the participants at seminars and workshops at Manchester University, Monash University, and the Performance Management Colloquium at University of Otago.

[1] Further information cannot be disclosed for confidentiality reasons.

senior and middle-level managers has been designed to give people rewards for high levels of performance, and half a managers' cash bonus was paid based on their performance evaluation rating.

The important features of Sensol's objective setting and performance evaluation processes for this paper is that despite the formal documentation appearing to be very objective and comprehensive, superiors had discretion over the choice and weightings of objectives and measures (e.g., financial, non-financial, personal factors) in subordinates' performance evaluations. While there is a substantial literature on organisational justice, what this paper contributes is a deeper understanding of the complexities (e.g., injustices) that we need to consider when linking subjective performance evaluations to cash bonuses and the consequences that result from the perceived injustices. The findings that show that the lack of procedural, distributional and interactional justice leads to consequences such as individuals not trusting the performance evaluation system, paying generous bonuses and not based on performance, gaming, and resource allocation issues (e.g., promotions). This has resulted in a number of unintended consequences including that people generally felt that they had little impact on their performance evaluation rating, and did not trust the system.

An important contribution of this paper has been to enable us to learn from evaluating the outcomes associated with a seemingly ineffective performance evaluation system. This is in contrast to papers that report on successful changes due to the introduction of a fair performance evaluation process (see Taylor et al., 1995).

The next section briefly examines literature. This is followed by the research method, background to the performance evaluation system, findings (e.g., procedural, distributive and interactional justice), discussion and conclusion.

Literature review

The use of pay-for-performance plans is growing.[2] This includes variable merit pay plans where individual performance evaluation ratings are linked to annual cash bonuses which need to be re-earned each year (Heneman, 2002). Merit pay requires developing comprehensive measures and standards, and a system that clearly ties individual performance to standards (Lawler, 2000). These processes often include the use of tools such as management by objectives (MBO) (Hubble, 1972; Merchant and Riccaboni, 1990; Pratt, 1991) and personal scorecards (Kaplan and Norton, 1996, 2001). While merit pay plans seem like a good idea, there has been little research examining their success (Levy and Williams, 2004). The use of performance evaluation processes in organisations is not new and can be traced back to the Wei Dynasty in China, although it was not until the 1950s that there was widespread use of performance evaluations for managers (Pratt, 1991).

The main advantages with individual merit-based performance evaluation processes include the line of sight (e.g., ability to influence) over individual performance and rewards (Lawler, 2000), and that they reward soft behaviors such as the personal factors, as well as results or output controls (Heneman and Gresham, 2002; Heneman, 2002). Behaviour controls are important where the results or output measures are uncontrollable (Heneman and Gresham, 2002). The main disadvantages include misleading data, inaccurate ratings, unrealistic goals,

[2] Other common terms for pay-for-performance plans include incentives, rewards, compensation, pay, bonus (Merchant et al., 2003).

lack of valid measures (Lawler, 1990, 2000). Performance evaluation ratings lead to rewards (e.g., bonuses, promotions), and so equity and justice perceptions are very important.

The study of fairness started with Adam's equity theory (Adams, 1965) and focused on fairness of outcomes (distributive justice). This paper focuses on how the operation of Sensol's performance evaluation processes impacts on organisational justice issues such as distributional, procedural and interactional justice (Chan Kim and Mauborgne, 2003; Greenberg, 1990; Heneman and Gresham, 2002).

Distributive justice is the perceived fairness of outcome distributions (Greenberg, 1990). For example, performance evaluation ratings for individuals should be based on the performance achieved and are consistent for people with similar performance (Greenberg, 1990). The theory predicts that employees will react to injustices by focusing on specific outcomes and can affect an individual's emotions, cognitions, and their behaviours (e.g., performance or withdrawal) (Cohen-Charash, 2001, p.280-1). This is an important issue in performance-based management systems.

One of the common ways for organisations to ensure distributive justice in terms of consistency in ratings across diverse business areas is to use an averaging or forced-distribution process. This is a guide to indicate to managers how many people should be rated in each category (Heneman, 2002; Lawler, 2000).

Equity research has moved towards studying procedural justice, since findings have suggested that the distribution of rewards was not as important as the perceived fairness of the process and policies used for decision making (Cohen-Charash, 2001, p.279; Greenberg, 1990). The theory predicts that employees who experience procedural injustice will react negatively to the organisation (Cohen-Charash, 2001, p.281). Research also shows that superiors and subordinates respond more favourably to fair performance evaluation system (Lawler, 1990; Levy and Williams, 2004).

Interactional justice is a recent extension of procedural justice. This relates to aspects of the communication process between the superior and subordinate and includes items such as politeness, honesty and respect (Tyler and Bies, 1990 cited in Cohen-Charash, 2001). The theory predicts that where a subordinate perceives injustice, he/she will react negatively to the superior, rather than to the organisation; but there is little evidence of the consequences of perceived interactional injustices (Cohen-Charash, 2001, p.281). A related point is the importance of giving employees a voice in the performance evaluation process that results in determining their individual ratings (Detert and Burris, 2007).

The following discussion focuses on the consequences of an ineffective performance evaluation system, its impact on perceptions of organisational justice, and the consequences that will likely result for the firm such as lack of motivation and impacts on performance (e.g., it is unlikely to realize superior performance in the workforce).

Research method

This interpretive case study was conducted in Sensol,[3] a large decentralized Australasian organisation which was established in the late 1980s and by 2000 had revenues of around $1

[3] Sensol is a fictional name used for confidentiality purposes.

billion, and 9,000 employees. Sensol was successful in the early 1990s in being innovative and creating new business opportunities from these innovations. By the late 1990s, Sensol's large core business was in a declining industry due to technological substitution and its strategy was to cut costs and to diversify into new high-end technology businesses. Falling revenues in Sensol's core business had a large impact on its profitability because a high percentage of its costs were fixed. This resulted in a 'shock' loss to the organisation at the beginning of the study. At the time of the study the business units reported to four operating groups, and there was a corporate group (i.e., support business units).

This study started with a broad question, examining Sensol's overall performance management framework (e.g., objectives, strategy, measurement, target setting, incentives) (Stringer, 2006). The problems with the performance evaluation process emerged as important, especially leading up to the end of the financial year in the first year of the study (2000-2001). These same problems were evident the following year (2002) by which time two external consultants were commissioned to review the performance evaluation and incentive systems. I was a participant-observer at this review.

While intensive and longitudinal field studies are rare, they are needed to gain a deeper understanding of complex phenomena, especially where there are sensitive issues such as performance evaluations (Lee cited in Merchant and Riccaboni, 1990). Spending time in Sensol has been important to gain an understanding of how performance evaluation processes are *used* by managers, and how this compares to the formal design.

The multiple sources of data for this 2½ year study include open-ended interviews (over 200 have been conducted), informal conversations, extensive periods of observation, some participant-observation, extensive documentary analysis, field notes, external consultancy reports, and performance data. The performance evaluation data includes three years of data 2000 to 2002 including data for the individual ratings (e.g., by grade), and objective ratings (value scorecard performance) for over 700 managers for 2000 and 2001, and data for 2002 is provided by internal reports. A feature of this interpretive study has been to use qualitative and quantitative data (often missing from interpretive studies, see Norreklit, Norreklit and Israelson, 2006; Mattimoe and Seal, 2009).

The interview data was arranged into broad categories (e.g., performance evaluation, performance measures, incentives) across interviews and by business unit using a matrix-type format similar to that suggested by Miles and Huberman (1994). Then, patterns and themes in the empirical data were identified across interviewees and business units to gain a deeper understanding of Sensol's performance evaluation processes (see Ahrens and Dent, 1998). The pattern model involves identifying patterns across interviewees and across business units to understand the dynamic nature of Sensol's PMC. An inductive, rather than a deductive approach has been taken. An inductive approach is commonly used in empirical studies (e.g., Llewellyn, 2003). Patton (1990, p.390) has described this as:

> *Inductive analysis means that the patterns, themes and categories of analysis come from the data; they emerge out of the data rather than being imposed on them prior to data collection and analysis. The analyst looks for natural variation in the data.*

Data collection, data analysis and interpretation were on-going. Rather than a linear process, this process has been iterative, in that issues would be raised and then more evidence would

be collected to clarify outstanding issues. At the end of each visit to Sensol there were long periods of reflection and discussion moving back and forward between the empirical data, the literature, and discussions with colleagues to try to understand the reasons for the observed patterns in behaviour. This approach is similar to what Alvesson and Skoldberg (2000) call reflexive interpretation. This inter-play between empirical data and theories has been on-going through the data analysis and write-up phases.

Background and performance evaluation design

During the late 1990s Sensol's board pushed for a move to a performance-based culture which a senior manager described as where "*we all knew the score*", what the "*game plan*" was, everyone was aligned, and there was "*a robustness and a transparency about the outcome.*" An important objective of Sensol's performance evaluation processes was to hold managers accountable for their performance. As this senior manager explained:

> *So we're trying to promote a culture of accountability, it's really around trying to move from a very strong trust loyalty service type culture to ... how do we get a bit sharper on some of the performance things and actually holding people accountable ... so let's get our expectation clear, let's give good quality performance information to people to assess how they're performing to those… then you go into assessment and appraisals and reviewing and looking forward and then importantly – and we're not always great at this – is the consequences. So that is about promoting those, it's about offering training and available opportunities; it's about other recognition systems. It's the remuneration impacts. Alternatively – moving people out.*

Another senior manager raised concerns over how much actual change had occurred, and whether these 'tools' were being used:

> *[Sensol] tends to be generous… I'm not trying to say that that the company is in anyway soft but the strict regiment around setting objectives, measuring objectives, reviewing of objectives and there being consequences as a result of those reviews, whether they be positive or negative, isn't as entrenched as one might expect or from senior management perspective one would desire. So we have a framework … while the tools are there, how well those tools are used? … The reviews are expected to take place but it's not inspected. Is everybody actually doing this? There are questions around all of that. (2001)*

Sensol's remuneration strategy aims to set managers' fixed pay at below market rates, and then to allow managers to reach market rates through annual bonuses paid when the target were achieved. Managers could also earn above the market pay rate if they exceeded their performance targets. Sensol had been increasing the percentage of variable pay over time.

The formal design of Sensol's annual incentive scheme for its top 700 managers was very complex and included two key parts. Fifty percent of a manager's bonus payment was linked to an individual's performance evaluation rating (i.e., variable merit pay), and the other 50% was formula-based and linked to achieving pre-set targets for the four perspectives of Sensol's value scorecard (similar to Kaplan and Norton's 1996 perspectives of financial, customer, internal, growth). The base amount of the bonus varied between 8-25% and the amount of bonus a manager had 'at risk' increased by grade. A manager's maximum (or

potential) bonus could equal four times the base amount, if the manager achieved 200% for individual performance, and 200% if Sensol reached maximum performance to the value scorecard targets. Potential bonuses could be substantial at higher levels. This paper focuses solely on the first part of a manager's bonus that was linked to individual performance evaluations.

Sensol's managers were expected to be evaluated based on results achieved compared to the objectives agreed at the beginning of each year with their superiors. This included around five output or result objectives that managers could influence and that were linked to Sensol's value scorecard (e.g., EVA®, customer satisfaction, health and safety, or value drivers such as % of cost reduction, project milestones),[4] as well as behaviours called personal factors (e.g., accountability, team work, enterprise, customer focus, valuing people, critical thinking). These objectives were to cascade down from Sensol's value scorecard, and were to be tailored to what individuals could influence.

At the end of the year, a manager and his direct superior would prepare for the performance evaluation discussion, and then meet to agree on the performance evaluation rating. Managers were evaluated based on the merit rating scale that ranged from 0 to 200%: a rating of outstanding (126-200%), achieves target (76-126%), acceptable (25-75%), unacceptable (0%), and too early. Human resources issued guidelines to assist managers in evaluating their staff within these broad bands. For outstanding performance the guidance stated, "You have significantly exceeded many of your objectives; and you have been highly effective in all aspects of personal factors."

Once agreement was reached, the direct superior discussed the evaluation with his superior, and the individual rating was signed off and sent to human resources. There was also a process whereby individuals could ask for a review of their performance ratings. The human resource managers controlled the process in several ways: they required copies of all objectives to be sent to them, provided training sessions and support, and required average distributions across the diverse business areas (i.e., forced distributions). The formal design was similar to the recommendations of the MBO and personal scorecards (Hubble, 1972; Kaplan and Norton, 1996).

Sensol's 'averaging' process (forced-distribution) ensures the ratings across the business units are consistent. The averaging process was also important to help keep the incentive payments within budget, because of the iterative effect on reducing profit, and these issues were discussed at the highest levels. The average across business units was expected to be around 120%.[5] The impact of the way the averaging process has been *used* in determining the performance evaluation ratings will be discussed under interactional justice.

Despite the senior managers stating Sensol's system had changed towards a performance-based focus, there appears to have been little change to the formal performance evaluation processes from the early 1990s until 2002. A manager reviewed his performance evaluation review back to the early 1990s and found there were still four key result areas in the early 1990s. A good representation of the current system is evident from the diagram presented in

---

[4] Sensol's balanced scorecard was called a value scorecard because 80% of the weighting was on EVA, 5% on revenue growth, and 15% on non-financial measures. Sensol's four perspectives have been changed to Kaplan's and Norton's perspectives to ensure confidentiality.

[5] Making the average 120% not 100% would take into account higher performance.

Appendix A (dated 1995) that shows the link between Sensol's scorecard, a manager's objectives, performance ratings, and the annual bonus.

Findings

The performance evaluation process was extremely complex, stressful, and time consuming process for superiors and subordinates, and the pressure has intensified because they are linked to cash bonuses. A model of the operation and consequences of Sensol's subjective performance evaluation process is presented in Figure 1 and the findings are discussed in terms of procedural, distributional and interactional injustices.

Insert Figure 1 about here

Procedural justice

Managers discussed procedural injustices in terms of the mixed practices, unclear criteria, and the financial focus despite having a range of financial and non-financial objectives.

Mixed practices

There were mixed objective setting practices across the business. Initially, managers told me they had objectives and some examples were given to show how these objectives cascaded down the hierarchy and linked to Sensol's value scorecard, similar to that outlined in the formal documents (see Appendix A). In some areas there was evidence that the performance evaluation practices cascade down the hierarchy, as this manager explains:

> *If you can imagine my goal, my [subordinate] managers' goals... To meet my goal their staff to meet their manager's goal and it's a pyramid upwards and it's at the best level of understanding of how they are. Their performance is so critical to the success of the overall and it's introducing the right sort of behaviours. At the same time it is fairly new still.*

However, a senior manager from an area who closely followed the formal objective setting guidelines observed that, *"other businesses don't comply ... and this is frustrating."* A new manager said the prior objectives for his team seemed to be more subjective than results based, and he was surprised to find that the objectives on file were different from those held by his subordinates. At the other extreme one group even had the managers' objectives published in a black book one year, but these managers had no objectives the following year due to major changes in the business. Often managers would say they had objectives, but few people showed them to me. Some managers said they would email them to me, but when few arrived I assumed it was because of the sensitive and personal nature of the information. Later on, human resources advised that they also received few copies of objectives. This situation is not surprising as the process seemed to be ambiguous and there was little in the way of clear guidelines and training for evaluators. I attended a training session for new managers that included the performance evaluation and incentive scheme. It was held late at the end of a three-day workshop, on a Friday afternoon when the new managers were exhausted.

There were also variations in the way performance evaluation processes were *used* across different business units. At one extreme, some managers observed that in their business unit the objectives signed at the beginning of the year were strictly adhered to in the performance evaluation process. They stated that it did not matter what had changed during the year (even

their reaction to an emergency situation), if something was not noted in a manager's objectives at the beginning of the year, then it was not "*counted*" in the performance evaluation process at year-end. This extract from an external consultant's report (2002) refers to a focus group session where managers also highlight the mixed practices in terms of budget emphasis across the business units:

> *[one group] in particular believed the budgets had far too great a weighting, were applied unfairly and impacted on making sensible spending decisions, i.e. decisions were often made with a very short term focus (to keep within budget at all costs) and not always in the best long term interests of the company some spending was beyond your control but no allowance was made if this took you beyond your budget. Actual against budget was strictly adhered to (at appraisal time) even if there were realistic reasons why the budget had been exceeded. Other members of the focus group said they did not have the same issue, their managers adopting a more realistic/flexible approach when it came to agreeing actual against budget at year end.*

There was also diversity in the timing of the performance evaluation meetings. Some managers described having six-monthly meetings with their managers, most managers said they had some type of performance evaluation meeting at year-end. A few managers described how they emailed their senior manager about what they had achieved and received back their personal ratings with no explanation or discussion. Another manager described a quick five minute meeting at an airport. There were similar variations at senior management level. I was present when some senior managers discussed receiving an email from the CEO requesting them to rate their own performance.

Another problem was that the performance evaluation meetings served two conflicting purposes. Firstly, they were to help with development of managers in terms of where were the problems, where did they need training, etc. Secondly, they were to agree on the individual rating for the cash bonus. The problem is that a manager is hardly likely to say I am having trouble in these areas, when at the same time he is arguing that he had been a good performer and deserving of a high(er) rating.

One reason for these mixed practices may be due to the fixed nature of the objective setting process that does not take into account the dynamic business environment. A senior manager said *"you give me a job description today, it's obsolete in three months... so you can't do objective setting once a year"* (2000).

The focus in this section is on detailing the case study, and the implications of these practices are discussed in the next section.

Unclear criteria: What is required to "walk on water"?

Managers generally said performance evaluation processes were subjective and that ratings depended on many issues, including their relationship with superiors (e.g., *"if you suck up to the boss*"), and other biases (e.g., recent events). Performance evaluation processes appeared to work differently in different business units, and managers talked about the variety of approaches taken by their superiors (e.g., "*hard"* versus "*soft"* evaluators).

Most managers did not know what type of performance was needed to get particular ratings. Similarly, an external consultancy report found that, "There is no clear differentiation between a "76% performance" and a "125% performance" and the rating bands are too broad". Individuals who received performance evaluation ratings over 125% were said to *"walk on water"*. These are the managers rated as "outstanding" (see row 1 in Table 1). While the ratings were secret, there was lots of gossip about the ratings 'other' people get and who *"walks on water"*. These were typical comments from managers:

> *Over 100% some must be sensational people ... Should you need to 'walk on water' be above 100%? (2002)*
> *It's difficult, paying you to do job should be 100% not 150%, people think if they've done their job well and met objectives, they should get a higher rating. There are problems with consistency between groups and individuals. Subjective, difficult to compare. (2002)*
> *To be above 165% CEO has to be involved, to 'walk on water', and he has strong views. (Senior manager, 2001)*
> *The individual multiplier is open to interpretation, what is outstanding performance?*

Insert Table 1 about here

Financial focus – "double dipping"

Financial performance (especially to budget) was very important, and managers' perceptions were that this emphasis had increased over time. This resulted in a situation called "*double dipping*" were managers' perceived that they were rewarded for good financial performance in the objective formula-bonus based on the scorecard and in their individual performance evaluation. A senior manager said, "*There are different views ... whether the individual rating is different to business performance. Subjective.*" Similarly, business unit managers generally perceived that they were evaluated mainly on the bottom-line performance of their business units. This is in contrast to the formal processes that required managers to be evaluated compared to the pre-set objectives that were to be linked to Sensol's value scorecard. Managers were supposed to have 5-6 financial and non-financial objectives (e.g., achieve project milestones), as well as 'other' factors (e.g., team work, accountabilities). The perceived dominance of financial performance and "double dipping" has led to considerable gaming (e.g., smoothing, target setting).

While it appears to be contradictory, managers also perceived that there were many factors that influenced their financial performance measures that they had little ability to control. These include arbitrary transfer pricing, charges for support costs, restructuring, strategy changes by senior management, degree of reliance on other business units, etc.

Distributive justice

The distribution of the individual multipliers for three financial years for 2000, 2001 and 2002 is presented in Figure 2. The peaks in Figure 2 showed how most people were rated around 100% and 115% in 2001; with 110% and 125% in 2002. It is also interesting that there seems to be some sort of ranking (e.g., 110, 115, 120). While managers could potentially be rated between 0-200%, Figure 2 also supports managers' perceptions that ratings of around 200% were "impossible".

Insert Figure 2 about here

The following discussion includes the way different levels of performers are deal with, stickiness in ratings, conservative versus tight budget targets, and higher ratings at higher levels.

Little differentiation between good versus poor performers

The aim of Sensol's performance evaluation processes is to reward good performers and encourage accountabilities, so a reasonable assumption is that the performance evaluation processes would differentiate between good and poor performers. During the initial phase of the study, a corporate manager stated that Sensol was, "*much more rigorous about managing those people out of the business than we have been to date.*" However, the extent to which things had changed is debateable. Another senior business unit manager concluded that:

> *The level of consequence of non-performance or over achievement is not what it ought to be either. So if someone is not performing well enough there is this generous sort of, almost paternal type behaviour that carries on and that person gets looked after*.

A major problem was that evaluators tended to avoid dealing with poor performers. One manager said these were "*weak knee-ed*" managers, and another suggested that managers who could not deal with poor performers should be fired. An external consultant also reported that some managers were "*soft*" appraisers and not prepared to make the "*hard*" decisions. An external report stated, "Some managers saw the individual multiplier as a way to "be nice" to certain staff members, while others liked having the discretion to differentiate based on individual performance." A newly appointed manager observed that Sensol, "*does not seem to have competency in the management structure.*"

Dealing with poor performers has been a problem and this was the consistent view of most managers, human resources, external consultants and the incentive review team. The archival data also supports the view that there is leniency in ratings as a high percentage of people are rated outstanding (18-21%) and the few people who are rated 'unacceptable' (2-3%) and that this is consistent over a three year period (see Table 1). The concentrated and skewed nature of the performance evaluation data was highlighted in internal reports (see Figure 2). Sensol's performance evaluation and incentive scheme report (2002) that stated, "The key issue may be not that still too many people are rated "Outstanding, but that a very low percentage are rated as anything below 'Meets target'. Is this reality?"

Stickiness in ratings

Related to the (not) dealing with poor performers, is what has been called the "stickiness" in the performance evaluation ratings. Where a manager will be promoted to a particular grade and once he/she receives a high rating one year, this is likely to stick at this rating, until he/she is promoted again, and so on. This quote illustrates the problem:

> *I've had HQ* [the old term for "walk on water"] *service since I've been here. Right – I'm telling the truth, honest. And I'm thinking to myself, well I'm just doing what I'm expected to do. Now, if I didn't get a HQ I'd freak out.* [laughter] *Serious! I would like, massively freak because I would think what*

the [x] *am I doing. At the same time though, it has not incentivised me to do any more than I am doing. If they wanted me to do more than that, what should they do?* (2000)

This situation meant that people assumed that they should get high ratings for doing what was "expected", rather than for good performance (more than just "doing your job"). This is more evidence that the type of performance should be rewarded in the incentive scheme seemed to be unclear across the organisation.

Another reason given for the stickiness in ratings was that it was an easy way to deal with aggressive subordinates. This quote indicates that some managers give individuals the same high ratings year after year to avoid confrontation.

We work quite hard to make sure [the old boy's network] doesn't happen. Yet it happens and in some cases some people who haven't been performing very well, because they've been strong personalities for example, they've actually got away with things and been recognised, it's easy to go and say "you're doing really well," when you're not doing really well, than confront you and have you go off your nut. And so difficult personalities have not got the right sort of feedback, fair feedback and that's caused anomalies further on down the track, when you turn around and say, "well you're not performing." They turn around and say, "Look I've had Grade A assessments for the last five years and I'll take a [personal grievance] out against you because you don't like me. You're just trying to get rid of me. I haven't done anything different this year than I've done for the last four years and I've got these great assessments in the past." And I just go "ah." (2001)

Conservative or stretch budget targets?

There were issues over the perceived unfairness in the degree of stretch in budget targets across the businesses. Two of the business groups reached their maximum budget targets in 2002 (i.e., good performers) and two of the business groups did not meet their budget targets in 2002 (i.e., poor performers). Many managers perceived that there was a lot of gaming going on, and some senior managers appeared to be able to consistently negotiate conservative (capped) targets.[6]

Many people also perceived that financial performance of their business unit impacted on their individual performance evaluation (i.e., "double dipping" discussed earlier). If financial performance did impact on individual performance then it could be expected that the average performance evaluation ratings should be higher in the business units that reached their maximum targets that was supposed to indicate outstanding performance. In contrast, Sensol's performance evaluation and incentive scheme review (2002) reports that the averages across five groups ranged from 110% to 119%. What was even more interesting was that, in 2002, managers in the two poorly performing business groups received (on average) the highest individual performance evaluations ratings (i.e., averages of 118% and 116.5%) compared to the average performance evaluation ratings for the other two good performing business groups (i.e., 114.25% and 115.75%). One reason for this could be that the high

[6] Incentive targets were capped with a threshold, minimum, budget targets (achieve 100% of base amount) and maximum (200% of base amount).

performance evaluations recognised good managers who were turning around poor businesses.

The general view was that the senior managers in the groups had used individual multipliers to increase or decrease managers' overall incentive payment (Sensol's performance evaluation and incentive scheme review, 2002). This may have been to reward managers for other factors in poorly performing businesses. On the other hand, the managers' individual ratings in the well-performing businesses were reduced. This may have been to ensure equity reasons with other businesses (i.e., distributive justice), or concerns over the generous nature of the overall bonuses that resulted from the "multiplier effect". The multiplier effect meant that if you received a high individual (value scorecard) rating this was multiplied by the value scorecard (individual multiplier), rather than being added to (see footnote for an example).[7]

Higher ratings at higher grades

The distribution of performance evaluation ratings up the hierarchy, also raises questions over the fairness of the ratings. Sensol's performance evaluation data shows that, on average, the senior managers receive the higher ratings, and lower-level managers the lower ratings (see Table 2). The process required superiors to rate the performance of their subordinates between 0 and 200%. For example, the average for managers at Grade A was 175% which translated to a multiplier of 1.75 that was used to calculate part of a manager's bonus in 2000. The higher management grades are shown at the top of the table (from highest to lowest). Table 2 also shows that the ratings for other specialists (professionals who do not have subordinates) were much lower than for managers, and the findings are consistent over the three-year period.

Insert Table 2 about here

Analysing the actual performance evaluation data for 2001 provides some additional insights into Sensol's performance evaluation processes. As a hierarchical organisation, Sensol had numerous grades for managers and specialists. Using the list of grades from Table 2, they have been classified into four categories: senior managers in grades A-F (category 1), lower level managers in grades G-L (category 2), highly paid specialists grades P-R (Category 3) and lower paid specialists grades S-U (Category 4) (see Table 2 for mean ratings by grade).

The Kruskal-Wallis tests shows that the mean ranks increase up the four categories, and Chi-Square shows significant differences at 0.000 level of significance between the categories.[8] Tests show that there are significant differences between categories one and two and all other categories at the .000 level for senior managers (category one) and lower level managers (category two).[9] The significance level between the two higher and lower specialist categories (three and four) are significant at the 10% level. The significantly lower ratings at lower grades are not surprising in hierarchical organisations like Sensol, where management grades are considered to be more important.

---

[7] An example of the multiplier effect is that a manager with a base amount of $10,000 with a formula based multiplier of 1.5 (based on scorecard measures), multiplied by the performance evaluation rating of 1.2 (as determined by the superior) would result in an annual cash bonus of $18,000). Whereas, A 10,000 bonus multiplied by 1.5 and again by 1.5 = $22,000 (called the multiplier effect).

[8] There is an in-balance in cells. Of the 747 individuals in the data file, there are 38 individuals in category 1, 303 individuals in category 2, 211 individuals in category 3, and 195 individuals in category 4.

[9] The results are consistent using the Jonckheere-Terpstra test, Tukey HSD, Scheffe, Bonferroni, Tamhane, Dunnette T3.

The results were consistent in 2002. An internal report stated that the average multiplier for the senior managers (category 1) was 141% (range 120-175%), and no one in this grade received less than 120% (compared to 00/01 average of 146%; and 99/00 average 143%). Sensol's performance evaluation and incentive scheme review (2002) stated, "There is still a skew of higher multipliers predominating at the higher grades… [and while lower in 2002] arguably, this average is still too high" (see Table 1). For lower level managers (category two) in 2002, the average multiplier was 113% (range 65-150), and for lower level specialists (category four) the average was 111% (range 75-133**%).** These are several possible reasons for the higher ratings at higher grades including the CEO expected the senior managers to get the higher grades or "*walk on water*", and one senior manager said the *"expectations of senior managers are higher."*

Over the two and a half year-period of this study, only a few people mentioned that the higher ratings were mainly given at higher grades. One senior manager said, "*You will find different averages in the bands I bet ... that the average score at the lower bands is lower than the highest bands.*" This is not surprising as the high level of pay secrecy at Sensol. For example, at the end of the study data on the distribution of the performance evaluation ratings was not allowed to be provided to managers who were part of the annual incentive scheme. However, anecdotal evidence suggests that this knowledge was embedded in the logic used to estimate the bonus calculations.

Interactional justice

The main issue here is the "*predetermined*" theory that the ways the performance evaluation ratings are determined and communicated have resulted in managers' perceptions that the system is "*managed*" and their perceptions that they have little ability to influence their individual ratings.

The predetermined theory

While the forced distribution process was designed to ensure fairness in average ratings across the businesses (as discussed above), the way this was the process was *used* in practice meant that it has led to injustices in the way the communication process worked including the discussion between managers and their subordinates in determining the final rating.

One of the reasons is that to achieve an average rating of 120% across managers in a business unit, senior managers had to "manage" the process. Usually, the superior managers doing the performance evaluations usually met with their own senior managers prior to meeting with their subordinates. A new manager advised that he had worked out the ratings for his subordinate managers, but had not disclosed the ratings to his subordinates as he needed to get approval for his ratings from the senior manager of the business unit who would check the average ratings across the business unit. This appeared to be typical across the business units.

This is in contrast to the way the process has been designed, as discussed earlier. One manager said that only ranges were discussed with superiors, and, *"If it is predetermined why have a meeting?"* However, the general perception was that the ratings were pre-determined and that any subsequent discussion had little effect on changing the ratings. As one external report stated, "Managers are asked for the appraisal ratings of their staff before the appraisals are completed and this further supports the predetermined theory".

Managers generally felt that because of the "predetermined theory", they felt they have little voice or ability to influence their individual ratings, resulting in a lack of trust in the system. One manager stated, "*Individual multipliers are predetermined by the General Managers [of the business units] and therefore it is impossible to influence the setting of your multiplier*". Some managers talked about this and that the ratings were not linked to performance.

This lack of voice in the performance evaluation meetings was also evident in other meetings. I attended a meeting Sensol's top 80 managers, and the CEO's session on strategy felt more like a lecture, to the extent that the facilitator told managers they were not "*engaged*" in the discussion. One manager commented that if a good idea was raised at a meeting by a lower level manager, it was more likely to be dismissed, than if it was raised by a more senior manager. Managers also talked of the importance of being "*closer to God*" (the CEO). The change to the group structure meant that only five senior managers reported directly to the CEO (instead of 17). This lack of a direct reporting line to the CEO was a concern to the other business unit managers. Another manager described Sensol as having a "*paternalistic culture, which essentially is at the core and then Dad or God will tell you the answer*".

Discussion and conclusion

Subjectivity in performance evaluation processes, for example, leads to feelings of injustices (e.g., mixed practices, little differentiation between poor and good performers, and the predetermined theory). These injustices have led managers to feel that the performance evaluation and incentive systems were not to be trusted and the results have been "managed" by senior management and head office to achieve a certain result.

Several key themes highlight the injustices managers felt about the subjective performance evaluation process. There are mixed practices across the businesses and by managers. Some areas were evaluated in a similar manner to Hopwood's (1974) budget constrained style, and others more like a profit conscious or non-accounting style. The criteria used to evaluate managers are often unclear and some people talk about it being a black box process. Managers' perceptions are that the individual ratings are financially focused, despite objectives including financial and non-financial measures. This perception has resulted in "double dipping" where managers' perceive that they were rewarded for good financial performance in their individual and scorecard ratings. This is consistent with Ittner et al. (2003, p.754) who found that the use of a scorecard in an organisation that had previously had a financial focus, resulted in financial measures continuing to dominate performance measurement, consistent with psychology-based predictions, despite the use of the value scorecard and rhetoric that non-financial measures are important.

There has also been little differentiation in ratings between poor and good performers. The compression of results has been found in prior studies (Moers, 2005). The problem is that "It in effect says to the high performer, 'Your performance doesn't make any difference'" (Lawler, 1990, p.74). It is not uncommon to find leniency in performance evaluation ratings (see also Gibbs, 1995; Merchant and Riccaboni, 1990; Moers, 2005; Otley, 1992). Erdogan (2002, p.574) argues that high performance ratings could lead to higher performance in the future, because of the need to reciprocate the superior. The problem is that if a superior gives unjustifiably high performance ratings to an individual, perceptions of accountability may suffer, and this can lead to reduced performance. Equity theory is about fairness and distributive justice and so it is a problem that managers' perceptions are that there was stickiness in the ratings. This

presents an opportunity for future research to see if this actually happens by tracking the ratings of managers over several periods.

The ratings across the business units are also interesting. There was debate about the degree of target difficulty across the businesses. It appears that people in business units with high performance (and high bonuses) received lower individual ratings (and vice versa). There is some evidence that this might be for equity reasons, or to counter the (already) generous nature of the bonuses. This suggests the ratings may be an adjustment (upwards or downwards) to a managers' cash bonus (see Stringer and Pillai, 2009).

The impact of the Sensol's averaging process has been that for some people to be rated high, others have to be rated low. It is unclear the reasons that senior managers get the higher ratings. It may be that the senior managers are the best performers. Lawler (1990, p.6) argues that higher level managers have a great deal of responsibility and their pay should be tied to organisational performance (Lawler, 1990, p.6). Or the ratings could be linked to market forces (e.g., used as a "fudge factor" where there is a retention issue), or it could be that the system is "managed" by those at the top of the hierarchy to give themselves the higher rewards.

While it could be argued that many of the problems could be overcome with improved training procedures for managers, it is also that people are busy and ran out of time to conduct the formal appraisals, another reason is the impact of organisational context (e.g., culture).

Sensol's hierarchical culture appears to have had a strong influence on the ratings. For example, the number of people who reported to a particular manager was important in terms of grade and pay, and also indicated the status of a manager. Lawler (1990) states that, "Traditional cultures emphasize top-down decision making, vertical communications, and clearly defined jobs, while involvement cultures emphasize shared decision making, lateral communications, and loosely defined roles" (cited in Heneman and Gresham (2002, p.92). The problem is that managers who want hierarchical structures are unlikely to be the people who operate effectively in a pay-for-performance culture which encourages the involvement of lower level managers (Lawler, 2000). A related issue is that Sensol's specialists (who have no subordinates) had lower pay, consistent with other hierarchical organizations as their jobs are less highly valued (Lawler, 2000).

Sensol's use of an averaging process further complicates the performance evaluation process in ways not intended by the designers. This rule creates a problem because the performance of most groups does not match the bell curve. Some organisations deal with this by telling low-rated employees that it is their turn this year, and the ratings are rotated (Lawler, 2000 p.186). The way this was "managed" by the senior managers has led to managers' perceptions that they have little voice in the performance evaluation process. The degree of voice in the process was affected by supervisory discretion on choice and weights (Detert and Burris; 2007, Cawley et al., 1998; Korsgaard et al., 1999).

The subjectivity in performance evaluations has resulted in managers' perceptions that the system was unjust and this has led to several interrelated consequences. The main consequence was the lack of trust in the performance evaluation process. People generally felt the system was "managed" by corporate and senior managers. This is a major problem for performance-based systems, as Chan Kim and Mauborgne (2005 p. 134) argue that, "Fair

process builds trust and commitment, trust and commitment produce voluntary cooperation, and voluntary cooperation drives performance, leading people to go beyond the call of duty by sharing their knowledge and applying their creativity"… and results in dysfunctional behaviours when people do not trust the system. It is also difficult to motivate people who do not trust organisational systems. These findings are in contrast to agency-based economic theory that suggests that people are mainly driven by their own interests, whereas findings from this case study and Chan Kim and Mauborgne (2005 p. 134) show that people do care about the outcomes and the processes the produce these outcomes.

Other studies have also found trust is important in performance evaluation processes. Mayer and Davis (1999) find trust is more likely where a subordinate believes the superior has the skills to properly appraise, takes into account the interest of the subordinate, and upholds the standards and values required. Agency theory suggests that as supervisors bear all the monitoring costs, but receive little of the benefit, this can also lead to lack of trust in superior's evaluations (Baker et al., 1988 ; Gibbs et al., 2004). Gibbs et al. (2004, p. 410-11) argue that the use of subjectivity may not always be effective because it includes discretion. "Subjectivity works well only if the supervisor makes fair, unbiased judgements and if the subordinate accepts the judgements and does not try to influence the supervisor inappropriately." They find that subjective evaluations improve performance only where there is trust between a superior and subordinate. The importance of trust may also explain why Gibbs et al. (2004) find that subjective bonuses have a positive impact where managers have had long tenure with a firm (their proxy for trust). This suggests that trust builds over time.

The generous nature of the bonuses was generally accepted by the end of this study. An external report noted with surprise that people complained about the performance evaluation and incentive systems, despite receiving generous bonuses. One reason may be that following the "shock" loss during the initial phase of the research meant that people were concerned with how long Sensol could continue to pay out generous bonuses. For example, one senior manager stated "*I thought it was a bit embarrassing – it didn't seem to be unreasonable in terms of sharing the benefit but there was an awkwardness about the size of it.*" An external consultancy report noted "… the [the incentive] scheme was too generous – making the payout more generous for those achieving more than 'just doing their job' ", rather than a pay-for-performance scheme. At the end of the day, the administrative and emotional cost of the performance evaluation system was extremely high, for little benefit to the organisation.

Managers generally perceived that they had little voice in the performance evaluation process, and hence, little ability to influence their individual ratings. Managers also raised the issue that there were a range of 'other' factors that they could not influence that affected their financial measures (e.g., Business unit EVA). This lead to gaming (e.g., targets, managing the financial numbers) to increase "performance". Gaming is not surprising as the ratings were linked to cash bonuses. Castellano and Lightle (2005) argue that MBO/Management by results systems necessarily require managers to "manage" the earnings. While the inability to influence individual ratings appears to be contradictory with gaming of the financial performance, Hopwood would argue that a budget conscious approach can lead to job stress and may encourage people to short circuit the system (Hopwood, 1974). The other reason is that while there was gaming of the financial results, the subjective performance evaluations were seen as another black box, and often decided by senior management or the corporate office.

Another consequence is that the 'performance' ratings can influence resource allocation decisions such as performance measurement and promotion. Poorly conducted performance evaluation processes can also worsen the relationships between superiors and subordinates and poor appraisals can lower motivation, and hence lower performance (Lawler, 1990). Other consequences are that the bonuses tend to be expected, and this is especially the case at achieving the budget target as this brings managers' pay up to market rates. Sensol's performance evaluation and incentive system was costly to run, and individual performance evaluations require a considerable investment in time (that many managers did not have to spare).

The use of subjectivity in performance evaluations requires an understanding of the organisational context and behavioural implications, rather than on the technical attributes of the formal system (Ittner et al., 2003). In addition to these perceived injustices, this longitudinal case study also contributes to the literature by focusing on the system in use (not the formal system), and show superiors are consistent in their ratings over the three-year period (consistent with Kane, Bernardin, Villanova, and Peyrefitte, 1995).

The findings highlight that the divide between human resource and performance management experts needs to be bridged. In Sensol, the human resources managers understood there were problems with the performance evaluation process (e.g., poor performance issue, biases), and the accountants understood the problems with using accounting measures (e.g., manipulation) when linked to incentives, but no one had an overall view of the implications of these problems on the operation of the whole system, and potential consequences.

This case study raised several issues for future research. Performance evaluation is a key issue for organisations need to grapple with. This case study highlights how we need to know more about when performance evaluation systems work and do not work, and how they interrelate with the broader performance management framework (e.g., culture, measurement, rewards). Future studies need to access archival data to check out the stickiness of ratings and whether managers are being paid for just doing their job, or for good performance. Another issue is that we need to know more about why performance evaluation ratings tend to increase up the hierarchy. Is it managed at the top? Do senior managers use it to their own advantage? Is it because the system works and the higher level and better performing managers get the higher rewards? Is it because they have more to play for (higher bonuses) at the higher levels? Another issue is that in a performance-based system we would expect the poor performers to be replaced, Sensol's archival data and internal consultancy reports suggest this is not the case.

The limitations of this case study are the same as for other intensive case studies. The findings are sensitive and this has required time to be left before these results can be disseminated. The main limitations are that not all the performance evaluation data is available. For example, I do not have the data file for 2002, and I do not have any way to identify specific managers from the 2000 to 2001 data set.

The serious consequences of Sensol's subjective performance evaluation process beg the question, "Why have pay-for-performance systems, if you are not paying for performance?" Ittner et al. (2003, p.754) also find that the subjective performance evaluations led to unclear and mixed practices and biases that led the organisation to question the value of the incentive plan, and it was subsequently abandoned.

It is important to remember that Sensol remains a successful company, despite having problems with the performance-based performance evaluation and incentive system which has subsequently been changed.

Bibliography

Adams, J.S. (1965). Inequity in Social Exchange, *Advances in Experimental Social Psychology*, 2, 267-299.

Ahrens, T. and Dent, J.F. (1998). Realizing the richness of field research, *Journal of Management Accounting Research*, 10, p.1-39.

Alvesson, M., and Skoldberg K. (2000). *Reflexive methodology: new vistas for qualitative research*, SAGE Publications: London.

Baker, G.P., Jensen, M.C., and Murphy, K.J. (1988). Compensation and incentives: practice vs. theory. *The Journal of Finance*, 43 (3), 593-616.

Banker, R.D., Chang, H. and Pizzini, M. (2004). The balanced scorecard: Judgment effects of performance measures linked to strategy. *The Accounting Review*, 79 (1), 1-23.

Castellano, J.F. & Lighte, S.S. (2005). "Using Cultural Audits to Assess the Tone at the Top", The CPA Journal, 6-11.

Cawley, B.D., Keeping, L.M. & Levy, P.E. (1998). Participation in the Performance Appraisal Process and Employee Reactions: A Meta-Analytic Review of Field Investigations, *Journal of Appplied Psychology*, 83 (4), 615-633.

Chan Kim, W., and Mauborgne, R. (2003). Fair Process: Managing in the Knowledge Economy, *Harvard Business Review*, January, 127-136. Reprinted from 1997.

Cohen-Charash, Y. & Spector, P.E. (2001). The Role of Justice in Organizations: A Meta-Analysis, *Organisational Behavior and Human Decision Processes*, 86 (2) 278-321.

Detert, J.R. & Burris, E.T. (2007). Leadership Behavior and Employee Voice: Is the Door Really Open? *Academy of Management Journal*, 50 (4), 869-884.

Erdogan, B. (2002). Antecedents and consequences of justice perceptions in performance appraisals, *Human Resource Management Review*, 12, 555-578.

Gibbs, M., Merchant, K.A., Van der Stede, W.A., Vargus, M.E. (2004). Determinants and Effects of Subjectivity in Incentives, *The Accounting Review*, 79 (2), 409-436.

Greenberg, J. (1990). Organizational Justice: Yesterday, Today, and Tomorrow. *Journal of Management*, 16 (2), 399-432.

Heneman, R.L., (2002). Merit Pay. In Heneman, R.L. (ed.), *Strategic reward Management: Design, Implementation, and Evaluation*, Information Age Publishing: Connecticut, p.379-393.

Heneman, R.L. and Gresham, M.T. (2002). Performance-based pay plans, 75-109. In Heneman, R.L. (ed.), *Strategic reward Management: Design, Implementation, and Evaluation*, Information Age Publishing: Connecticut.

Hopwood, A.G. (1974). *Accounting and Human Behavior*, Prentice Hall: New Jersey.

Hubble, J. (1972). *Management by Objectives*, Gower Press: England.

Ittner, C. D., Larcker, D. F. and Meyer, M.W. (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard, *The Accounting Review*, 78 (3), 725-759.

Kaplan, R.S., and Norton, D.P. (1996). *Translating strategy into action, The Balanced Scorecard*, Harvard Business School Press: Boston, Massachusetts.

Konovsky, M.A. (2000). Understanding Procedural Justice and Its Impact on Business Organizations, *Journal of Management*, 26 (3), 489-511.

Korsgaard, M.A. & Roberson, L., (1995). Procedural Justice in Performance Evaluation: The Role of Instrumental and Non-Instrumental Voice in Performance Evaluation Discussions. *Journal of Management.* 21 (4), 657-669.

Lawler, E. E. III. (1990). *Strategic Pay, Aligning Strategies and Pay Systems*, Jossey-Bass Inc: San Francisco.

Lawler, E. E. III. (2000). *Rewarding Excellence, Pay Strategies for the New Economy*, Jossey-Bass Inc: San Francisco.

Lee, R.M. (1993). *Doing Research on sensitive topics*. Sage Publications, London.

Levy, P.E. and Williams, J.R. (2004). The Social Context of Performance Appraisal: A Review and Framework for the Future. *Journal of Management*, 30 (6), 881-905.

Llewellyn, S. (2003). What counts as ‘theory” in qualitative management and accounting research. *Accounting, Auditing and Accountability Journal*, 16 (4), 662-708.

Mattimoe, R. & Seal, W. (2009). *Pricing in a service sector context: accounting and marketing logics in the hotel industry*, seminar paper presented at Department of Accountancy & Business Law, March.

Merchant, K. A., and Riccaboni, A. (1990). Performance-based management incentives at the Fiat Group: a field study. *Management Accounting Research, 1*, 281-303.

Merchant, K. A., Van der Stede, W., and Zheng, L. (2003). Disciplinary constraints on the advancement of knowledge: the case of organizational incentive systems, *Accounting Organizations & Society*, 28, 251-286.

Mayer, R.C. and Davis, J.H., (1999). The effect of the performance appraisal system on trust for management: a field quasi-experiment, *Journal of Applied Psychology*, 84 (1), 123-136.

Miles, M.B. and Huberman, A.M., (2004). *An expanded Sourcebook, Qualitative Data Analysis*, Second edition, Sage Publications: California.

Moers, F. (2005). Discretion and bias in performance evaluation: the impact of diversity and subjectivity, *Accounting, Organizations & Society,* 30, 67-80.

Nørreklit, L. Nørreklit, H. and Israelsen, P. (2006). The validity of management control *topoi*. Towards constructivist pragmatism, *Management Accounting Research 17*, pp. 42–71.

Otley, D. (1992). United Bank: A Case Study on the Implementation of a Performance-Related Reward Scheme, Chapter 4, 97-121, In Bruns, W.J. Jr. (ed.) *Performance Measurement, Evaluation and Incentives,* Harvard Business School Press, Boston, Mass.

Otley, D. (1999). Performance management: a framework for management control systems research. *Management Accounting Research,* 10, 363-382.

Otley, D.T. (2002). Measuring performance: The accounting perspective. In A. Neely (ed), In *Business performance measurement, Theory and practice*, Cambridge University Press: UK, p. 3-21.

Ott, J.S. (1989). *The organizational cultural perspective*. KY Dorsey Press: Florence.

Patton, M.D. (1990). *Qualitative Evaluation and Research Methods*, 2nd Edition, Sage Publications: California.

Stringer, C. P. (2006). *Performance management: An empirical study.* Dunedin, New Zealand: University of Otago, 465p. Unpublished PhD Thesis. Reference Number: 0020060331.

Stringer, C.P. & Pillai, P. (2009). *The use of objective and subjective measures: implications for incentive system design.* Paper to be presented at the Performance Measurement Association Conference, Dunedin, 14-17 April, 2009.

Taylor, M.S., Tracy, K.B., Renard, M.K., Harrison, J.K., Carroll, S.J. (1995). Due Process in Performance Appraisal: A Quasi-experiment in Procedural Justice. *Administrative Science Quarterly*, 40, 495-523.

Tyler T.R., & Biers, R.J. (1990). Beyond formal procedures: The interpersonal context of procedural justice. In J.S. Carroll (Ed.) Applied social psychology and organizational settings (77-98) Hillsdale, N.J., Erlbaum.

	2002	**2001**	**2000**
Outstanding (126%+)	18%	21%	18%
Meets target (75-125%)	80%	77%	79%
Minimum acceptable (25-75%)	2%	2%	3%
Unacceptable (>25)	0%	0%	0%

Source: Sensol's performance evaluation and incentive scheme review, 2002.

Table 1 The percentage of individual ratings per category

Grade	2000	2001	2002
Managers			
A	1.75	1.47	N/A
B	1.55	1.7	N/A
C	1.39	1.38	N/A
D	1.4	1.47	1.44
E	1.51	1.38	1.32
F	1.39	1.27	1.35
F+	1.43	1.46	1.41 (120-170)*
G	1.29	1.23	1.25
H	1.24	1.17	1.2
I	1.10	1.18	1.19
J	1.15	1.2	1.16
K	1.14	1.13	1.115
L	1.08	1.16	1.12
J-L			1.13 (54-150)
Professionals			
P	1.07	1.14	1.145
Q	1.09	1.11	1.155
R	1.06	1.1	1.15
S	1.07	1.1	1.09
T	1	1.06	1.15
U	.99	1.12	1.14
S-U			1.11 (75-135)
Total Average	1.09	1.15	1.14

Note: * The range is in brackets. This table shows the average performance ratings per grade for managers and professionals. Archival data has been provided for 2000 and 2001. The performance evaluation data for 2002 has been collected from reports presented at Sensol's performance evaluation and incentive scheme review, 2002.

Table 2 Average performance evaluation ratings by grade

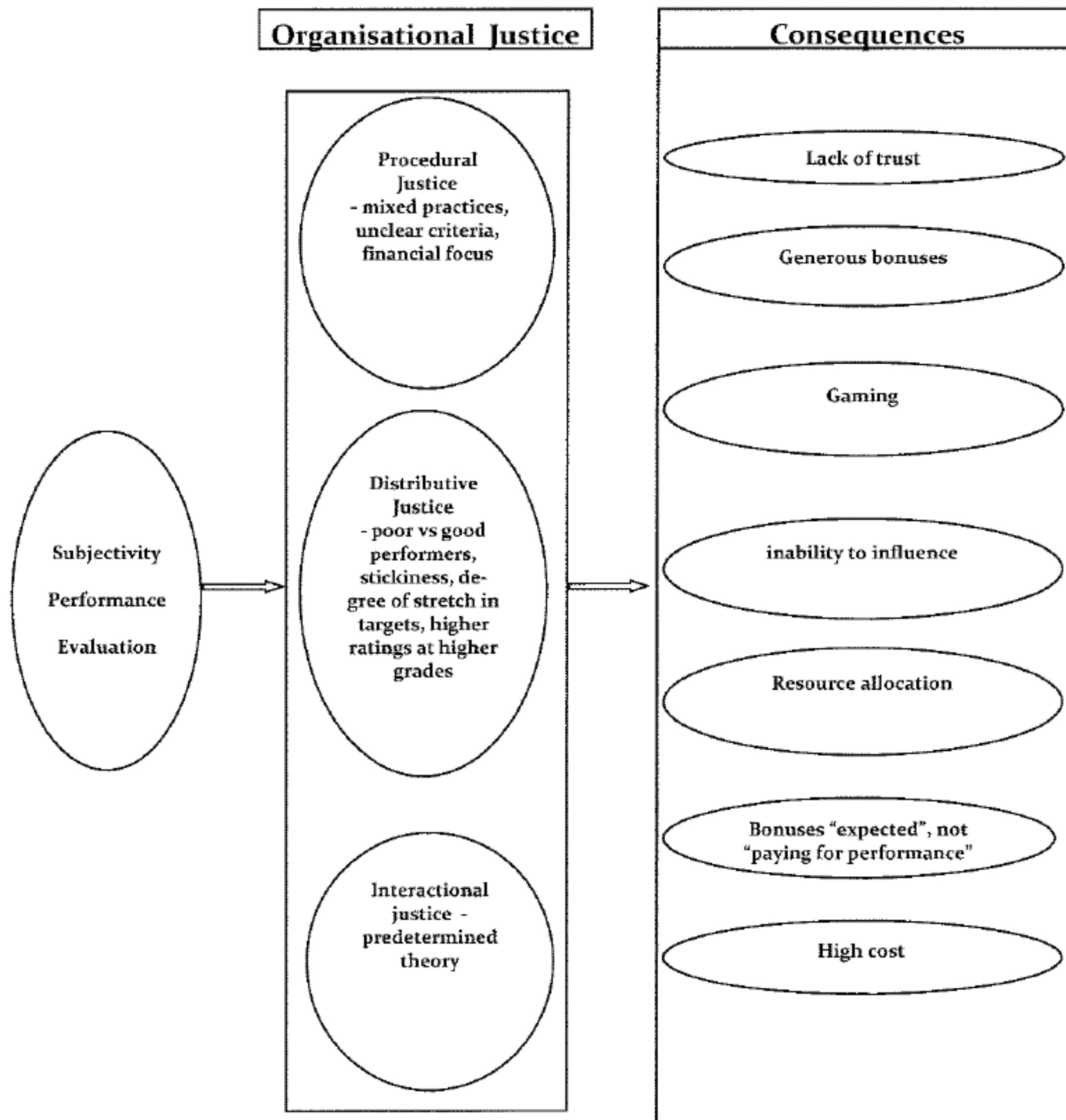
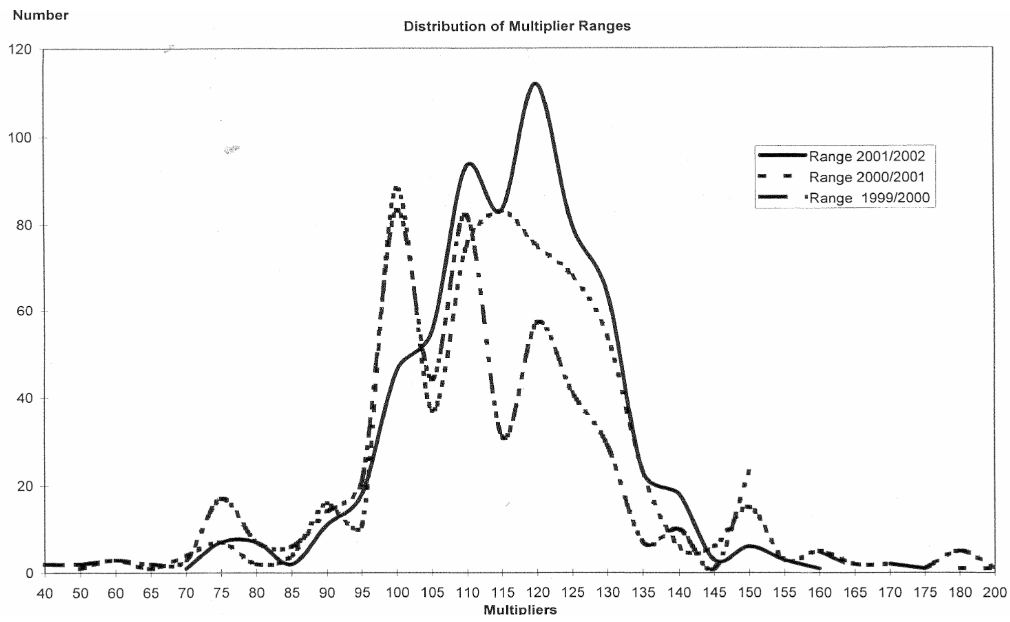


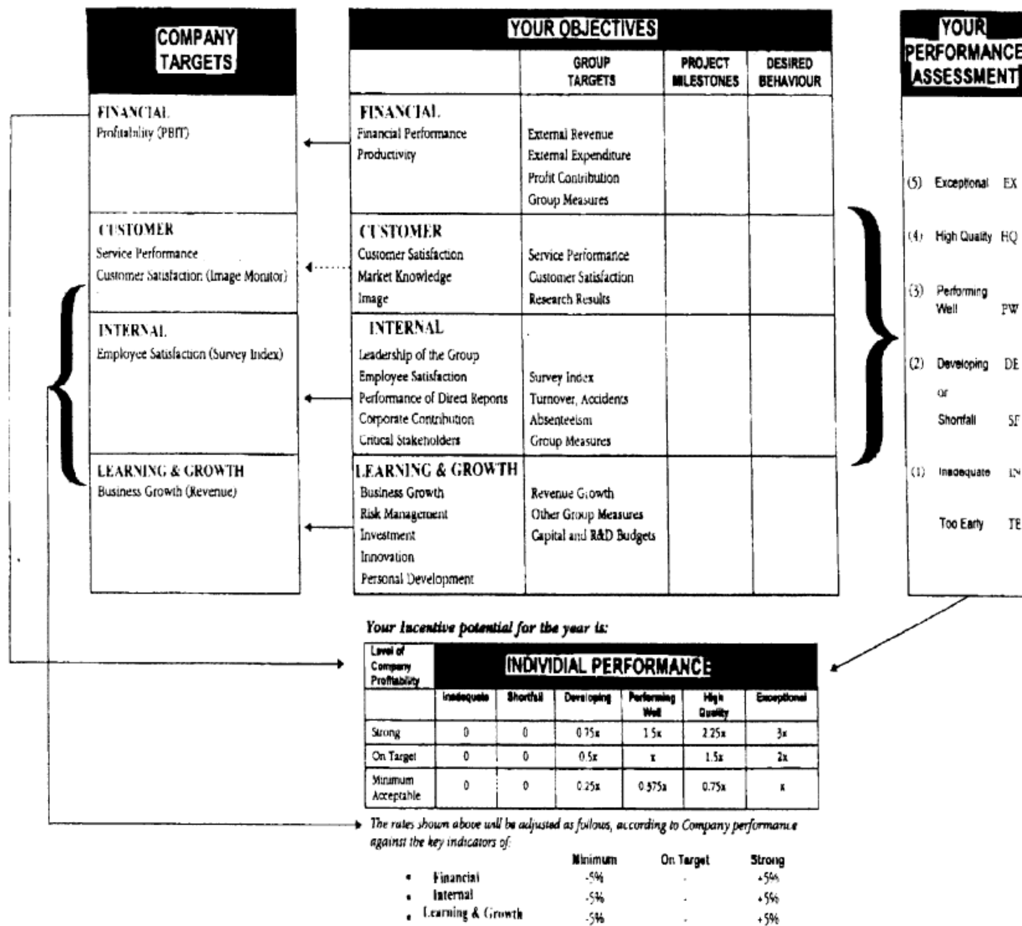
Figure 1 Subjectivity in Performance Evaluations



Source: Sensol's performance evaluation and incentive scheme review, 2002.

Figure 2 Distribution of the annual incentive scheme multipliers

Appendix A Your Performance → Your Bonus (1995)



Your Performance → Your Bonus

COMPANY TARGETS

COMPANY TARGETS
FINANCIAL Profitability (PBIT)
CUSTOMER Service Performance Customer Satisfaction (Image Monitor)
INTERNAL Employee Satisfaction (Survey Index)
LEARNING & GROWTH Business Growth (Revenue)

YOUR OBJECTIVES

	GROUP TARGETS	PROJECT MILESTONES	DESIRED BEHAVIOUR
FINANCIAL Financial Performance Productivity	External Revenue External Expenditure Profit Contribution Group Measures		
CUSTOMER Customer Satisfaction Market Knowledge Image	Service Performance Customer Satisfaction Research Results		
INTERNAL Leadership of the Group Employee Satisfaction Performance of Direct Reports Corporate Contribution Critical Stakeholders	Survey Index Turnover, Accidents Absenteeism Group Measures		
LEARNING & GROWTH Business Growth Risk Management Investment Innovation Personal Development	Revenue Growth Other Group Measures Capital and R&D Budgets		

YOUR PERFORMANCE ASSESSMENT

(5) Exceptional EX
(4) High Quality HQ
(3) Performing Well PW
(2) Developing DE or Shortfall SF
(1) Inadequate IN
Too Early TE

Your Incentive potential for the year is:

Level of Company Profitability	INDIVIDIAL PERFORMANCE					
	Inadequate	Shortfall	Developing	Performing Well	High Quality	Exceptional
Strong	0	0	0.75x	1.5x	2.25x	3x
On Target	0	0	0.5x	x	1.5x	2x
Minimum Acceptable	0	0	0.25x	0.375x	0.75x	x

The rates shown above will be adjusted as follows, according to Company performance against the key indicators of:

	Minimum	On Target	Strong
• Financial	-5%	-	+5%
• Internal	-5%	-	+5%
• Learning & Growth	-5%	-	+5%

Source: Internal performance evaluation document from 1995. Kaplan and Norton's (1996) four generic perspectives have been used for confidentiality reasons.