

# “Is This Correct?” Understanding the motivation for manually acquiring data

Roger Molina

Kerrie Unsworth  
School of Business

Melinda Hodkiewicz  
School of Engineering

Tino Galati  
CEED Client: Water Corporation

## Abstract

*This project seeks to understand the motivational factors that contribute to the quality of data being collected by operators. The importance of knowing the motivational drivers for data collection is that organisations will be able to strategically direct their intervention efforts into improving data collection activities, and hence reduce the need for resource depleting data cleansing. The project does this by focusing on goal hierarchy theory and how goals both at the abstract level and concrete level can affect an individual's performance and intention to collect data accurately. It assesses the individual operators from a water utilities company, focusing on identifying their goals and intentions. This was done with the use of semi-structured interviews and will be further investigated using surveys with field operators within the water services function of the company.*

*Phase 1 of the study has found that there is a positive relationship between mastery approach and motivation to collect data, that there is a positive relationship between managerial pressure and motivation to collect data which is influenced by the contractor identity and further that there is a positive relationship between motivation and collecting high quality data influenced by familiarity with technology, work pressures and peer pressure.*

## 1. Introduction

The overall problem tackled by this project was to identify how to improve the data quality of manually acquired data. Data quality (DQ) is when data is considered “fit for the purpose” (Wang & Strong 1996). However, this definition is not as simple as it seems because the concept of data quality is quite complex (Ballou et al. 2004). In the past it was believed that DQ and data accuracy were interchangeable terms, however in more recent papers it is argued that DQ comprises of many more dimensions than just accuracy (Wang & Strong 1996; Lin et al 2006). Most academics agree that the four main indicators of DQ are accuracy, timeliness, completeness and consistency of data (Lin et al 2006).

Strategic asset management (SAM) is tasked with the job of ensuring that the business has infrastructure assets available to effectively and efficiently deliver the required levels of service. Therefore DQ, for SAM, is the ability to utilise data in a way to extrapolate the life expectancy of assets, the expected performance of assets, and the maintenance of assets over

their life time to ensure that they are able to deliver the required level of service. If the data are not collected or are not of sufficient quality the overall result is the inability for asset management to properly plan, execute and evaluate projects to ensure that corporate goals are met.

The financial implications of the problem of ensuring DQ are immense, particularly in the studied organisation, because the Water Corporation (WC) has \$17 billion dollars worth of assets underground. Furthermore, the industry in which the WC operates in is highly regulated, adding to the pressure of ensuring that data are of good quality.

Past literature on DQ suggests that the main solution to the problem is data cleansing. Although data cleansing is recognised as being an essential component of ensuring DQ, it is a time consuming and costly process (Lee 2004). Even with data cleansing in place it is evident that data quality is still a problem at asset intensive organisations, costing up to 5% of the annual turnover of an organisation (Wallsgrove 2004). The predominant strategies in place for ensuring DQ take place in the data storage and usage stages and thus there is a lack of strategies that looks at the original source of the data that is being collected, the data operators.

Therefore this research project seeks to understand the motivation behind manually acquiring data by asking “what are the factors that most significantly influence the motivation of operators to collect and feedback high quality data in the water industry?”

## **1.1 An Alternative view: Motivation and Goal Hierarchy**

The theoretical framework that this project uses to examine the motivational factors that influence the behaviours of operators to collect data is Goal Hierarchy theory (GH). The main overarching assumptions of GH are that goals are used to direct individual behaviour and that this is done by comparing current states to desired states (Cropanzano et al 1993). If any discrepancies occur then individuals modify their behaviours to minimise these discrepancies. Unlike goal setting theory which focuses on only one goal at a time, GH theory suggests that there is a matrix of goals which an individual attempts to fulfil in their lives.

The basis of goal hierarchy theory is that individuals have higher level goals. These higher level goals are abstract and guide the behaviours of individuals throughout their lives (DeShon & Gillespie 2005). High level goals are hard to measure progress against and thus to measure self performance more easily individuals create sub-goals which also get further broken down into sub-goals. Each level of sub-goals becomes more concrete and proximal and therefore easier to measure (Cropanzano et al 1993).

According to GH theory higher level value goals (e.g., wisdom, respect) are broken down into self identities which refer to both the actual roles taken by individuals as well as the roles that they wish to take in order to fulfil their abstract higher level goals (e.g., Water Corporation employee, father; Cropanzano et al 1993). However, this level is still abstract because we hold many identities and these often conflict with each other, therefore it is difficult for individuals to measure their progress towards their higher level goals. Instead, self identities are used as a systematic way of organising more proximal goals (Cropanzano et al 1993). For this reason self identities are broken down into other sub-goals. The next level of GH is referred to as personal projects (DeShon & Gillespie 2005). Personal projects are more concrete than identities, because they are specific goals which individuals allocate resources to complete in order to fulfil their identities (e.g., getting a promotion). Though not

exceptionally proximal, personal projects are more concrete than the higher levels of the hierarchy. The final level of the GH is known as operational goals; these are the most proximal of all the goals in the hierarchy. These operational goals are the most immediately attended to goals and guide the most recently observable behaviour of an individual (Cropanzano et al 1993).

GH can and will be used as the theoretical base in the study to assess and explain the factors that motivate operators to collect quality data. This is because, for example, let us say that an operator for the WC can take one of two contradicting identities: a subordinate WC identity or a contractor identity. The “subordinate identity” is characterised by the perceived power of supervisors to affect the outcomes of an individual in the desired direction of the supervisor or organisation. Individuals who embody strong orientations to this identity are more likely to be motivated to collect high quality data because it will help to fulfil this WC identity. On the other hand, let us imagine that the individual instead has a “contractor identity”. The “contractor identity” is the extent to which operators see themselves as working for a separate contractor who is an entirely different entity to the WC. In this instance, they do not need to fulfil a “WC subordinate identity” and are less likely to behave in ways that align with the WC’s goals (they will instead behave in ways that align with the contracting organisation’s goals). Therefore, unless the contracting organisation has a desired goal of collecting data they are less likely to collect high quality data.

## **2. Process**

Multiple methods were used in this study. The first phase of the research was exploratory and determined the factors that influenced operator motivation to collect quality asset data. Interviews with 20 water service contractors in the Perth North region were conducted. The interview questions were constructed using the critical incident technique, which involves identifying events that have occurred in the past and recording the responses and outcomes that were generated from these events (Flanagan 1954). The second phase of the study tests the hypotheses that emerged from the results of the interviews through a survey of all WC operators across WA. The second phase is ongoing at the time of writing.

## **3. Findings to date**

Sixty-five percent of interviewees indicated that they were unable to identify any time in the past year or so when they or their co-workers had inaccurately or incompletely recorded work order data. A further analysis of these positive responses have shown that over 70% of these operators have a strong fear of coercive action by their supervisors and 80% have indicated that they are highly committed to the organisation, also stating that ensuring data are properly recorded is part of their job. Interestingly a large proportion only made a superficial connection between data and their job, stating that without the work order data they would not be able to finish their jobs and not linking the importance of the DQ beyond just their immediate personal financial goals. Provided that the self reported performance indicates the actual quality of the data collected, this finding is inconsistent with the managerial belief that DQ is low because operators do not understand the follow on effects of data throughout the rest of the organisation. Only a small minority indicated that doing so lead to the meeting of customer service KPIs or because they genuinely wanted to excel in their work.

Preliminary analysis of the information gathered in this phase indicates that the reason for interviewees reporting high levels of DQ can be linked to personal goals held by the individual operators. The data indicates that the operators who reported high levels of DQ

feared the disciplinary action of their supervisors because they tended to have personal goals linked to security and safety. One interviewee put it this way, "I do my job well, if I am good by them (supervisors), they leave me alone, I'm good."

The operators who stated that they regularly collected high quality data could be classified into two distinct groups: those who had worked for the corporation previously before becoming contractors and those who had only been working with the organisation for 6 months or less. The contractors who had previously worked for the WC had implicit knowledge of the importance of DQ which was gained through their experience, whilst those operators who had worked for a shorter period of time collected data to improve their personal performance. One interviewee said, "I want to be cross skilled and become an expert in my area, so I collect job order data so I can work better." This mastery orientation was shared with many of the younger contractors who reported collecting accurate data.

Conversely, those who had reported inaccurate data could be grouped into three broad groups: those who felt it was in their control, those who felt it was in the control of others or those who felt it was in the control of external factors. Despite the finding that those who felt that choosing to collect inaccurate and incomplete data was within their control were few, this group was the most alarming. A member of this category put the reason behind his neglect to collect meter reading data bluntly, "I do not collect meter readings all the time, it is not part of my job, the WC has their own employees doing that...why should I help them out?" Not surprisingly the individuals who fell into this group tended not to have ever worked for the corporation before and showed goals geared towards autonomy, with very few goals related towards skill mastery or even stability in their job as a contractor.

The largest group of operators who reported not to collect data accurately felt ensuring data quality was within the control of other individuals. This group could then further be divided into two smaller groups: those who felt data quality was the responsibility of the "system" and those who felt it was the responsibility of their co-workers.

The operators who felt the "system" was to blame for the inaccuracies in data collection were characterised by their cynicism towards the authority of the WC and the efficiency of the Mobile Computing System (MCS) which was used to feedback the data. These operators felt that the improvements in technology actually lead to a reduction in their performance to ensure DQ. They felt that the system was restrictive and slow with one operator stating that "Scrolling through the PDA (Personal Digital Assistant) options takes longer than actually writing it down...sometimes some parts are not even in the drop down menu." Other members of this group stated that despite writing down accurate data the data storage system was corrupt and that was what was making the end user data inaccurate. This view was explained by one operator saying that he did not believe that his co-worker forgot to close off on a job, and it was the systems fault that the job was reallocated to him. Conversely, there was a larger number of operators who felt that the system made ensuing DQ easier. One operator pointed out that "the new system still needs getting used to, but it is the future and I can see why we need to use it." Therefore it is not surprising that the "anti-system" group of operators consists of operators who felt that training in regards to using the PDA was inadequate. Interestingly these individuals tended to be older operators, which could be indicative of a difference in technological adaption of older individuals. With this stated, the majority of the anti-system operators actually hold similar abstract goals to those operators who collected data accurately. This indicates that if not for a lack of knowledge about emerging technology most "anti-system" operators would collect accurate data.

The operators who felt that their choice to collect accurate data relied on co-workers' approval are characterised as having a high dependence on their co-workers, and tend to have personal goals leaning towards collective values and a fear of conflict. These operators in most cases hold similar goals to those who collect accurate data regularly, but with the addition of goals that indicate a need to affiliate or be part of a group. Operators in this group reported that they do not collect accurate data when told not to by co-workers, especially co-workers whom they consider to be their senior. One operator gave the example of reporting faulty pipes, he stated that because he was once in a position where he had to complete a job and the planning was "dodgy" due to incomplete data being reported that if he come across a fault once he had his own truck he would report it. However, when asked if he currently reports faults the same operator said "Nah, (be)cause I do not have my own vehicle yet I do not report faults, I do what the older guys say, they are smarter and have been doing this a long time...you know I do not want to start anything...if they say not to report it then I don't." This strong co-worker attitude was common amongst those who worked in teams and on trucks, with one newer contractor even stating that he "didn't like being paired with some drivers...the drivers really makes a difference, you learn a lot from some of the drivers." This shows the strong influence that senior operators have on the ability of other operators to collect accurate data.

The final group of operators who reported to have been unable to collect accurate information believed that it was external factors that influenced their ability to do so. One operator said that it was due to a "slip of the mind," elaborating on this he went on to say that some jobs require a lot of physical and mental work and there are times when such pressures prevent operators from collecting accurate information. This was also the main reason behind the 90% of operators interviewed who openly stated that they have regularly inaccurately recorded job safety assessment data (JSA). One operator gave the following example "You get to site and a pipe is overflowing, the customer is stressed and water is going everywhere...there is no time to do the JSA." It is evident from these interviews that although operators believe that work order data are reported accurately in most cases, additional data are not.

From this initial phase the following preliminary hypothesis can be tested using the survey phase of the study:

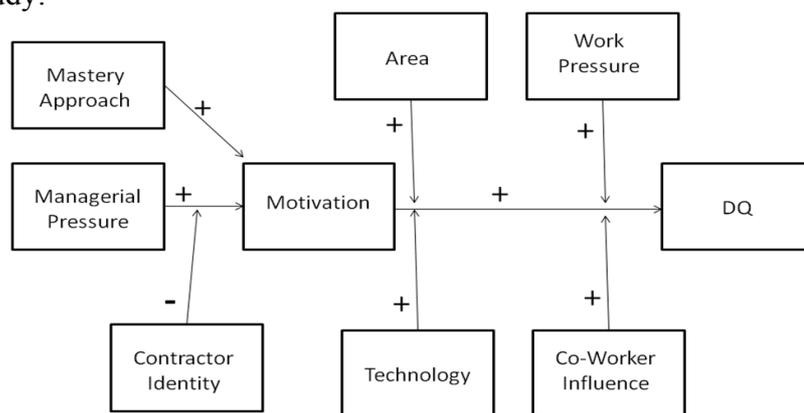


Figure 1. Hypothesised relationships to be tested

- H1: There is a positive relationship between mastery approach and motivation to collect data
- H2: There is a positive relationship between managerial pressure and motivation to collect data which is influenced by the contractor identity
- H3: There is a positive relationship between motivation and collecting high quality data influenced by:
  - a) Technology in a positive manner
  - b) Area in a positive manner

- c) Work Pressures in a negative manner
- d) Co-worker in a positive manner

## 4. Conclusions and Future Work

So far the study has been successful in identifying possible factors that lead to operator intentions to collect high quality data, in particular work order data and JSAs. However, the findings of this study must still be statistically tested in order to validate the relationships found in the interview process. This is planned to be completed in phase 2 of this study. Once these relationships have been validated, it would then be possible to formulate strategic interventions to tackle the DQ issue at the WC. The findings of this study could also be further enhanced in future studies which could look to measure the relationship of the above mentioned factors to the actual DQ of data collected.

With this stated given the possible relationships so far, possible solutions to the DQ problem include; the regular auditing and monitoring of collected data, coupled with disciplinary action when data is not of the standard required. This could be further enhanced with a contractual element that objectively states the level of DQ to be collected by each operator. Another possible solution is regular training sessions to ensure the correct use of the PDA, and procedural policies to standardise the collection of data. Included in these training sessions should also be the communication of the importance of such data to the corporation as a whole as well as to the individual operator beyond just the usual “to get the job signed off.”

## 6. References

Ballou, D., Mandrick, S. & Wang, R. (2004) “Special Section: Assuring Information Quality.” *Journal of Management Information Systems*, 20, pp.9-11

Cropanzano, R., James, K., & Citeria, M. (1993) A Goal Hierarchy of Personality, Motivation and Leadership. *Research in Organisational Behaviour*, 15, pp.267-322.

DeShan, R.P & Gillespie (2005) “A Motivated Action Theory Account of Goal Orientation,” *Journal of Applied Psychology* , Vol. 90, No. 6, pp.1096–1127

Flanagan, J.C (1954) The Critical Incident Technique. *Psychological Bulletin*, 51(4), pp.327-358

Lee, Y (2004) “Crafting Rules: Context reflective data quality problem solving,” *Journal of Management Information Systems / Winter 2003–4*, Vol. 20, No. 3, pp. 93–119.

Lin, S., Gao, J. & Koronios, A. (2006) A Data Quality Framework for Engineering Asset Management. *WCEAM*, 079, pp. 1-10

Wang, R.Y & Strong, D.M (1996) Beyond Accuracy: What Data Means to Data Consumers. *Journal of Management Information Systems*, 12(4), pp.5-33

Wallsgrave, R. (2004) The Real Cost of Asset Information: Better Costs Less. *The Institute of Asset management*, ICOMS, Sydney.