Guidelines for Setting Organizational Policies for Data Quality

Rajiv Dewan and Veda Storey
University of Rochester and Georgia State University

Abstract:

From a process perspective, the tasks that individuals carry out within an organization are linked. These linkages are often documented as process flow diagrams that connect the data inputs and outputs of individuals. In such a connected setting, the differences among individuals in preference for data attributes such as timeliness, accuracy, etc., can cause data quality problems. For example, individuals at the head of a process flow may bear all the cost of capturing high quality data but may not receive all of the benefits although the rest of the organization benefits from their diligence. Consequently, these individuals, in absence of any managerial intervention, may not invest enough in data quality. In this research, solutions to this and similar organization data quality problems are proposed. The solutions focus on principles of reengineering, employee empowerment, decentralization of computing, and mechanisms to measure and reward individuals for their data quality efforts.

Introduction

Data quality research often treats the organization as a monolithic entity with no differences in incentives or preferences. In reality, however, data quality is not valued identically by all users, nor are the value and cost of data quality distributed evenly across the organization. These differences, combined with the interdependence of departments on common business data, create a number of management problems when setting organizational data quality policies.

The objectives of this research are to*: 1) identify and analyze sources of common data quality problems that arise in organizations; 2) suggest policies for organizations to follow when setting data quality standards; and 3) provide solutions for data quality management problems.*

Data quality is a multi-dimensional concept. Ballou and Pazer (1995) identify four dimensions of data quality: accuracy, completeness, consistency, and timeliness. For the purposes of this research, we adopt the definition of Orr (1988) who defines data quality as the measure of agreement between the data views presented by an information system and that same data in the real world. Orr argues that the quality of data should be accurate enough, timely enough, and consistent enough for the organization to survive and make reasonable decisions.

Setting reasonable organizational data quality goals and auditing their implementation is difficult. The major source of this difficulty is the fact that data quality depends on the whole business process and all users of the data are affected. For example, data quality decisions made by the department that is responsible for capturing and maintaining the data limit the quality of the data for other departments. The interdependence of departmental data quality decisions is best examined from a process perspective. A typical business process [Davenport, 1993] begins with a customer that may be an entity inside or outside the organization. Any inaccuracies in data capture at this stage will affect decisions down the line. Although some errors at this stage can be detected and corrected through the use of semantic and referential integrity constraints in databases and applications, many will persist throughout the process and result in incorrect decisions and actions. It then proceeds through a number of business departments that may use, update, or augment the data gathered from previous steps in the process.

If a manager of a particular department in the work decides to take extra care in training its workers and build extra checks and balances into its data capturing process, the whole organization will benefit from these efforts while the manager's department bears all the cost. A manager making a decision on levels of diligence will probably under invest in them from an organizational perspective. This situation is modeled in the next section.

An organization may solve this data quality problem in a number of ways. Solutions range from principles of reengineering [Hammer, 1990], decentralization of information systems to departmental client-server systems, and putting into place a program for measuring and rewarding for data quality [Redman, 1995]. A more conventional solution is illustrated by the *Quest* program of AT&T [Redman, 1995], which involves setting up a procedure to measure and reward data quality. Decentralization of data resources, as implemented in departmental client-server systems also helps in alleviating data quality problems as the department that values the data most is given ownership of the data.

This paper is divided into four sections. In Section two we develop a model for predicting quality choices made by individuals and organizations. This model is applied, in Section three, to develop a set of organizational policies to solve data quality problems. In Section four we discuss the managerial implications of the solutions and conclude the paper.

1 A Model of Organizational Data Quality

This section presents a model to analyze organizational data quality problems. Consider a process flow diagram, such as that shown in Figure 1. The following notation describes the diagram and other parameters:

i index of activities, 1,..., N

P_i set of activities, excluding i, whose quality levels affect the value of activity i. Nominally, it may include all the activities that precede i in the process flow. However, it may be empty or as large as a set of all activities other than i.

q_i quality level chosen by worker performing activity i

$c_i(q_i)$ cost of quality level q_i to the worker

$v_i(q_i, q_j, j \in P_i)$ This is the value to a worker who performs activity i. It depends on i and on the activities in the set P_i.

V Value to the organization.

The assumptions needed for the model are listed and discussed below.

1. *Costly data quality effort*: The cost of quality level q_i of activity i to the worker who performs it (and to the organization), $c_i(q_i)$, is twice differentiable, convex and increasing in q_i. It is natural to expect that higher levels of quality cost more than lower levels. If this were not so, it would be in the worker's self interest to choose higher levels of quality without any intervention from management and the problem would be trivial. Furthermore, it is natural to assume that a unit increase in quality at higher levels of quality costs more than a similar change at lower levels. This could arise for numerous reasons. Suppose a worker has a number of means available to improve data quality. The worker would "cherry-pick the low hanging fruit" first; that is, first implement the quality improvements that are easiest to implement.
2. *Value of Data quality:* The value of an activity to a worker is increasing and concave in the level of quality of all activities that affect it. The concavity assumption captures the fact that the value of decisions may increase with increases in the quality of the data but at a decreasing rate. For example, in determining shipping rates, knowing the region, state, county, zip code and street address would offer increasing precision and quality of decisions but the successive impact of increases in precision on the accuracy of shipping costs decreases relatively.
3. *Increase in data quality by a worker helps others:* The cross partial derivatives of the value of an activity to a worker is positive on activities that affect it, i.e., $\frac{\partial^2 v_i}{\partial q_j \partial q_k} > 0$ for $j, k \in \{i \cup P_i\}, j \neq k$ and zero otherwise. This implies that the quality efforts by workers are complementary, i.e., increases in the quality of one activity makes the quality of others more valuable. To illustrate, consider a simple organization in which worker 1 observes and captures the data and worker 2 processes it. Processing the data more carefully may be worthwhile only if it has been captured accurately in the first place ("garbage in - garbage out.") Conversely, it is worth more to capture data accurately if it is used with greater precision for decision making.
4. *The impact of data quality on value is larger for the organization than for individuals:* The marginal value of quality to the organization is strictly greater than the marginal value of quality to any one worker, i.e., $\frac{\partial V}{\partial q_i} \geq \sum_j \frac{\partial v_j}{\partial q_i}$ for all i.
   This can occur in a variety of ways. One common scenario is that the quality of work

done in a task in a process affects other processes within the company. Consequently, the marginal value for the organization that cares about the value of all work flows is greater than that of the worker performing the task at that one station.

With these assumptions in place, consider first a situation where each worker selects the quality that is best for him or her. Since the value of quality to the worker depends on the quality level chosen by others, the worker has to consider another's decision while making his or her own. The *Nash Equilibrium*, is described below.

Each worker i, picks quality level q_i^* to:

$$q_i^* \in \arg\max_{q_i} v_i(q_i;\, q_j^* \in P_i) - c_i(q_i)$$

The first order conditions for the above are:

$$\frac{\partial v_i}{\partial q_i} = \frac{\partial c_i}{\partial q_i}$$

Second, consider a situation where the organization can and does select the quality level for each worker.

$$\max_{q_1 \dots q_N} V(q_N)$$

The first order condition for the organization's optimization problem is:

$$\frac{\partial V}{\partial q_i} = \frac{\partial c_i}{\partial q_i} \text{ for all } i.$$

Let $q_1^{**},\dots,q_N^{**}$ be the organization's choice.

Theorem 1: *Individuals under-invest in quality from an organizational perspective, i.e., individual workers voluntarily select a quality level lower than the level optimal for the organization.*

The proof is straightforward and is omitted here. This theorem highlights the quality management problem that exists in organizations. If managers do not intervene, then the workers, left to select the quality level themselves, will not choose one that is the best for the organization. This kind of situation is common whenever the actions of an individual create a benefit for others. Such actions are said to create a *positive externality*. The individual making his or her own decision ignores the value created for others. As a result, the individual compares only the marginal cost of an additional unit of quality to the marginal value to the individual alone, rather than for the individual and the rest of the organization. It is this self-interested behavior that leads to a less than optimal solution for the organization. Many similar situations occur, such as the "Tragedy of the Commons" [Milgrom and Roberts, 1992].

3 Organizational Policies to solve Data Quality Problems

Changes in information technology, especially, development of group-ware for collaboration among teams (both local and virtual) such as Lotus Notes; enterprise wide integrated data warehouses and planning systems such as SAP, and other personal computer systems provide a number of solutions to the problems of data quality in organizations. We identify five approaches to solving the organizational data quality problems:

Reengineering Based:

1. *End user computing*: the decision maker processes the information. This is related to *employee empowerment* in which the worker at the head of the process, closest to the customer, makes decisions.
2. Teams: use multi-skilled teams with richer interaction to perform the process jointly.

Organization Theory Based:

3. Setup a program for data quality measurement and incentives for key workers in the process.
4. Change data ownership.

These solutions depend as much on organizational theory as they do on technology. Problems with under-investment by individuals have been addressed by setting up incentive schemes [Jensen, 1983] or changing ownership rights to projects [Coase, 1960, Brickley, Smith and Zimmerman, 1996 and Milgrom and Roberts 1992]. Relevant socio-behavioral theories include enhanced participation in decision making, hierarchy of authority [Deshpande 1982], top management involvement [Bentley] and [Halloran et. al., 1978], establishing leadership in data quality [McGee], [Wang & Kon, 1993], and others [Wang, Storey, and Firth, 1995].

Each of the solutions is analyzed next.

3.1 End-user Computing and Employee Empowerment

End-user computing relies on technology for integration, ease of computing, etc. and it solves economic problems by allocating decision rights to the person to whom it matters. It is commonly used for reengineering business processes**.** Historically, business processes have wended their way through many departments, each of which collect, modify and update data until it reaches a decision maker who uses the data. Long data flows describing such processes are common. Quality decisions, made earlier by workers and managers who may not themselves use some of the data collected, impact the decisions made at the end of the flow. This causes a number of problems related to data quality.

First, a long data flow that separates data capturing and processing from the final decision making also puts the workers at the early parts of the process at an information disadvantage. Since they do not make the decision, they do not know the value of data qualities such as accuracy and timelines. This can, obviously, be a source of data quality problems.

Consider a data entry clerk who is entering data into an electronic form while talking to a customer. Eventually, a decision may be made that depends upon this data. The worker capturing the data may not know the importance of the quality of different data elements in different cases. For instance, the model year of a car may be critical for diagnosing a certain kind of failure but not as relevant for others. This kind of knowledge is useful in motivating additional effort for data quality, but hard to judge by any one but the decision maker.

Second, worker motivation can be a problem when the quality investment has to be made by one worker and the benefits, from better decision making, accrue to someone else. In such situations, the organization has to craft an incentive mechanism that measures and accordingly rewards workers; for example, AT&T's *Quest* [Redman, 1995].

One solution to these data quality problems is *task consolidation* in which the initial tasks of data capturing and processing are consolidated with decision making. Indeed, this is one of the principles proposed by Hammer [1990]: "*Have those who use the output of the process perform the process*." Task consolidation also offers many benefits such as hand off and cycle time reductions [Dewan, Seidmann, Walter, 1997].

The value of end-user computing, that is, having the decision maker perform the data processing tasks, is analyzed in the example below, based upon the model presented earlier.

Example 1. Consider a simple workflow in which a worker 1 in the Inventory Control department counts and enters the inventory levels, which is used for materials planning by worker 2 in the Manufacturing department. Let the value of data quality to worker 1 be $v_1 = q_1(a_1 - b_1 q_1)$ where q_1 is the level of quality picked by worker 1, $0 \le q_1 \le 1$, and $a_1 > 2b_1 > 0$. It is easily verified that v_1 is increasing and concave over the range. The cost of achieving the quality level is $c_1 q_1^2$, for $c_1 > a/2 - b_1$.

Worker 2 decides on the quantity to order for each part using the inventory data collected by worker 1. The value of purchase planning is affected by the quality of inventory data capture, q_1, and the quality of materials planning, q_2.

$$v_2 = q_2 (a_2 - b_2 q_2) + \varepsilon\, q_1 (b_2 q_2 - q_1)$$

where ε is small enough so that $\varepsilon < \frac{a_2}{2} - b_2$; $\varepsilon < 4/b_2$.

The cost of data quality is $c_2 q_2^2$. This is easily verified by v_2 being concave and increasing in q_2.

Assume that the organization accrues the total value of $v_1 + v_2$ and the total cost of the effort. First, consider the case when each worker individually picks the data quality. Workers choose a quality level that maximizes the net benefit to them. It is easily verified, for the parameter ranges listed above, that the solution is interior and first order conditions are necessary and sufficient. These are:

$$a_1 - 2b_1 q_1 = 2c_1 q_1$$
$$a_2 - 2b_2 q_2 + \varepsilon\, b_2 q_1 = 2\, c_2 q_2$$

Solving these simultaneously, we get the Nash solution:

$$q_1^n = \frac{a_1}{2(b_1 + c_1)},$$

$$q_2^n = \frac{a_2}{2(b_2 + c_2)} + \varepsilon \frac{a_1 b_2}{4(b_1 + c_1)(b_2 + c_2)}$$

Obviously, by applying Theorem 1, the organization would prefer that individual workers did not set the quality level.

Corollary 1: *Using* v_1, v_2, V*, and costs defined above, the workers individually will under-invest in quality from an organizational perspective.*

It is easily verified that the conditions in Theorem 1 hold when $\varepsilon < \frac{4}{b_2}$. This is explored further by determining the quality levels the organization would select:

$$q_1^0 = \frac{2a_1(b_2 + c_2) + a_2 b_2 \varepsilon}{4(b_1 + c_1)(b_2 + c_2) + 4(b_2 + c_2)\varepsilon - b_2^2 \varepsilon^2}$$

$$q_2^0 = \frac{2a_2(b_1 + c_1 + \varepsilon) + a_1 b_2 \varepsilon}{4(b_1 + c_1)(b_2 + c_2) + 4(b_2 + c_2)\varepsilon - b_2^2 \varepsilon^2}$$

Consider an alternative process organization in which worker 2 picks the quality levels and performing both of the activities. To obtain the solution, assume that worker 2 receives a total value of $v_1 + v_2$. The worker needs to spend q_1^2 and q_2^2 hours respectively for quality q_1 and q_2 on inventory audit and planning, respectively. The cost to the worker (and the organization) is $c_2 (q_1^2 + q_2^2)$.

In Figures 1 and 2, the quality levels picked by workers 1 and 2 are plotted against increasing interaction, ε. We have set $a_1 = 3$, $b_1 = 1$, $c_1 = 8$ and $a_2 = 10$, $b_2 = 8$, $c_2 = 9$.

In Figure 2 the quality of inventory audit picked by worker 1 is lower than that preferred by the organization. In fact, while the organization prefers to have an increase in quality of inventory audit as the interaction level increases, the worker continues to pick the same low level. If we follow the principle above and have worker 2 performs inventory audit and planning, then the quality level of inventory audit selected by this worker more closely tracks the organization's desired quality level and increases as the interaction level increases.
In Figure 3 the quality level of planning picked by worker 2 is lower than that preferred by the organization. Note that this is the case even though the value of inventory audit is not affected by the quality of planning. When worker 2 performs the both the tasks, we see that he picks a higher quality for planning (and for inventory audit, as shown earlier). This is because the higher level of quality of inventory audit makes an increased investment in the quality of planning more worthwhile. From a process flow perspective, worker 2 decides to process the data more carefully when the data capturing is done with higher quality.

This analysis indicates the presence of quality problems for the organization. Having worker 2 perform both of the activities greatly alleviates the problem. This is true even when worker 2 is more skilled and better paid than worker 1 for larger interaction levels. This can be seen in Figure 4 where the net value to the organization from different processes is plotted, for $c_2 > c_1$, i.e., worker 2's time is more valuable than that of worker 1. For large dependence, when ε is larger, the organization would prefer worker 2 do the job of the less skilled worker 1, in addition to his own job. In fact, we find that worker 2 picks a higher quality level for *both* the activities if the worker does the whole process. In the case of $c_1 < c_2$, this preference is true *à fortiori*.

3.2 Use of Teams

Increased use of teams to replace parts of a hierarchical organization is another business reengineering practice. Numerous cases (Hallmark, Inc., Bell South, Inc., and others) have been documented [Hammer & Champy 1993; Davenport 1993] with the use of teams having been called "one of the most powerful enablers of structural change." Teams offer new interactions, often much more involved and numerous, between individuals than does a hierarchical organization. This increased level of interaction and awareness of each other's quality decisions changes the choice made by an individual to one that more closely reflects the choice of the organization. Teams have become the favored method for organizing large projects such as Ford Car Brake design and Pentium Microprocessor design [Eppinger et al., 1994].

The value of teams for improving data quality is illustrated below.

Example 2: This example continues the scenario explored in Example 1, which shows that worker i's self-motivated choice was to pick a level of quality q_i such that:

$$a_i - 2 b_i q_i + \varepsilon b_i \sum_{j \in P_i} q_j = 2 c_i q_i.$$

Assume that individuals are similar and have similar interactions. Let each individual interact, on average, with m other individuals. Invoking symmetry and simultaneously solving the conditions for each worker, we obtain:

$$q^* = \frac{a}{2(b+c) - m\varepsilon b}.$$

The case where the organization picks the quality level is:

$$q^{**} = \frac{a}{2(b+c) - m\varepsilon b - m\varepsilon(b-2)}$$

For $b > 2$, the denominator for q^{**} is smaller by $m\varepsilon(b-2)$, so, for this case, $q^{**} > q^*$. The same relationship would have been obtained by employing Theorem 1.

Corollary 2 *Using the value and cost functions of Example 2, workers under-invest in quality, i.e.,* $q^{**} \geq q^*$.

It is interesting to examine the performance of self-managed workers as the interaction level increases. This is explored in Figures 6 and 7.

The solid line is the benchmark case in which the organization dictates the quality level for all workers, and the value of each worker depends on six others. The benchmark case is represented by a solid line. The dashed line is for a self-managed organization with levels of interaction ranging from 1 (star organization), 2 (straight line or hierarchy of any span), and larger teams. We see that the individual quality levels and organization value increases as teams get larger. The difference at interaction level 6 is the difference between dictated levels and voluntary choice (q^{**} - q^{*}).

3.3 Incentive Schemes

Setting up business procedures and incentive schemes to co-ordinate choices solves data quality problems within organizations. There are three parts to this process: 1) set clear individual and group goals for data quality; 2) set incentives that are tied to success in meeting goals; and 3) build-in mechanisms for measuring and compensating workers based on performance.

3.3.1 Setting Data Quality Goals

First, clear data quality goals must be set. From a process perspective, the choice of data quality goals is dictated by the goal of maximizing net value to the customers of the process. Applying this to the work flow in Figure 1, the key "customers" are the manufacturing department that schedules its activities based on parts availability and the supplier who must be paid for goods and services rendered. The manufacturing department would benefit greatly from an accurate inventory count and reduced cycle time for the parts ordering process. The supplier would like a short cycle time from shipping to payment. Additionally, there could be organizational goals for control and accounting.

Information systems standards, which are a part of the control mechanism for information systems [Dewan, Seidmann and Sundaresan, 1996] can play a key role in formulating and communicating data quality goals. Programming standards (mandated structured walk-throughs, documentation requirements, etc.) are an example of IS standards that help improve the quality of data by reducing errors from software bugs, improving relevance of data, etc. Standards play a role similar to that of budgets in financial management. They set the minimal acceptable level that all participants must meet to qualify for incentives.

3.3.2 Setting Incentives

Differences in costs and benefits from data quality are a source of difficulty in managing data quality within the organization as discussed earlier. Consequently, it is only natural to expect incentives, which alter a decision makers' preferences, to play a role in providing a solution. Group and individual incentives that are tied to clear organizational goals can be very successful. An article on Gordon Bethune, CEO of Continental Airlines, [*The Wall Street Journal* May 15, 1996] chronicled the turnaround that he achieved through incentives. Continental Airlines was at the bottom of the list of on-time performance published monthly by the US Federal Aviation Administration. Presumably, this is an important quality measure of airline performance. Bethune instituted a group incentive plan which give a bonus of $65 to each employee every month the airline was listed in the top half of the on-time list. The airline made the top half of the list in the second month of the program. This incentive not only changed the on time performance, it also brought large changes in employee morale and cooperation.

3.3.3 Measuring Performance

Setting goals and incentives work only if they are credible. Credibility requires a commitment to reward performance. This in turn requires that performance be measured in clear and direct terms to give feedback to employees. Fortunately, information systems are uniquely suited to facilitate measurement. Numerous tools are available to measure human and machine performance. The key is to choose one that makes measurements that are directly connected to data quality goals. For example, if cycle time is an important goal, the information system can be enhanced to use the time stamps on customer orders and bills of lading to regularly print out the average cycle times. An example of such measurement is the *percentage scanned report* that is regularly generated at large retailers to measure the performance of point of sale personnel. A certain amount of skill and diligence is required to reliably scan items at the point of sale terminal. The system keeps track of the percentage of items that the clerk scans instead of typing-in the product code. While some items are hard to scan, scanning an item code, in general, improves the productivity of the clerk and reduces the check out time for the customers. Hence percentage scanned is a quality measure and one that is directly measured by the information system used by the retailers. Clerks who get the highest scanning percentages are recognized and given a bonus. Clerks who are at the bottom of the list are offered additional training.

3.4 Data Ownership Policies

Data ownership polices involve a decentralization of information technology resources, departmental computing, etc. (In some sense it resembles end-user computing, except that it is at the departmental level.)

Coase suggested that ownership or allocation of decision making rights can achieve a similar purpose [Coase, 1960]. The work flow described in Figure 1 can be used to illustrate Coase's approach. First consider a situation in which the inventory control department does not report to the manufacturing department and maintains its data in its own or corporate database system. Provided the manufacturing department does not have direct control and incentives are not used, the inventory department manager will under invest in data quality for data capture from the perspective of the manufacturing manager. This was shown in Figure 2. One solution for the problem is to make the inventory department report to the manager of the manufacturing department who would then set the data quality standards. A similar solution is proposed in [Van Alstyne, Brynjolfsson and Madnick, 1995]. While this approach may seem cumbersome using legacy technology, current client-server technology is uniquely suited to take advantage of the ownership approach to solving data quality problems.

Decentralization of data ownership and management is the key of client server technology. Instead of having a centralized database, the data is decentralized and managed in server systems owned by the departments. Users access the data through clients. In cases where there is a clear beneficiary from the quality of certain data, Coase's approach requires that the department that derives the most benefit from the data quality also owns the data. It is then free to manage the data server and set usage policies that are aligned with its data quality requirements. This completely obviates the externality problem described in sections 2 and 3. In cases where there is not a clear beneficiary of data quality, a combination of ownership policies and incentive schemes may be used to manage the data quality of the organization.

4 Managerial Implications and Conclusion

Data quality characterizes the whole business process rather than just the data present in the databases. Each step in the process, from data capture to processing for decision support, has an impact on the final quality of the data. This creates interdependencies in the organization where the net value that an individual or department receives from data quality depends on the choices of others. This is a source of data quality management problems in an organization. The problems may manifest themselves as under investment in data quality enhancing activities by individuals as they do not see the value they create for others in the organization. Even if managers are diligent in enhancing data quality, the multi-attribute nature of data quality poses additional problems. Managers make tradeoffs between data quality attributes that suit their decision making but not necessarily the organization as a whole.

Solutions to these problems require attention to both the human and machine portions of the information system. Data quality problems and solutions must be considered as early as the design stage of the information system. To the extent that data quality problems can be anticipated, checks and balances such as referential and integrity constraints can be built into database systems. With these constraints in place, all applications that read or update data will transparently receive the benefit of these checks and balances. It is also easier to build in data quality performance measurements into the information system at the design stage. It would be useful to analyze the decisions and determine the impact of data quality on them. Then a measurement system can be built to measure these key data quality attributes.

Data quality considerations should also play a role in the design of client-server systems. Data ownership is a key decision. The department that owns the data also manages the data server and sets policies for data use update. Hence, it is beneficial if the department that derives the most benefit from the data be the one to own and manage it. This reduces the interdependency problem and the remaining issues can be dealt with by the incentive system.

Finally, the human part of the information systems deserves as much attention as the machine part to improve the data quality. Determining key data quality characteristics, setting clear data quality goals and building incentive systems to reward individuals who perform well are essential parts of organization architecture.

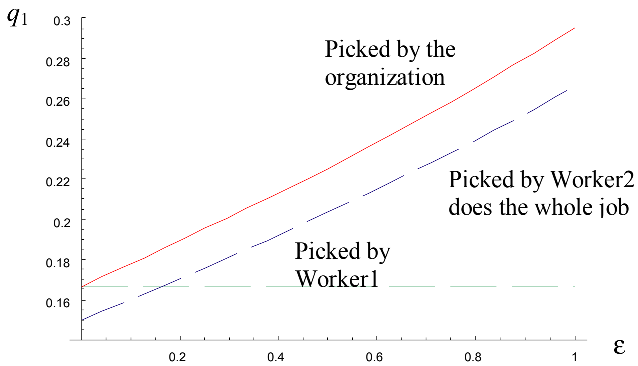

Figure 1: Quality of activity 1 (q_1) Versus Interaction level (ε)

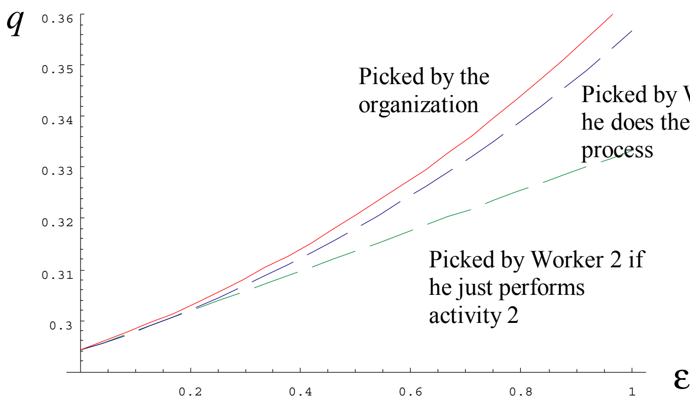

Figure 2: Quality level of activity 2 (q_2) Versus Interaction Level (ε)

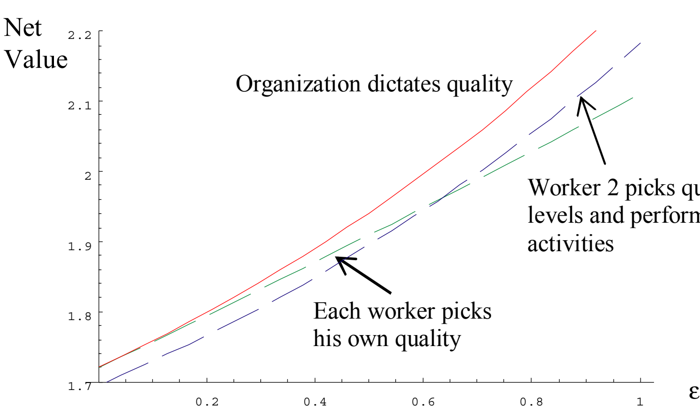

Figure 3: End-User Computing: Organizational Net Value Versus Interaction Level (ε)

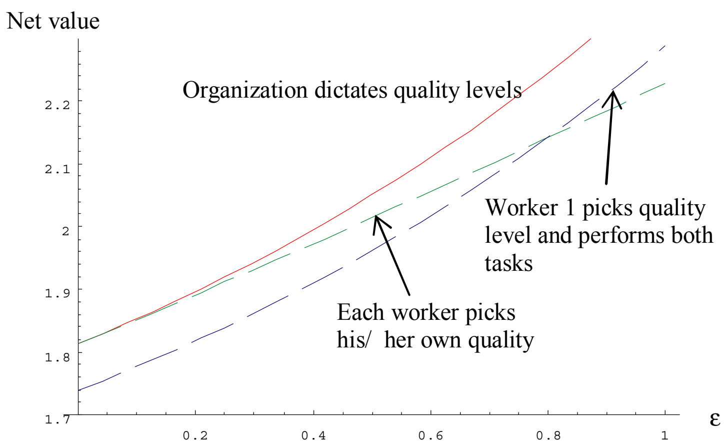

Figure 4: Employee Empowerment: Net Value Versus Interaction Level

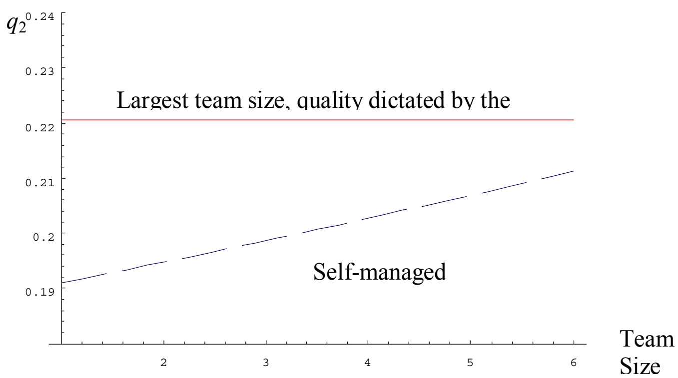

Figure 5: Impact of Team Size on Quality Level

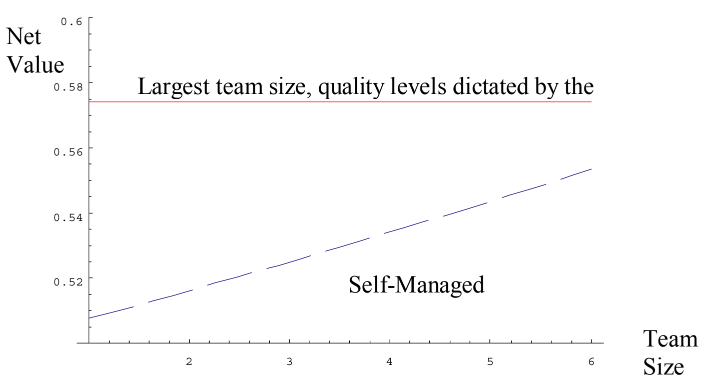

Figure 6: Net Value Versus Team Size

Bibliography

1. Ballou, D.P., and Pazer, H.L., “Modeling data and process quality in multi-input, multi-output information systems”, *Management Science*, Vol.31, No.2, 1985, pp.150-162.
2. Benjamin, R.I. “Information Technology in the 1990s: A Long Range Planning Scenario.” *MIS Quarterly* 6, No.2, June 1982, pp. 11-31.
3. Brickley, J.A., Smith Jr., C.W., and Zimmerman, J.L., *Organizational Architecture,* Irwin, 1996.
4. Coase, R., “The Problem of Social Cost”, *Journal of Law and Economics*, Vol.3, 1960, pp.1-44.
5. Davenport, T.H., *Process Innovation: Reengineering Work Through Information Technology,* Harvard Business School Press, 1993, pp. 5- 9.
6. Deshpande, R. “The Organizational Context of Marketing Research Use.” *Journal of Marketing*, 1982, pp. 91-101.
7. Dewan, R.M., Seidmann, A., and Sundaresan, S., “Organizational Determinants of Corporate Information Systems Standards Policies,” in the *Proceedings of the First Informs Conference on Information Systems and Technology,* Washington, DC, 1996, pp. 233-235.
8. Dewan, R.M., Seidmann, A., and Walter, Z. “Workflow Redesign Through Consolidation of Information Intensive Business Processes.” Forthcoming in *Proceedings of ICIS*, 1997, Atlanta, GA.
9. Eppinger, S.D., Whitney, D.E., Smith, R.P. and Gebala, D.A. “A Model-Based Method for Organizing Tasks in Product Development.” In *Research in Engineering Design*, 1994, pp. 1-13.
10. Gray, P., King, W.R., McLean, E.R. and Watson, H.J. *MOIS: Management of Information Systems*. Dryden Press, 1989.
11. Halloran, D., et al. “Systems Development Quality Control.” *MIS Quarterly*, vol. 2, no. 4, 1978, pp. 1-12.
12. Hammer, M. “Reengineering Work: Don’t Automate, Obliterate.” *Harvard Business Review*, July-August 1990, pp. 104-112.
13. Jensen, M., “Organization Theory and Methodology”, *The Accounting Review*, Vol.58, 1983, pp.319-339.
14. Milgrom, P., and Roberts, J*.,* *Economics, Organization and Management*, Prentice Hall, 1992.
15. McGee, A.M., and Wang, R.Y. “Total Data Quality Management (TDQM): Zero Defect Data Capture.” *The CIO Perspectives Conference*. Tucson, Arizona, 1993.
16. Orr, K. “Data Quality and Systems Theory”, *CACM*, February, 1998, Vol.41, No.2, pp.66-71.
17. Redman, T.C., “Improve Data Quality for Competitive Advantage,” *Sloan Management Review*, Winter 1995, pp.99-107.
18. Takayama, A, *Mathematical Economics,* Dryden Press, 1985.
19. Tayi, G.K., and Ballou, D.P., “Examining Data Quality,” *CACM*, Feb. 1998, Vol.41, No.2, pp.55-57
20. Van Alstyne, M., Brynjolfsson, E., and Madnick, S., “Why Not One Big Database? Principles for Data Ownership”, Decision Support Systems, Vol.15, 1995.
21. Wang, R., and Kon, H.B., “Towards Total Data Quality Management (TDQM),” in *Information Technology in Action: Trends and Perspectives*, Prentice Hall, Englewood Cliffs, NJ, 1993, pp.179-197.
22. Wang, R., Storey, V.C, and Firth, C., “Data Quality Research: A Framework, Survey, and Analysis”, *IEEE Transactions on Knowledge and Data Engineering*, Vol.7, No.4, August 1995, pp.835-842.
23. Wang, S., “OO Modeling of Business Processes,” *Information Systems Management,* Spring 1994, pp. 36-43.