How Important is Data Quality for Evaluating the Impact of EDI on Global Supply Chains?

Bas H.P.J. Vermeer
b.h.p.j.vermeer@tm.tue.nl

Abstract

In practice a growing number of companies report negative impacts of using EDI, which are attributed to the impact of poor data quality. This result is in contrast with current EDI theory, which shows that EDI leads to less data entry errors and hence to a positive role of data quality in EDI. Based on the current literature on data quality and communication, this paper develops a theory on the role of data quality in an EDI enabled business process. The theory hypothesizes that EDI has a positive effect on the reliability and relevance of the EDI information. However, EDI has a negative impact on the context of the EDI information. This means that EDI requires a higher degree of context alignment between the sender and receiver of the EDI messages. If this higher degree of context data quality is not established, EDI will lead to a negative impact on process performance, because the poor data quality in the context leads to more errors or more prevention in the ordering process.

1. Introduction

Using Electronic Data Interchange (EDI) in interorganizational business processes promises many positive impacts. EDI results in more *efficient* business processes through reduction in paper handling related activities, less data-entry errors, elimination of the data-entry function, a reduction of throughput time and hence in a reduction of inventory costs [1]. Furthermore, EDI results in more effective business processes because EDI generates new process design options, which lead to more process productivity [2, 3].

Although several studies regarding operational costs and benefits of EDI confirm these positive effects [4, 5, 6, 7, 8] several others do not [9, 10, 11, 12, 13]. For instance, Benjamin et. al. [9] conclude that: "Although the odds are better than the lottery, the prospects of hitting the jackpot with EDI application are still slim."

In several studies, poor data quality (DQ) is mentioned as an important cause for these negative effects of EDI. In a study of the impact of EDI in the European automotive industry, Reekers and Smithson report several negative impacts that could result from poor DQ: "in many cases EDI data insufficiently reflects reality".

Ribbers [14], found that Technische Unie, a large Electrotechnical wholesaler, experienced major EDI implementation problems such as: (1) Problems with translation of article codes between suppliers, (2) Interpretation problems, because a large supplier did not make a distinction between different truck loads, and (3) Data alignment problems, because not all codes of suppliers were known to Technische Unie. These problems are similar to the problems that were reported in a field study in the food sector, which concluded that: "many problems with scanning and EDI emerge from insufficient DQ between the databases of suppliers and retailers" [15]. A case study between a pharmaceutical supplier and wholesaler confirmed that: "many EDI problems resulted from major inconsistencies in the article data of the participants' databases" [16].

In this paper we will investigate the role of DQ in explaining the impact of EDI on interorganizational process performance. In section 2, we will first examine the DQ literature to find possible explanations for the role of DQ in explaining the impact of IT initiatives (such as EDI) on process performance. Since the DQ literature does not offer a sufficient explanation, we extended our search to include the communication literature, because communication theory is also concerned with data quality. The communication view on data quality is described in section 3. In section 4, we will combine the quality views of the data quality and communication literature into a theory that explains the impact of DQ on process performance. In section 5, we will test this theoretical model in a case study, after which conclusions follow.

2. Data Quality literature

Through the years, the concept of DQ has shifted from the reliability of data to the relevance of data. We will discuss both views and then use each view in an attempt to explain the role of DQ in the impact of EDI on a business process.

2.1. Reliability view of data quality

The concept of DQ is first presented in the accounting and auditing community, where it is part of the system of internal control. Internal Control comprises: "the plan of organization and all of the coordinate methods and measures adopted within a business to safeguard its assets, *check the accuracy and reliability of its accounting data*, promote operational efficiency, and encourage adherence to prescribed managerial policies." [17]

In this traditional accounting view, DQ is primarily concerned with the Reliability of the data, which means: the degree of agreement of account entries with what happened in reality [18, 19]. Reliability consists of three dimensions [19]:

- *Accuracy*, which relates to the degree that account entries conform with real events in reality.
- *Completeness*, which relates to the degree that an account balance contains all events that occurred in a certain time period.
- *Timeliness*, which relates to the degree that the time of entry of account entries is in conformance with the real time an event occurred.

The traditional view of DQ is shown in figure 1. In this view, an accounting system processes data. The quality of this data is primarily defined as the degree of conformance with reality.

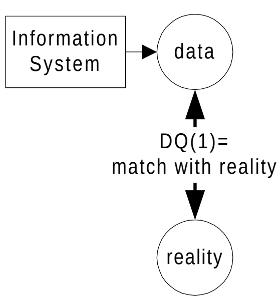

Figure 1. Reliability view of data quality

Same or similar views on DQ can be found in ([20], pp. 851-877 and [21], pp. 75).

2.2. Relevance view of data quality

Starting in the eighties three new models of DQ appeared, which we will discuss below.

2.2.1. Relevance model. In 1980, the Financial Accounting Standards Bord publishes the *hierarchy of accounting qualities* [22]. This hierarchy introduces the *users* of accounting information, who evaluate accounting information on its decision usefulness. Decision usefulness consists of two dimensions [23]:

- *Relevance*, which is defined as the capacity of information to make a difference in a decision by helping users to form predictions about the outcomes in the past, present and future or to confirm or correct prior expectations. (predictive value), which consists of timeliness (is the information on time for a decision?), predictive value and feedback value.
- *Reliability*, which is defined as the quality of information that assures that information is reasonably free from error and bias and faithfully represents what it purports to represent.

Evaluating the relevance model, we see that the concept of relevance is added next to reliability.

2.2.2 'Fitness for use' model. In the beginning of the nineties, a similar approach is adopted in the information systems (IS) community. Several studies in the IS community showed that large organizational problems were the result of poor DQ [24, 25, 26]. To solve these problems, Wang & Strong [27] propose a different approach to DQ. Following the quality literature, they argue to take the consumer viewpoint of 'fitness for use' in conceptualizing the underlying aspects of DQ. In contrast with the reliability view of DQ, the fitness for use approach argues that DQ cannot be assessed independent of the people who use data- the data consumers [28].

However, since there are many users, each with their own definitions of DQ, this results in many dimensions of data [29, 30, 31]. Therefore Wang & Strong developed a conceptual framework for DQ, based on a two-stage survey and sorting study. This framework describes four dimensions for DQ: *Internal DQ* (Believability, Accuracy, Objectivity, Reputation), *Contextual DQ* (Value-added, Relevancy, Timeliness, Completeness, Appropriate amount of data), *Representational DQ* (Interpretability, Ease of Understanding, Representational consistency, Concise representation) and *Accessibility DQ* (accessibility, access security).

The salient part of this study is that data attributes were collected from actual data consumers instead of being defined theoretically or based on researchers' experience.

Evaluating the 'fitness for use' model of DQ we see that in this view the user determines DQ.

2.2.3. Teleological model. The teleological model of DQ is introduced in the nineties in the Dutch accounting literature [32]. According to van der Pijl, two views on the quality of information exist: the causal model and the teleological model of information quality. In the causal model, information quality is the result of a chain of activities, that together influence the quality of information. These activities can be grouped in two phases: the information system development phase and the information system operation phase. The importance of the causal model of information quality is that it is not possible to measure all aspects of the quality of information only from the information itself. The reliability of information is also depending on the measures that are taken in the IS development and operational phase.

In the teleological model the quality of information is determined by the objective for which the information is intended to be used. Van der Pijl argues that information depends on personal objectives which in their turn (partly) depend on organizational objectives. The importance of the teleological model is that it introduces organizational objectives next to personal (e.g. user) objectives in the concept of DQ.

Evaluating van der Pijls' approach we see that the causal model is similar to the reliability concept of DQ, while the teleological model is similar to the relevance of fitness for use concept of DQ.

2.2.3. Evaluation of all three approaches. Evaluating all three approaches, we find that they all include the users' perspective to explain DQ. Since data is used, the quality of the data largely depends on the specific use of the data, apart from the intrinsic quality of the data. This 'relevance' view on DQ is shown in figure 3.

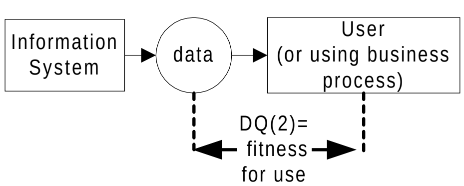

Figure 3. Relevance view of data quality

2.3. Connection with EDI impact theory

Using both the reliability and relevance view of data we will attempt to explain the impact of EDI on a business process. Firstly, if we look at the reliability aspect, we see that EDI leads to less data entry errors in order messages [33, 8]. Hence EDI has a positive effect on the reliability of the order message.

Secondly, if we look at the relevance aspect, we see that EDI leads to a *complete* specification of the information that is needed. This in contrast with manual order handling, where an electrotechnical installer appeared to use an order form where essential information to process the order was omitted [34]. Furthermore, EDI leads to faster communication, and therefore to an improvement of the timeliness dimension. Hence EDI also has a positive effect on the relevance of the order message.

Evaluating DQ on reliability and relevance we conclude that EDI has a positive impact on both the reliability and relevance of the order message, which leads to a positive impact on the interorganizational process performance. Since both DQ views do not explain the experienced negative role of DQ in practice, these views are not suitable.

3. Communication literature

Since data quality theory cannot give a sufficient explanation of the role of DQ in the impact of EDI on an interorganizational business process, we will have to look in other scientific fields. Therefore we chose the communication literature, firstly because communication and information (or data) are closely related. Klooster et. al. [35] discuss several authors who argue both that an information and a communication theory are distinct theories or should be considered as one single theory. Secondly, because EDI is a communication process, and the problems that resulted from poor DQ, were related to EDI. We will give an overview of the communication literature, after which we will show that in this literature a third view on DQ exists. This view will again be connected to the EDI impact theory.

3.1. Objective view of communication

In 1949, Shannon & Weaver introduce a mathematical theory of communication, in which they provide a probabilistic treatment of the transmission of signals through noisy channels [36]. They give a symbolic representation of the communication *system*, which is shown in figure 4.

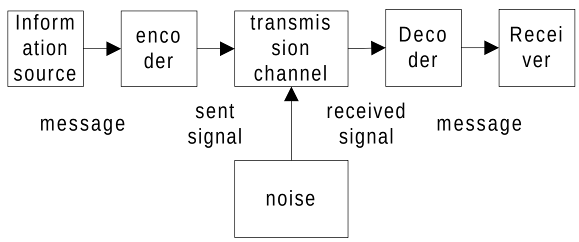

Figure 4. The communication system

In this system, communication is seen as the transmission of messages in the form of signals through a noisy channel. The objective of the communication system is to reconstruct the original message as good as possible at the receiver.

In the eighties, the International Standardization Organization introduces the Open System Interconnection Reference model, (the ISO/OSI model, see for instance [37]). This model provides a layered architecture for electronic interconnection over computer networks. Each layer provides a specific service, independently of the other layers. In the lower layers, the physical connection between computers is established, protocols handle the links between hubs in the network, error detection and correction is provided, and messages are routed through the network. On the higher levels of the architecture, the encoding and decoding of messages using a standard syntax is handled.

The ISO/OSI model is an important extension of the communication theory, because the layered architecture creates transparency for each higher layer. This means that on each layer different protocols can be implemented, without having to worry about the impact of the specific protocol on the other layers.

An important characteristic of both theories is that both the content of the message (what does it mean?) and the objective of the message (What reaction is expected from the sender?) are not really considered. Both theories focus on transmitting the message without errors, assuming that the interpretation of the message is unequivocal, and hence objectively given.

3.2. Intersubjective view of communication

3.2.1. The semiotic framework. In 1992, based on the theory of Signs (Semiotics), Stamper introduces the Semiotics [38, 39]. This framework describes six levels of communication between two subjects (persons or organisations) that operate in a different context. When messages (for instance, EDI messages) are exchanged, agreements on the first five levels are necessary to guarantee successful communication (on the sixth level). A slightly adapted version of the Semiotic Framework is shown in figure 5.

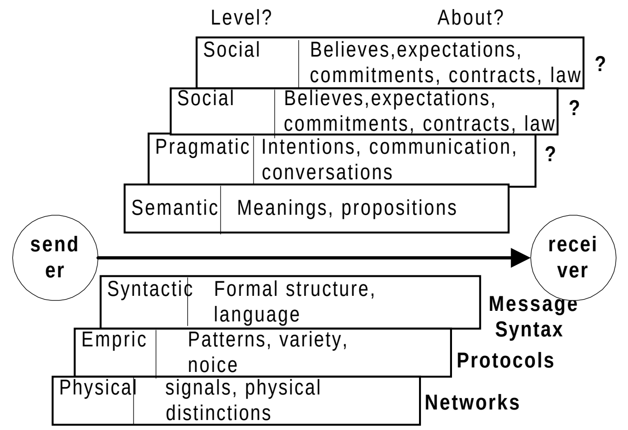

Figure 5. The semiotic framework (adapted from [39])

We will discuss each of the levels shortly. On the *first* level the physical connection is established. For electronic communication, on this level agreements about the type of network connection that is used are established. On the *second* level, the communication of a signal over the physical connection is established. For electronic communication, on this level agreements about the network protocols that are established. On the *third* level, agreements about the syntax of the signal (e.g. the message) are established. Syntactics relate to the structure of a message: which data-elements can be used, and how and in which order are they displayed . The syntactic level can be compared with a grammar dictionary that specifies which words are available in a language and how sentences are constructed. On the *fourth* level, agreements about the semantics of the message are established. Semantics relate to the *meaning* of the data in the message. Here, the relation between data elements with other elements is specified and constraints are defined. This level can be compared with data dictionaries in Database Management Systems, where the conceptual data model is constructed and agreements about for instance the range of product numbers or the definition of turnover are established. On the *fifth* level, agreements about the *pragmatics* of the message are established.

Pragmatics is concerned with the *intention* the sender has with sending the message. In electronic communication, these intentions are normally specified in procedures. Thus, on this level, agreements about procedures are established. Finally, on the *sixth* level, the message should result in a change in the social world, e.g. a change in the believes, expectations, commitments, contracts, laws or culture of which the social world is constructed.

3.2.2. Importance of Stampers' model. The importance of Stampers' communication model is not that he adds three new layers to the ISO/OSI model (the semantic, pragmatic and social level). The importance is that Stamper views information as intersubjectively defined, and not as an objective representation of reality.

To Stamper, information is a sign, which stands for a physical reality, *according to an interpretant* (see figure 6).

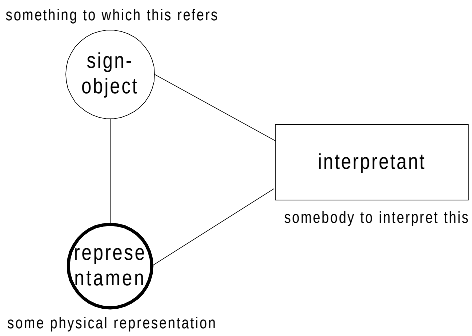

Figure 6. The theory of signs

An interpretant interpretes a sign following the norms and rules of the social group the interpretant belongs to. Examples of such social groups are: a work department, an organization, the company soccer team, a dinner table group, etc. The norms of the social group are implicit rules that describe how things are (semantics) and how things should be done (pragmatics).

When two parties are communicating (which means that they exchange signs), the success of the communication depends on the degree that a shared norm system is established. Only if the communicating parties share the same meanings (semantics) and intentions (pragmatics), a sign can be faultlessly interpreted by each party. Thus, in this view, information (i.e. a sign) is not the representation of an objective reality, but depends directly on the shared norm system (which is intersubjectively defined). Because the shared norm system is established in the context of the communication process, we will refer to the shared norm systems as the context (of the communication).

3.3. Context view of data quality

Since, according to Stamper, information depends directly on the communication context, the quality of information will depend directly on the quality of the communication context. We will define this quality of the communication context as the degree that the contexts of the communicating parties are *aligned*. This new view of DQ is shown in figure 7.

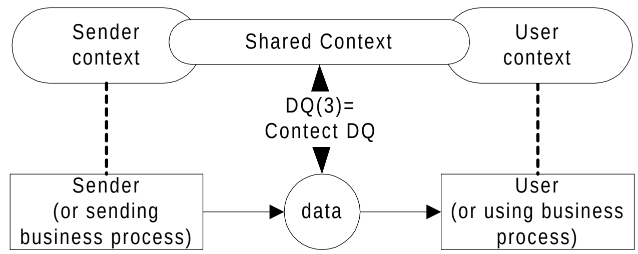

Figure 7. Context [1] view of data quality

As we can see from figure 7, in this model DQ is defined in relation to the shared context. The question is how this view relates to the relevance view on DQ: What is the difference between both views?

In the relevance view, the point of reference is the user (or the using business process). This means that DQ is judged 100%, when the information is exactly what the user wants. The context view recognizes that, even if the information is exactly what the user wants, there may still be problems with using the information because of interpretation problems. Thus, the quality of the information is not 100%.

3.4. Connection with EDI impact theory

Using the context view of DQ, we will again attempt to explain the impact of EDI on a business process.

As defined earlier, the context DQ of an EDI order is defined as the degree of context alignment between the communication parties. This means that we do not look at the order information itself, but at the information where the order information refers to, which is the product information and the identification information of sender and receiver. The degree that product information and

[1] This concept of Context data quality should not be confused with the concept of Contextual data quality of Wang & Strong. Our definition of context data quality focuses on the context of the communication, while Contextual data quality focuses on the context of the user.

identification information between communication parties is aligned, determines the context DQ of the EDI order. Since EDI implementation leads to higher requirements on product DQ (in terms of aligning definitions between sender and receiver and sending updates) this leads to a negative impact of EDI on context DQ, which leads to a negative impact on process performance through preventing and/or correcting errors in EDI orders.

Thus, with context DQ we have found a concept that can explain the negative impact of EDI in a business process.

4. A theory on the role of data quality in EDI

Based on the results of our review of the DQ literature and the communication literature, we developed a theory about the role of DQ on the impact of EDI on process performance. This theory is shown in figure 8.

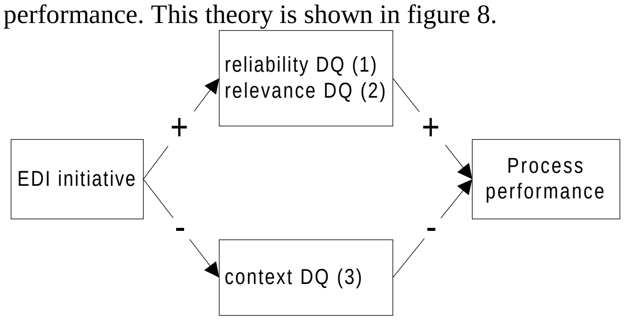

Figure 8. Theory on the role of data quality on the impact of EDI on process performance

Figure 8 shows that we hypothesize that EDI has a positive impact on the reliability and relevance aspect of DQ (of the EDI order), which leads a positive impact on process performance. As discussed earlier, this hypothesis has already been extensively validated. However, the bottom of figure 8, shows that we hypothesize that EDI has a negative impact on the context aspect of DQ, because EDI leads to higher DQ requirements for the context of the communication. This means that EDI requires a higher degree of context alignment between the sender and receiver of the EDI messages. If this higher degree of context DQ is not established, EDI will lead to a negative impact on process performance, because the poor DQ in the context leads to more errors or more prevention in the ordering process.

5. Testing the DQI model: The Schwartz case

The case was conducted at E. Schartz B.V, which a Dutch wholesaler that supplies ball bearings and other technical materials to the industrial market. E. Schwartz BV is a typical medium sized company with 40 employees and a yearly turnover of 13 million US dollars. In the Dutch technical materials sector, where EDI is virtually non-existent, E. Schwartz BV successfully implemented EDI with two large suppliers and several customers.

5.1. The DQI method

Based on the DQ impact theory, we developed the DQ Impact (DQI) method, a method to measure the impact of DQ on an EDI enabled business process. In this section we will describe the DQI method.

5.1.1 Measuring context data quality. As described earlier, the context of the EDI order is formed by the information the order refers to, which is the product information and sender/receiver information in the databases of both the sender and receiver. The context DQ of the EDI orders is defined as the degree that the contexts of both sender and receiver are aligned. In the DQI method, we chose the degree of alignment of the product information between the senders' and receivers' database as the measure for context DQ. We excluded the senders' and receivers' information in the case, because this information (such as sending and receiving address) changed very infrequently in this case. Hence we assumed that the quality of this context information is 100%.

We developed two levels of testing: the internal quality of the product data and the dependent DQ of the product data. The first level test measures the internal DQ of the product data, which means that the conformance of the data with the data model is tested.

The second level test compares the data in the database with the source data. This means that data in the receivers' database is compared with the corresponding data in the senders' database.

5.1.2 Measuring EDI impact. To measure the impact of EDI, we chose not to measure the intermediate effect of EDI on the reliability and relevance of the EDI orders, since many studies already provided positive evidence of these effects. Hence, we chose to directly measure the impact of EDI in terms of the difference in performance in the ordering process.

Although many different performance measures exist, several authors agree on three performance measures for an ordering process: (1) The labor time spent on processing an order, (2) the throughput time of an order, and (3) the remaining number of errors in the process (the quality of the process), see [40, 8] In the DQI method, we define the performance of the business process as the labor time spent in the ordering process and the resulting number of errors. Hence, we chose to exclude the throughput time, because in this case, the ordering process is extremely simple, thus resulting in hardly any throughput variations.

5.1.3. The steps in the DQI method. This resulted in a five steps approach for measuring the impact of context DQ on an EDI enabled business process.

- Assess the level 1 (or internal*)* DQ through comparing the data with the internal data model.
- Assess the level 2 (or independent) quality of the data in the article database through comparing all external fields in the article database with the source data.
- Define the different steps in the ordering process.
- Measure the total time spent on labor and the percentage of time spent on checking activities and calculate the impact on process performance through multiplying the total process time spent on labor with the percentage of time spent on checking activities.
- Establish the remaining error rate through measuring the percentage of invoice errors resulting directly from the order cycle.

5.2. Case design

5.2.1. Objective and proposition. In this case we wanted to test the DQ impact theory, and specifically the bottom part of the theory, which explains the impact of context DQ on an EDI enabled business process. Hence, the objective of the case is to demonstrate that there is a negative impact in process performance because of the impact of poor context DQ.

The theory explains that if the context DQ is not sufficient, there will be a negative impact on process performance. The question is how process performance will be influenced. We hypothesize that poor context DQ will lead to either a longer processing time due to more prevention effort, or more undetected errors after the ordering process. This leads to the following proposition: “With comparable DQ levels, the use of EDI results in a shorter processing time. However, the processing time will be less short due to an increase in prevention effort. Otherwise, the number of corrective errors will have increased”. Using the DQI method we should be able to show these differences.

5.2.2. Design. The best way to test the impact of context DQ, would be to measure the DQ at a company and then determine what the process performance would be under varying DQ levels. However, this is rather difficult, since we can not vary the DQ level within one company. Therefore, we chose to examine two nearly identical ordering processes, which differ on only one aspect, namely the use of EDI in the ordering process.

The units of analysis are two mini-ordering cycles between Schwartz and an EDI and non-EDI customer. Therefore, the case design could be characterized as a quasi-experiment, where all factors are kept equal except the use of EDI. This should result in measurable differences in processing time (due to EDI), prevention effort (due to DQ) and number of corrective errors.

We chose to include the customer of Schwartz within the ordering cycle (and hence the case) because the measured DQ is a result of a database comparison between both supplier and customer. Thus, errors on both sides result in the measured DQ.

The ordering process was kept as small as possible, starting at the moment that an order was generated at the customer, till the moment that the order was accepted and stored in the order database of Schwartz. We did this to assure that there is only one contact point at the customer and Schwartz, where the ordering cycle and the database, and hence where DQ influences the order processing time.

5.2.3. Criteria for interpreting findings. These criteria should indicate under which circumstances we are willing to falsify the impact of context DQ on an EDI enabled business process. This would be the case, if the DQI method would not find any evidence of changes in the amount of prevention and/or correction. Therefore, we identified the following criterion: “The DQI method should be able to explain the difference in processing time between the EDI and non EDI ordering process (each having approximately the same DQ), through a substantial relative increase in processing time due to more prevention effort and/or a substantial increase in the number of corrective errors after the order is processed”.

5.3. Results

5.3.1. Selection of the two customers. We selected two customers for the case, a cigarette manufacturer and a shiphandler. The cigarette manufacturer, is the customer who uses EDI to process their orders. They order about

4000 different articles (such as maintenance and spare parts) using a reorderpoint strategy. This means that orders are completely electronically generated. The ship handler is the customer who is not-EDI enabled. They order about 2000 different articles that they supply to ships that enter the harbor and have to reload. Since every purchasing order of the ship handler is preceded by a quotation, all information for placing the order is already available in the ordering process. Therefore, the ship handler has started to implement EDI in its ordering process, which should result in direct orders that are driven by the quotation process.

5.3.2. Assessment of Data Quality. The quality of the data was tested through comparing the article data in the database of E. Schwartz BV with a subset of the article data of the EDI customer (the cigarette manufacturer) and the non-EDI customer (the ship handler). The overall results of the DQ tests on the two quality levels are displayed in Table 1.

Table 1. Context data quality for the ordering process with 2 Schwartz customers

Level	Important characteristics	Schwartz & EDI customer	Schwartz & non-EDI customer
1	Duplicates in key	0,0%	0,1%
2	Not matched	18,0%	4,0%
	Net price differences	5,8%	6,8%
	% of quality problems	23,9%	10,9%
	Context Data Quality	**76,1%**	**89,1%**

For the internal DQ tests (level 1 tests) the two customers provided the article data they normally requested from E. Schwartz BV on floppy disk. Firstly, we tested if duplicates existed in the article numbers for E. Schwartz BV, or the two customers. Table 1 shows that both the EDI customer and the non-EDI customer had (almost) no duplicates in their records. The article database of E. Schwartz BV also had no duplicates (not shown here). Next we checked whether the field contents complied with the field constraints. For E. Schwartz BV and both customers no important deviations were found.

In the level 2 tests, we compared 11 with the corresponding fields in the Schwartz article database. Although sometimes considerable differences were found (especially in field descriptions) we have concentrated on the fields that are indispensable for order processing. These were the article numbers and the net prices. In the net price check we considered differences due to price unit errors and gross price errors (the non-EDI customer did not use a net price, but calculated the net price through applying a discount percentage to the gross price. Table 1 shows that 18% of the article numbers in the database of the EDI customer were not available in the Schwartz database. This was considerably more than the 4% in the database of the non-EDI customer. Finally, 5.8% of the net prices of the EDI customer were different, compared to 6.8% of the net prices of the non-EDI customer.

Table1 shows that in total 23.9% of the shared article numbers with the EDI customer contained quality problems. This means that the DQ with the EDI customer is 76.1%. In a similar way the DQ with the non-EDI customer is 89.1%. Thus, the quality of the product data with the non-EDI customer appeared to be better than with the non-EDI customer.

5.3.3. Assessment of impact on ordering process. To quantify the processing time in the ordering process, we first described the complete ordering processes between E. Schwartz BV and the EDI customer and non-EDI customer respectively.

Next, we assessed the impact on labor time through comparing the differences in processing time and prevention. The total processing time was measured using time registration forms at E. Schwartz BV and through time estimations provided by the customers. The results are shown in Table 2.

Table 2. Positive effect of EDI and negative effect of preventive checking

EDI customer	**EDI**	**checking**	**Total**
Generate order advise and checking	2	26	28
Manual transport from PC to AS/400	3	0	3
Repairing missing order lines in order	0	7	7
Checking & repairing prices	0	11	11
Total time per order line (sec)	5	44	49
Non EDI customer	**non EDI**	**checking**	**Total**
Manual entry quotation in ordersyst.	30	0	30
Answer phone & checking order	1	22	23
Data-entry Sales	22	0	22
Total time per order line (sec)	53	22	75
Effects			
Positive EDI effect	**-64%**		
Negative extra prevention effect		**30%**	
Total effect			**-34%**

Table 2 shows a processing time difference of 26 seconds (=34%) between the EDI enabled and traditional ordering processes. As we can see from table 2, the positive effect of using EDI leads to a decrease of 64% in processing time. However, since extra checking activities are necessary, an increase of 30% is measured in the EDI enabled process.

Finally, we assessed the impact through detecting differences in error correction. To do this, we looked up all invoices of the passed year and analyzed all invoice errors that had occurred. Then we classified these in errors that had surfaced in the ordering process and errors that occurred in other processes, such as the delivery process, or the invoicing process itself. The results of this analysis are shown in table 3.

Table 3. Analysis of remaining errors

EDI customer		
Coupling error customer	0,29%	
Wrong net price	0,21%	
Return of inventory	0,25%	
Coupling error Schwartz	0,14%	
Data-entry error Schwartz	0,07%	
Total % errors in ordering process		**0,97%**
Errors in other processes	0,57%	
Non-EDI customer		
Wrong product ordered	0,05%	
Wrong net price	0,05%	
Wrong delivery address	0,16%	
Data-entry error	0,11%	
Total % errors in ordering process		**0,38%**
Errors in other processes	0,33%	

Typical errors with the EDI enabled customer were: *Coupling errors,* which are the result of article number matching problems between the article database of the customer and E. Schwartz BV. *Wrong net prices,* because of price differences. *Return of inventory errors,* which resulted because the customer stored two equal products under different private product numbers. These errors are especially expensive. *Data entry errors*, when the Sales department of E. Schwartz BV accidentally miscorrected some prices. The typical errors for the non-EDI customer speak for themselves.

Table 3 shows first of all that the number of errors in both ordering processes is lower than 1% of all order lines, which is low (compared to a normal industry average of 2%). Furthermore, the number of errors without EDI is more than twice as low as the number of errors with EDI. The main reason is that net prices, coupling errors at the customer and return of inventory errors are not detected through preventive checking. However, these errors do not occur in the non-EDI process, since in that process the buyer and the seller talk with each other on the phone. We learned from interviews that in this conversation, problems with prices and mismatches in product numbers are detected before they can really occur.

5.3.4. Case conclusions. The comparison of the EDI-customer with the non-EDI customer lead to the following results:

1. The measured DQ with the EDI customer was 76% while the DQ with the non-EDI customer was 89%.
2. The total effect of EDI on processing time was –30% (decrease in processing time). The positive effect of using EDI was –64%. However, the negative effect of extra prevention activities was +34%.
3. The number of errors in the non-EDI process was 0,38% compared with 0,97% in the EDI process.

6. Conclusions

The objective of this paper was to explain the role of DQ in the impact of EDI on a business process. Especially we were interested in embedding the practical experience of problems with EDI because of poor DQ, in a theoretical framework.

The paper demonstrates that if we extend current DQ theory with the concept of context quality from the current communication theory, we are able to explain the negative role of DQ in EDI enabled business processes. The Schwartz case shows that the implementation of EDI leads to a direct positive effect on processing time of 64%. However, the extra prevention activities in the ordering process due to poor context DQ resulted in a negative effect of –34%. Also the number of remaining errors increased, although the effect is marginal. Thus, this case collaborates the bottom part of the theory about the role of DQ in explaining the impact of EDI on a business process.

Despite these findings, one case is still a small basis for collaborating this theory. Hence more extensive research is required to provide firm evidence.

References

[1] Emmelhainz M.A., *EDI: A Total Management Guide*, Van Nostrand Reinhold, USA, 1993

[2] Davenport T.H., *Process Innovation, Reengineering Work through Information Technology*, Harvard Business Press, Boston MA, 1993

[3] N. Venkatraman, “IT-enabled Business Transformation: from Automation to Business Scope Redefinition”, *Sloan Management Review*, Winter 1994, pp. 73-87

[4] S. Kekre and T. Mukhopadyay, “Impact of Electronic Data Interchange Technology on Quality Improvement and Inventory

Reduction Programs: A field Study", *International Journal of Production Economics* (28:3), December 1992, pp. 265-282
[5] D.R. Mackay, "The Impact of EDI on the Components Sector of the Australian Automotive Industry", *Journal of Strategic Information Systems*, 2:3, 1993, pp. 243-263
[6] C.D. Wrighly, R.W. Wagenaar, and R. Clarke, "Electronic Data Interchange in the International Trade: Frameworks for the Strategic Analysis of Ocean Port Communities", *Journal of Strategic Information Systems*, 3:3 1994, pp. 211-234
[7] T. Jelassi and O. Figon, "Competing through EDI at Brun Passot: Achievements in France and Ambitions for the Single European Market", *MIS Quarterly*, (18:4), 1994, pp. 337-352
[8] T. Mukhopadhyay, SW. Kekre and S. Kalathur, "Business Value of Information Technology: a Study of Electronic Data Interchange", *MIS Quarterly*, (19:2), 1995 pp. 137-155
[9] R.I. Benjamin, D.W. De Long and M.S. Scott Morton "Electronic Data Interchange: How Much Competitive Advantage?", *Long Range Planning* (23:1), 1990, pp. 29-40
[10] J.R. Carter, "The Dollars and Sense of Electronic Data Interchange", *Production & Inventory Management Journal*, 31:2, 1990 pp. 22-26
[11] T. McCusker, "How to Get More Value from EDI", *Datamation* (40:9), May 1994, pp. 56-60
[12] T. Riggins and T. Mukhopadhyay, "Interdependent Benefits from Interorganizational Systems: Opportunities for Business Partner Reengineering", *Journal of Management Information Systems*, (11:2), 1994, pp. 37-57
[13] N. Reekers and S. Smithson, "The Role of EDI in Interorganizational Coordination in the European Automotive Industry", *European Journal of Information Systems* (5), 1996, pp. 120-130.
[14] P.M. Ribbers, "Purchasing Through EDI- The Case of Technische Unie in The Netherlands", in: *EDI in Europe*, H. Krcmar, N. Bjorn-Andersen and R. O'Callaghan (eds), John Wiley & Sons, Chichester, 1995
[15] Vermeer B.H.P.J., *Articles are not alike! An investigation of problems due to poor synchronization of article master data in the grocery sector*, Report EUT/BDK/77, Eindhoven, The Netherlands, 1996 (In Dutch)
[16] Vermeer B.H.P.J., *The information distribution center: distribution mechanism for disitributing article master data in the pharmaceutical sector*, Neprofarm, Utrecht, The Netherlands, 1996 (In Dutch)
[17] Committee on Auditing Procedure, American Institute of Accountants, *Internal Control*, American Institute of Certified Public Accountants, New York, 1949.
[18] Starreveld R.W., H.B. de Mare and E.J. Joëls, *Bestuurlijke Informatieverzorging, Deel 1: Algemene Grondslagen,* Samson Bedrijfsinformatie, Alphen aan den Rijn, 1976 (first print; in Dutch)
[19] See [18], third print, 1994
[20] Weber R., *Information Systems Control and Audit*, Prentice Hall, Upper Saddle River, 1999
[21] Cushing B.E., *Accounting Information Systems and Business Organizations,* Addison-Wesley, Reading MA, 1982
[22] Financial Accounting Standards Board (FASB), "a Hierarchy of Accounting Qualities", published in: Statement of Financial Accounting, Concepts No. 2, *Qualitative Characteristics of Accounting Information*, may 1980
[23] A. Chandra, R. Krovi and B. Rajagopolan, "Flow Parameters and Quality in Accounting Information Systems", In: *Proceedings of the 1998 conference on IQ*, Chengular-Smith I, Pipino L.L. (eds) Cambridge MA pp. 194 - 201
[24] K. Laudon, "Data Quality and Due Process in Large Interorganizational Record Systems", *Communications of the ACM*, (29), 1986, pp. 4-11
[25] T. Percy, "My Data, Right or Wrong", *Datamation* (32,11), 1986, pp. 123-128
[26] T.C. Redman, "Improve Data Quality for Competitive Advantage", *Sloan Management Review*, Winter 1995 pp. 99-106
[27] R.Y. Wang and D.M. Strong, "Beyond Data Accuracy: What Data Quality Means to Data Consumers", *Journal of Management Information Systems*, (12:4), 1996, pp. 5-34
[28] D.M. Strong, Y.W. Lee and R.Y. Wang "Data Quality in Context", *Communications of the ACM*, (40,5), 1997, pp. 103-110
[29] R. Zmud, "An Empirical Investigation of the Dimensionality of the Concept of Information", *Decision Sciences* (9), 1978, pp. 187-195
[30] C. Fox, A. Levitin and T. Redman, "The Notion of Data and Its Quality Dimensions", *Information Processing & Management* (30,1), 1993, pp. 9-19
[31] Y. Wand and R. Wang, "Anchoring Data Quality Dimensions in Ontological Foundations", *Communications of the ACM* (39), 1996, pp. 86-95
[32] Van der Pijl, G.J., *Kwaliteit van Informatie, In theorie en Praktijk,* Dissertation, Katholieke Universiteit Brabant, Tilburg, 1993
[33] M.R. Hoogeweegen and R. Wagenaar, "A Method to Assess Expected Net Benefits of EDI Investments", *International Journal of Electronic Commerce* (1:1), 1996, pp. 73-94
[34] MONTFORT L. van, *Business Process Redesign of the ordering process of an electrotechnical installation company*, Graduation Report University of Eindhoven, Eindhoven, 1997
[35] Klooster, A.J. van 't, and J.A.M. Oonincx, *Leerboek Informatiesystemen*, Samson, Alphen aan den Rijn, 1978
[36] Shannon C.E. and W. Weaver, *The Mathematical Theory of Communication,* University of Illinois Press, Urbana, 1949
[37] Tanenbaum A.S., *Computer Networks,* Prentice Hall International, London, 1989
[38] R.K. Stamper, "Signs, Organisations, Norms and and Information Systems", in: Proceedings of the third Australian Conference on Information Systems, University of Wollongong, Australia, 1992
[39] Huang K., *Organizational Aspects of EDI: a Norm-oriented Approach,* Dissertation Enschede, University of Enschede, The Netherlands, 1998
[40] D.M. Strong, "IT process Designs for Improving Information Quality and Reducing Exeption Handling: A Simulation Experiment", *Information & Management,* (31:5), 1997