

**Effects of reflective thinking and professional experience on sourcing decisions with
inaccurate cost information**

Sebastiaan Morssinkhof
University of Twente
s.morssinkhof@utwente.nl

Marc Wouters
University of Twente
m.j.f.wouters@utwente.nl

Luk Warlop
Katholieke Universiteit Leuven
Luk.Warlop@econ.kuleuven.ac.be

July 2006

Effects of reflective thinking and professional experience on sourcing decisions with inaccurate cost information

ABSTRACT

In this manuscript we investigate decisions that are informed by reports on the total cost of ownership (TCO) of sourcing alternatives. TCO aims to provide a monetary quantification of points of difference of such alternatives, and this enables comparisons between non-commensurate differences in alternatives. However, a “perfectly” accurate monetary quantification of all attributes is unrealistic. We consider three ways in which an attribute may be inaccurately included in the TCO information: as a minimum cost, a maximum cost, or not at all. We investigated the weight of such an attribute in sourcing decisions, and how this depends on the type of inaccuracy, reflective thinking and experience.

We conducted experiments with a total of 529 participants. As expected we found that when TCO translation of a target attribute was represented as a minimum cost, reflective thinking increased the weight for that attribute. We found no support for the hypothesis that reflective thinking would result in less weight for a target attribute that was represented as a maximum cost. We also found that professional experience affects the weight of an attribute that was not included in the TCO number. We found no support for the hypothesis that accountability, as a way to stimulate reflective thinking, would increase the weight of attributes that were not included in the TCO number.

These results suggest that TCO information, which is inherently incomplete, should be provided with care. TCO information, especially if the costs are likely to be underestimated, should be provided in a decision context that motives decision makers to reflect on the information provided. Experience seems to affect the use of TCO information, but experienced decision makers may not always more carefully consider the incompleteness of the provided information. Furthermore, accountability, which may be thought of as the natural way to stimulate reflection in practice, may not have the desired effects. Furthermore these results suggest that the introduction of TCO information should be complemented with careful clarification of the precise meaning, completeness, and limitation of this information. Decision makers with more general business experience in particular may need such training to reduce their tendency to pay less attention to attributes that are not included in the TCO information.

1. INTRODUCTION

Accounting supports sourcing decisions by providing data for the economic evaluation of sourcing alternatives. An example would be when a manufacturer of electronic products compares alternative suppliers of standard electrical components such as resistors, transformers, and printed circuit boards (Degraeve et al. 2005). Suppliers may differ in terms of the purchase price per unit, discount structures, lead-times, or quality. Purchase price and discounts are easy to compare, but how could the firm tradeoff purchase price against lead-time or quality? “Total cost of ownership” (TCO) is a term used in accounting for the monetary quantification of all costs associated with acquiring and using a particular choice alternative, such as transaction costs related to purchasing activities (e.g., ordering, freight, quality control) and the costs related to poor quality (e.g., rejection, rework, and warranties) or delivery performance (e.g., inventory, transportation costs) (Carr and Ittner 1992; Ellram 1995). As in activity-based costing, cost drivers can be at various levels, such as unit level (e.g., purchase price, quality control cost when each item must be inspected), batch level (e.g., costs of creating a purchase order, inspecting an order received), supplier sustaining level (e.g., cost of identification and certification of a supplier), and product or part sustaining level (e.g., cost of maintaining technical product information).

The key element of TCO is monetary quantification: some attributes of a choice alternative are initially not expressed in a financial metric. The financial consequences of these attributes have to be computed—a translation of quantitative non-financial or qualitative information into financial information. As a result, all financial information can easily be aggregated into an overall number, on the basis of which alternative offerings may be compared.

The monetary quantification of quantitative, non-financial attributes or qualitative attributes is persuasive and likely to have a significant impact on the decision-making process. A financial metric draws attention, suggests accuracy, is easy to communicate, and makes comparisons of alternatives easy. In one of the few empirical studies explicitly focusing on the persuasiveness of monetary quantification, Kadous et al. (2005) found that it suggests a greater competence of the manager who prepared it and a higher subjective plausibility of a favorable outcome.

The persuasiveness of a financial metric is understandable, because multi-attribute decisions are difficult for human decision makers. Comparing several alternatives and considering attributes that are expressed on different dimensions is a very challenging cognitive task (Bettman et al. 1998; Bonner 1994). Such decisions become more difficult as

the amount of information increases (e.g., more alternatives, more attributes, more difficult tradeoffs) and the conditions become more demanding (e.g., time pressure) (Bettman et al. 1998). Decision makers are flexible: they adjust their decision strategy depending on the decision, the situation, and other factors (Payne et al. 1993). Managerial decision makers are likely to make use of provided accounting information, because it helps to simplify the multi-attribute decision task by reducing the amount of information that must be considered (e.g., diverse information is accumulated into an overall “net present value” or “cost per unit” summary measure per alternative).

However, the monetary quantification of non-financial attributes can also be problematic, at least for some attributes. Accounting information does often not reflect all operational effects of these attributes (Galbraith 1973; Chambers 1996; Chapman 1997; Lillis 2002) and is in that sense inaccurate. However, little is known about decision-making processes that involve the combination of accounting information and other decision-relevant information. How much importance will decision makers attach to available and relevant information that is not reflected in the financial quantification? This setting must be typical for many complex decisions in which accounting information is a key, but not the only consideration, such as sourcing decisions, capital budgeting decisions, or cost-benefit analyses in public administration.

This study focuses on multi-attribute decisions and monetary quantification. As a first contribution, we conceptualize three ways in which monetary information can ignore decision-relevant information. As a second contribution, we empirically investigate how reflective thinking and general business experience affect the weight of the attribute that is inaccurately quantified. Previous empirical studies in accounting have demonstrated that characteristics of accounting information, decision context, and decision makers have effects on decision-making processes. Yet, the interplay of non-financial attributes, monetary quantification of these attributes, and inaccuracy of that quantification has received little attention.¹

¹ For example, it was found that complexity and amount of information provided may impact decision-making strategies (Shields 1980), decision quality (Isselin 1988), and information selection and judgment accuracy (Shields 1983). Groups may process information better compared to individuals (Stocks and Harrell 1995). Financial information tends to carry more weight compared to non-financial information (Reck 2001; Schiff and Hoffman 1996). Recent studies showed that unique measures may receive less weight compared to common measures when comparing the performance of organizational units (Lipe and Salterio 2000). This effect may also depend on the presentation of performance measures in categories (Lipe and Salterio 2002), subjectivity of the performance measurement system (Ittner et al. 2003), accountability (Libby et al. 2004), and on links between measures and the unit’s strategy (Banker et al. 2004). Other studies have investigated how accounting knowledge (Dearman and Shields 2001) and accounting systems (Briers et al. 1997; Cardinaels et al. 2004; Gupta and King 1997) may impact the use of accounting information, in particular how decision makers consider

We expect that reflective thinking, which in a practical situation may be induced by accountability, will stimulate the decision maker to process the accounting information more carefully. This will affect how the decision maker considers the inaccuracy of the monetary quantification, which will affect the weight of the inaccurately quantified attribute. We expect that general business experience also affects the processing of accounting information, but we do not predict a unidirectional effect. To the extent that experience reflects expertise and skill, its effects should mirror those of reflective thinking. Expert decision makers should be more mindful of the inaccuracies in the provided data. But experience may also induce routine and habitual thinking, and may make decision makers less inclined to adjust their decision strategy to the accuracy of the information provided.

The remainder of this paper is structured as follows. In the next section, total cost of ownership is discussed, and various ways in which such information may ignore decision relevant elements. Section 3 is about the effect of reflection on the weight of an attribute that is inaccurately quantified as part of TCO information provided, and section 4 is about the effect of experience on the decision weight. In both sections, hypotheses are developed and the experimental design and results are presented. Concluding remarks are in section 5.

2. INACCURACY OF TCO INFORMATION

We investigate three forms of inaccuracy, which we will discuss in this section. An example of TCO information used in one of the experimental tasks for this study is presented in figure 1. Information is provided about six attributes of three alternative brands of a machine. For five of these attributes (lifetime, inspection time, the components of the purchase price, chemicals consumption, and energy consumption) the calculation of the cost per hour is straightforward. But not so for the attribute ‘uptime’. We consider three ways in which relevant uptime (downtime) information can be ignored in a TCO report.

[Insert figure 1 about here.]

First, the financial translation can be downwardly biased. Part of an attribute’s financial implications is included in the financial number, but more ambiguous implications are not included. For example, downtime of production equipment may lead to costs for repair and extra labor costs, and these costs may be accurately quantifiable. Downtime may also lead to production orders being to late and, as a consequence, disappointed customers.

the inaccuracy of cost information. In sum, the accounting literature has shown that judgment and decision-making processes are central to understanding the impact of accounting information on the outcomes of decision-making processes. However, such processes are not well understood when it comes to multi-attribute decisions and the role of imperfect monetary quantification of information that is initially non-financial.

Such costs may be much more difficult to estimate. If, in this example, the monetary quantification of the attribute downtime only includes repair costs and labor costs but not the hard-to-quantify cost of lateness, then the TCO number is negatively biased. In other words, the TCO number can be seen as a kind of minimum. This type of inaccuracy is investigated in the first experiment, in combination with the effect of reflective thinking.

Second, the financial translation can be upwardly biased. Rather than leaving out the hard-to-quantify elements, the maximum costs that could be incurred by the firm may be estimated and included in the monetary quantification. In the example, all remotely possible costs associated with lateness can be included, together with the costs for repair and extra labor, in the monetary quantification of the attribute downtime. The financial translation of this attribute is upwardly biased. This type of inaccuracy is also investigated in the first experiment, and also in combination with reflective thinking.

Third, attributes that are difficult to estimate may be completely left out of the calculation of TCO numbers. The monetary quantification of such attributes is calculated as zero, and it is left to the decision maker to consider the attribute in non-financial terms next to the TCO information provided. In this example, there may no cost of downtime included in the overall TCO number. That means that the TCO number, which includes the financial impact of all attributes except downtime, must be traded-off against the downtime of each alternative. This kind of inaccuracy is studied in the second experiment, and we investigate the effects of accountability and experience.

3. REFLECTIVE THINKING

This section contains the motivation of hypotheses, experimental design and results of first experiment. Based on these results, the second set of hypotheses and findings will be presented in the next section.

Hypotheses development

Human decision makers often adopt a strategy of least effort. They want to finish decision-making as soon as possible and are not willing or not able to adopt a demanding cognitive strategy such as specified in normative models (Dewey 1933; Simon 1979; Tetlock 1985). Reflective thinking may enhance the quality of decisions (Langer 1978). Reflection may help decision makers to judge the meaning, relevance, and quality of the available information, and such judgments may help decision makers to get more insight in the data, and bring together facts that seem to be incoherent and disconnected (Dewey 1933).

Reflective thinking is often considered to be beneficial by increasing self-knowledge, getting

a better understanding of the information, and making better decisions (Raiffa 1968; Keeney 1977; Langer 1989).²

We hypothesize that reflection will increase the awareness of the inaccuracy of the TCO information provided. Participants who are motivated to reflect will be more conscious of the complexity of the multi-attribute purchasing problem and think and weight the information more carefully. We expect that they will take the inaccurately quantified attribute into consideration and adjust the weight for the inaccuracy. The implication is that providing TCO-numbers as minimum or maximum costs might result in opposite effects of reflection. Reflective decision makers may attach more weight to inaccurate TCO-numbers if these numbers are estimated as minimum costs, while they may attach less weight to inaccurate TCO-numbers if these numbers are estimated as maximum costs. Minimum TCO-numbers are likely to be underestimations, so the relative weight of the inaccurately translated attribute will increase if a decision maker considers that the costs can actually be higher. The opposite will be the case if costs are estimated as maximum TCO-numbers; these numbers are likely to be overestimations, and so the relative weight of the attribute may decrease when the decision maker takes this estimation error into account. The arguments given above lead to the following hypotheses:

H1: Minimum TCO estimation: Reflection will increase the weight of the attribute that is inaccurate included in the TCO numbers.

H2: Maximum TCO estimation: Reflection will decrease the weight of the attribute that is inaccurate included in the TCO numbers.

Research method

Participants were undergraduate students who were familiar with the basic concepts of cost accounting. They were provided with several alternative purchasing options from which they had to select one. These options concerned alternative brands for production equipment. Information was presented on the same attributes for each brand. Some information was financial (i.e., expressed in Euros) and other information was quantitative non-financial. The

² The process through which reflection can improve decision-making may involve steps such as problem recognition, thinking through the problem and generating possible causes and solutions, building hypotheses and testing these, and evaluation and revision of possibilities (Baron 1981). Empirical studies have demonstrated that reflective thinking may improve the quality of decision-making, which has been measured in various ways, for example by less discrepancy between answers and actions (Pryor et al. 1977), greater predicted correctness of answers (Koriat et al. 1980), and higher correspondence between attitudes and behavior (Snyder and Kendzierski 1982).

dependent variable was the choice for a particular alternative. The experiment was run as a 2×2×2 between-subjects design.

The uptime percentage of Brand C (*Uptime C*) was an independent variable at two levels (99% and 96%).

The second independent variable was reflection on the information provided (*Reflection*), which was manipulated at two levels (no-reflection, reflection). Participants in the no-reflection conditions were only asked to choose one of three alternatives. Participants who were encouraged to reflect were asked to answer the following two questions *before* choosing one alternative:

“Please, explain below how you make a choice for one of the three brands and which trade-offs you make.”

“Please, explain below to what extent you are able, on the basis of the information provided above, to make a choice without further analysis.”

TCO information was provided as an estimation of either the *minimum costs* or the *maximum costs*. In the minimum-cost condition, participants learned that TCO-numbers for downtime were difficult to estimate, and they were instructed that costs for downtime could be higher:

“Other downtime costs, for example costs of not reaching delivery times or additional transports, are highly uncertain and are therefore not included in the table below.”

Figure 2 shows the text introducing the experimental task in the first experiment, and figure 1 shows the table that follows after this text (for the condition *Uptime C* = 99% and no-reflection). In the maximum-costs condition, participants were instructed that costs for downtime could be lower than estimated. The experimental task was the same, except for the note placed below the table, which was formulated as “*Maximum downtime cost estimation”. Furthermore, the sentence on the estimation of downtime cost was different. Instead of the last sentence shown in figure 2, it was stated:

“Other downtime costs, for example costs of not reaching delivery times or additional transports, are highly uncertain. The number in the table below is an estimation of the maximum downtime costs that might occur, however these might be lower.”

[Insert figure 2 about here.]

The interaction of the two independent variables *Uptime C* and *Reflection* allows testing H1 and H2, based on a logistic regression³ with the following specification:

$$\text{Choice C} = \beta_1 + \beta_2 \text{Uptime C} + \beta_3 \text{Reflection} + \beta_4 (\text{Uptime C} \times \text{Reflection})$$

If coefficient β_4 is significant and positive, then the effect of *Uptime C* on the dependent variable will be stronger if participants have to reflect compared to when participants do not have to reflect.⁴ A negative coefficient indicates the opposite, a lower weight. Note that the preference for Brand C is expected to decrease as the uptime of Brand C decreases. If the preference for Brand C decreases more strongly in the high accountability condition compared to the no accountability condition, which is indicated by the interaction term *Uptime C* \times *Reflection*, the weight of the attribute uptime is higher with than without reflection.

³ Actually we should use multinomial regression to analyze the results in this experiment, because more than two choices could be made. However, only 3 out of 148 participants choose Brand B. We decided to eliminate these 3 observations for several reasons from further analysis. First, we have two cells with zero frequencies (not one participant in the conditions with 99% Uptime choose Brand B). Cells with zero frequencies do create problems in the calculation of ratios used in (multinomial) logistic regression models (these ratios become infinite if divided by zero). Merging categories to avoid this, would not be a suitable solution, because we would not be able to investigate the impact of interaction variable β_4 . Secondly, choosing Brand B is illogical; attribute values for Brand B are equal or worse compared to attribute values for Brand A or C, and also total costs per hour for Brand B are above total costs per hour for Brand A and C (even in case downtime costs become €0).

After eliminating choices of Brand B, a model with two (i.e., dummy) dependent variables remains. In case of a model with two possible outcomes, formulas for the ratios in the multinomial regression and binary logistic regression methods are equal; results in multinomial models are expressed in terms of a logistic regression model (Borooah 2001).

⁴ In the logistic regression analyses, Uptime was transformed from a percentage into hours of uptime to avoid multi-collinearity problems due to high correlations between Uptime and interaction effects with Uptime. Note that the independent variables in our models (e.g., Uptime B, Experience, and TCO info) are not related, because these are independently manipulated. However, in models with interaction terms (as in our regression models) the evaluation of the interaction term might be undermined if the interaction term (e.g., XZ) is highly correlated with either X, Z or both. In our models, Uptime indicates the percentage time a machine was available for production, and this was always high (e.g. 96% or 99%). The categorical variables (Experience, TCO info) were either 0 or 1. Multiplying these values in interaction terms gives 0, 1 or close to 1, leading to a high risk on multi-collinearity.

This was indeed found. Although it is not possible to test for multi-collinearity in the SPSS Logistic Regression menu, tolerance levels for multi-collinearity statistics can be found by calculating a linear regression model in SPSS (all other output is not relevant). The tolerance statistic is $1 - R^2_x$, where R^2_x equals the variance of each independent variable explained by all of the other independent variables. Tolerance levels below .10 almost certainly indicate a serious collinearity problem, while tolerance levels between .10 and .20 are causes for concern (Menard 1995). The results indeed indicated multi-collinearity problems. Computer algorithms to isolate the relevant standard errors might fail if multi-collinearity with the product term is high (Jaccard 2001).

Although multi-collinearity is easy to detect, there are few acceptable remedies for it, and no really satisfactory solutions. "What to do about it if it is detected is problematic, more art than science" (Menard, 1995, page. 71). A solution for high collinearity among interaction terms and independent terms is the transformation of the continuous independent variable by centering it before forming the multiplicative term (Cronbach 1987), also suggested by (Jaccard et al. 1990). Hence, we transformed the uptime percentages in uptime hours by multiplying the uptime percentage by lifetime hours (e.g., Uptime = 99%, Lifetime = 2,800 hours, Uptime hours = $2,800 \times 99\% = 2,772$ hours). Subsequently, uptime hours were centered as suggested by Cronbach (1987). As a result, the values of the interaction terms 0, 1 or around 3,000; tolerances for the independent variables did not indicate any (serious) risk on multi-collinearity (always above .10 and mostly above .20).

Results and discussion

Table 1 shows the number of participants and their choices in the various conditions of the first experiment.

[Insert table 1 about here.]

Results related to H1 are in figure 3a and in table 2. We expected that, if the TCO information for uptime was provided as an estimation of minimum costs, the impact of uptime on the preference for Brand C would be larger if participants are encouraged to reflect compared to when they only have to choose. Figure 3a shows that, consistent with this hypothesis, the effect of *Uptime C* on *Choice C* was moderated by *Reflection*. The coefficient for the interaction term was significant ($p = .043$).

[Insert table 2 and figure 3 about here.]

Results for H2 are in figure 3b and table 2. This hypothesis is not supported. We expected that, if the TCO information on uptime was provided as an estimation of maximum costs, the impact of uptime on the preference for Brand C would be smaller if participants are encouraged to reflect compared to when they only have to choose. Figure 3b suggests that the weight was the same whether or not they were encouraged to reflect. The coefficient for the interaction term between *Reflection* and *Uptime C* was not significant ($p = .806$).

The difference between results under the minimum versus maximum conditions is relevant to discuss further. It could be that the provision of TCO as a minimum estimation has more impact than a maximum estimation, and that might explain that the effect of reflection is stronger for the minimum condition compared to the maximum condition. Representing costs as a minimum or maximum implies that decision makers have to make a decision under uncertainty; “real” costs might deviate from the TCO numbers provided in the tables. Prospect theory (Kahneman and Tversky 1979) describes how human decision makers attach more weight to negative consequence (losses) than to the equivalent positive consequences (gains). Reflective decision makers considering a minimum cost may realize that costs could also be higher, which will be experienced as a relative loss, and receive a high weight in the decision. On the other hand, a maximum cost representation indicates that “it can only get better”, which will be experienced as a relative gain. Therefore, decision makers could have attached relatively more weight to a cost attribute represented as a minimum compared to costs represented as a maximum.⁵

⁵ A second argument supporting this idea is based on how decision makers tend to deal with ranges of outcomes. Formulating costs as a minimum implies larger range of possible outcomes compared to formulating costs as a maximum. Decision makers tend to give a higher weight to larger ranges (Fischer 1995). For example, the

This first study suggests that reflection matters for the way in which TCO information is used to support purchasing decisions, and also the type of inaccuracy of such TCO information may have an impact. However, the findings are based on experiments pertaining to business decisions taken by students, and so a logical next step is to involve more experienced research participants. This also raises the question of how reflective thinking may be achieved in real-life situations, and in the next experiment we will create conditions of accountability to invoke more careful and reflective decision-making.

4. EXPERIENCE

In this section, we study accountability in relation to TCO numbers from which some attributes are totally excluded (the third type of inaccuracy described above). The effect of the decision makers' experience is considered.

Hypotheses development

Experience of the decision maker may also affect the weight of the inaccurately quantified attribute. Experience in this study refers to general business experience, and not to specific experience with and knowledge of accounting or purchasing tasks. We will study experience in relation to TCO numbers from which some attributes are totally excluded (the third type of inaccuracy described above). We posit two effects of experience, without predicting a directional effect; experience might be helpful, on the other hand it might hinder decision makers to improve in their judgment and decision-making.

On the positive side, experience may result in expertise, defined as a general ability to solve domain specific problems (Alba and Hutchinson 1987). For example, more experienced decision makers appear better able to distinguish relevant from irrelevant information (Shelton 1999). Broad domain knowledge (e.g., knowledge of basic accounting principles that accounting professionals have who are working in specialized accounting areas), might help them to recognize problems and to further investigate relevant information, even if they are not directly familiar to a specific topic (Vera-Muñoz et al. 2001). Experience is especially helpful if tasks become more complex; and general training (such as education in business administration) and experience helps decision makers too outperform less experienced decision makers (Chang et al. 1997). In our study, we therefore expect that experience can help decision makers to assess the quality and completeness of the provided cost information.

downtime cost of Brand C in figure 1 is €1.41. When this is an estimation of the minimum cost, the range is, in principle unlimited: €1.41 to infinity. When this is an estimation of the maximum cost, the range is €1.41 (at most €1.41 and at least €0). Decision makers are therefore likely to attach more weight to the downtime cost in case these costs are represented as a minimum compared to when these costs are represented as a maximum.

More experienced decision makers should better understand that some attributes are not included in these cost numbers, and they might better realize that it is important to make a deliberate trade off between the TCO numbers and the non-included attributes.

However, experience may also have the opposite effect. It may be associated with habit, routine, automaticity, and superficial thinking. Several studies found unfavorable effects of experience. Experienced decision makers may find it difficult to adopt another strategy, and they may unconsciously adopt the same decision strategy over and over again unless something stops them from doing this (Alba and Hutchinson 1987). Also, a task that deviates from the task structure experts are used to might make it difficult for experts to excel (Nelson et al. 1995). In another study (Vera-Muñoz 1998) it was found that information that is highly relevant according to economic theory was ignored by more experienced decision makers, because they were used to analyze information by decision rules that attached less importance to this particular type of information. Marchant et al. (1991) found dysfunctional effects of high-accounting knowledge as a result of inflexibility to change decision strategies. These studies indicate that expert knowledge might result in inflexible decision strategies. For our study, we expect that more experienced decision makers may more or less ignore information not included in TCO numbers.

We posit no directional effect, because of these contradictory empirical indications and because not many studies have specifically investigated the effect of general business experience, making the effect even more unpredictable. Thus, we formally hypothesize:

H3: Experience will affect the weight of the attribute that is not included in the TCO numbers.

H4: Provision of TCO info will affect the weight of the attribute that is not included in the TCO numbers, and this effect will be moderated by experience of decision makers.

In this experiment, we include reflective thinking in a different way compared to the previous experiment. How do companies stimulate people to be reflective? We posit that accountability will play an important role: people will think more about their decisions, if they expect having to justify their decisions. Accountability is the explicit or implicit feeling and pressure decision makers undergo when they have to justify their actions and decisions (Curley et al. 1986; Tetlock 1992). Accountability is important in business, because organizational buyers have to be able to provide analysis and give a more formal justification than consumers (Sherlock 1991). Being able to give sound justifications is also important for managers, because convincing others is an important managerial ability that is close related to

managerial success (Tetlock 1985). Decision makers might want to avoid “loosing face” when they feel they are held accountable, and this may change their decision processes (Carnevale 1985). This may be even more so in threatening situations; if the decision maker’s explanation is monitored by others and mistakes might result in sanctions (Fox and Staw 1979). Therefore, accountability might motivate decision makers to adopt a strategy that will result in profound analysis and evaluations (Tetlock 1983). For example, Doney and Armstrong (1996) found that purchasers who are accountable for their decision-making process analyze information more extensively than purchasers who are accountable for the outcome of decisions.

We will study accountability in relation to TCO numbers from which some attributes are totally excluded (the third type of inaccuracy described above). We expect that accountable decision makers will be more motivated to make a cognitively challenging tradeoff between the TCO numbers and the attribute that is not included in the TCO numbers. Thus we hypothesize:

H5: Accountability will increase the weight of the attribute that is not included in the TCO numbers.

Research method

In this experiment, the dependent variable was again the choice for a particular alternative, for brand B in this case (*Choice B*) because participants were given two alternatives instead of three. A 2×2×2 between-subjects design was used, and *Experience* was added as a measured variable. Some participants (N = 199) were undergraduate students similar to the first experiment, and another group of participants (N = 128) consisted of practitioners who were visitors and exhibitors at a large international maritime trade fair in Rotterdam. The median years of working experience was 16.5 years, and practitioners were on average 41 years old. Each participant participated only once in the experiment. In table 1 the number of participants in each condition is specified.

The uptime percentage of brand B (*Uptime B*) was set at two levels (99.5% and 96.0%).

The second independent variable was the availability of TCO information (*TCO info*), which was either available or not. In conditions without TCO information, a column with a calculation of the “Cost per hour” such as in figure 1 was not provided. In conditions with TCO information, data such as in figure 1 were provided. However in this experiment the inaccuracy of monetary quantification was different: there was no number for the cost per

hour related to uptime (while this was included as a minimum or as a maximum cost estimation in the first experiment).⁶

The level of accountability (*Accountability*) was the third independent variable, and it was manipulated at two levels (no-accountability versus high-accountability). Participants in the no-accountability condition participated anonymously and were only asked to choose one alternative. These participants received a one-sided printed sheet on which the experimental task was printed (the other side was blank). Participants in the high-accountability condition received a double-sided printed sheet. Participants had to fill-out their name, e-mail address, and phone number on the front page (the page that was blank in the no accountability condition). Students could read on the front page that we needed those data to organize a meeting where we would discuss their choices. Practitioners could read on the front page that we needed their contact information to call them for additional questions concerning their choice. Practitioners in the high accountability condition were also asked to think aloud while conducting the experimental task.⁷

The coefficient of the interaction term of *Uptime B* \times *Experience* is estimated for testing H3. Note that the preference for Brand B is expected to decrease as the uptime of Brand B decreases. If the preference for Brand B decreases more strongly for Practitioners compared to Students, which is indicated by the interaction term *Uptime B* \times *Experience*, the weight of the attribute uptime is higher for Practitioners than for Students. Note that H3 applies only to conditions in which TCO info was available.

The coefficient of the interaction term is estimated using a logistic regression, with either of the following two specifications:

$$\begin{aligned} \text{Choice } B &= \beta_0 + \beta_1 \text{ Uptime } B + \beta_2 \text{ Accountability} + \beta_3 \text{ Experience} + \\ &\quad \beta_4 (\text{Uptime } B \times \text{Accountability}) + \beta_5 (\text{Uptime } B \times \text{Experience}) + \\ &\quad \beta_6 (\text{Accountability} \times \text{Experience}) + \\ &\quad \beta_8 (\text{Uptime } B \times \text{Accountability} \times \text{Experience}) \end{aligned}$$

$$\text{Choice } B = \beta_0 + \beta_1 \text{ Uptime } B + \beta_2 \text{ Experience} + \beta_3 (\text{Uptime } B \times \text{Experience})$$

Note that in the first specification, the coding of categorical variables affects the results for the second-order interaction effects. The coding of *Accountability* (high-accountability = 1,

⁶ In this experiment, participants had to choose between two alternatives (Brand A and B), two attributes shown in figure 1 for the first experiment were not included (Inspection time, and Chemicals), and the purchase price was provided as a single figure.

⁷ We recorded what they said only for the purpose of increasing the sense of accountability. The sound quality of the recordings also would not enable further analysis of the content of participants' considerations.

low-accountability = 0; or alternatively: low-accountability = 1, high-accountability = 0) affects the results for the interaction term *Uptime B* × *Experience*. For that reason, we also estimate a model without the variable *Accountability* (the second specification) and using the χ^2 -difference we test whether this variable can be excluded (Pampel 2000). If the χ^2 -difference is significant, the first specification must be estimated, so this must be done with both versions of the coding of *Accountability*, which essentially means: separately for high and low accountability.

For testing H4, we will investigate the coefficient of the three-way interaction of *Uptime B*, *TCO info*, and *Experience*. Note that the preference for Brand B is expected to decrease as the uptime of Brand B decreases. If the preference for Brand B decreases more strongly with TCO info compared to without TCO, then the provision of such info affects the weight of the uptime attribute. This is measured through the coefficient for the interaction term *Uptime B* × *TCO info*. If this effect also depends on the level of experience (as hypothesized), there should also be a significant coefficient for the three-way interaction of *Uptime B* × *TCO info* × *Experience*. We estimate a logistic regression including the four main effects, all two-way and three-way interactions, and a four-way interaction. We also estimate a logistic regression without considering *Accountability*, and a χ^2 -difference test is used to determine which empirical specification to estimate.

The coefficient of the interaction term of *Uptime B* × *Accountability* is estimated for testing H5. The reasoning for this is similar to the test for H3. Note that this hypothesis applies only to conditions in which TCO info was available. Alternative empirical specifications, with and without inclusion of the independent variable *Experience*, are estimated, for similar reasons as discussed above for H3.

Results and discussion

Results for H3 are in table 3 and in figure 4. Note that for this hypothesis we look only at conditions in the second experiment in which TCO info was available (N = 154). We expected that experience would change the weight of the attribute that is not included in the TCO numbers. The coefficient for the two-way interaction term *Uptime B* × *Experience* was significant (p = .020), hence we find support for this hypothesis. As figure 4 shows, students put a higher weight on the attribute uptime, which was not included in the TCO number in this experiment, compared to practitioners.

[Insert table 3 about here.]

[Insert figure 4 about here.]

Results for H4 are in table 4. The coefficient for the interaction term *Uptime B* × *TCO info* × *Experience* in column 3 was highly significant ($p = .006$), providing strong support for H6. As a result of providing TCO information, students gave a higher weight to *Uptime B*, while practitioners gave less weight to it. See also figure 5, which indicates that the interactive effects of *Uptime* and *TCO info* on decisions (*Choice B*) in figure 5a and 5b are different for Students and Practitioners.

[Insert figure 5 about here.]

Results for H5 are in table 3. We expected that accountability would increase the weight of the attribute not included in the TCO numbers (if TCO info is available). The coefficient of *Uptime B* × *Accountability* is not significant. Hence, we find no support for our hypothesis.

The lack of support for the effect of accountability in this experiment is surprising. We assumed that accountability would increase the weight of the attribute that was not included in the TCO numbers (H5). One possible explanation could be that the experimental task might have been too simple. The inaccuracy might have been too obvious to ignore, regardless of the manipulation of accountability. Maybe participants made the same trade off in all conditions between total costs and the attribute that was not included in this TCO information. A higher task complexity, for example through more attributes per alternative, might have been required to see an effect of accountability. For a more difficult task, decision makers who are not pushed to think more carefully (because they are not held accountable) might ignore the inaccuracy in the TCO numbers and attach less weight to TCO numbers, compared to decision makers who are held accountable.

Another explanation could be that selection bias in the experimental design may have countered the effect of accountability. Participants were asked to provide their contact details, as explained above. Their responses were excluded if they refused to provide such information. As can be seen from table 1, the number of participants in the high-accountability conditions is smaller than in the no-accountability conditions.⁸ As a result of

⁸ A total number of 350 students were asked to participate in this experiment. They were randomly assigned to one of the four conditions, but the willingness to participate and provide contact details was lower in the high-accountability conditions. This yielded 141 participants in the no-accountability conditions and 58 participants in the high-accountability conditions. For practitioners, the differences in the numbers of practitioners in these conditions were smaller than for students (70 and 58 participants). Practitioners participated sequentially (students simultaneously). The researcher kept searching for a practitioner who was willing to answer a randomly selected condition before changing to a new randomly selected condition. It was much more time consuming to find a practitioner who was willing to participate in the high-accountability condition, compared to finding a practitioner who was willing to participate in the no-accountability condition, because some practitioners who were willing to participate refused to continue the experiment after learning that they had to provide contact details and possibly had to answer questions afterwards.

self-selection, it could be that the resulting participants in the high-accountability conditions were more confident or were otherwise not particularly concerned about how their decisions would be judged. In other words, for those participants we may have failed to actually create accountability and anxiety for losing face. However, this limitation is not exclusive to our study, although few studies of accountability discuss it in great detail. For example, in accountability studies conducted by Antonioni (1994) and Schwartz et al. (2004) the number of participants in the high-accountability conditions was lower compared to non-accountable conditions. As in our study, participants could drop out of the experiment at any time, and they did more so under high accountability. Antonioni (1994) asked subordinates in an insurance company to evaluate their managers. Participants admitted that they preferred anonymous (non accountable) appraisals because reprisal might be costly. Participants in the study by Schwartz et al. (2004) were family physicians who selected a patient treatment. In the high-accountability condition, the physicians provided a written justification on their decision, which also disclosed the research participant's name. In both studies, it is possible that self-selection factors were stronger in the accountable condition than in the non-accountable condition.

5. CONCLUDING REMARKS

Sourcing decisions like the ones we described in this paper are cognitively challenging. The alternatives can be described by several attributes; some are financial, while other attributes describe are non-financial and describe, for example, quality. Humans will adapt their decision-strategy to the specific situation, and try to do so with minimum cognitive effort (Payne et al. 1993). Relying on TCO info is tempting, because monetary quantification facilitates tradeoffs among attributes that are non-comparable, as the measurement units of these attributes were initially not commensurable. However, the monetary quantification may often not be capable to fully capture the operational considerations, and “decisions can no longer be taken and actions can no longer be chosen at a distance via the abstract language of numbers. Operational considerations must to some extent be involved in the decision process, not coded, but in their original form.” (Chapman 1997, page 202) We investigated three variables that may affect how carefully decision makers consider the inaccuracy of the TCO info, and, hence, how much weight they gave to the inaccurately quantified attribute: reflective thinking, accountability, and experience.

Participants in our experiments had to choose between alternative offerings for replacing a production machine, each characterized by several attributes. One attribute

(downtime of the machine) was problematic in terms of providing a monetary quantification. In some conditions in the first experiment, the cost associated with downtime was included in the TCO info as minimum cost estimation. As hypothesized, it was found that the weight attached to the inaccurate included TCO numbers increased as a result of reflective thinking. In other conditions, the maximum cost associated with downtime was estimated and included the TCO info. We found no support for the hypothesis that the weight attached to this attribute was reduced as a result of reflective thinking. In the second experiment, downtime was completely excluded from the TCO calculation. We found no support for the hypotheses that accountability would affect the weight of this attribute. We found strong support for the hypotheses that experience would affect that weight of this attribute.

The results of this study may have managerial implications for the introduction and use of TCO information, and other costing information that aims to better capture the indirect costs of cost objects, but at the same time leaves out some important characteristics of these objects. Improved costing information that captures financial impact and aggregates financially quantifiable attributes may be helpful for the decision maker, but care should be given to unintended affects—that those elements that are not included in the new costing information get less attention as a result of providing the information. The results suggest that when introducing cost information, it is important that the limitations should be explained very clearly to managers, so that they will take a closer look at the accuracy of the cost information. It is probably also important to give a prominent position in costing reports to the excluded attributes. This might attract attention and result in judgment and decision-making where decision makers take these results into account.

Several limitations have to be mentioned. Conducting experiments on computers is preferable to running experiments by paper. A computer would have provided the opportunity to measure process variables such as decision time. Although we are aware of this limitation, we choose to conduct experiments on paper for several practical reasons. We wanted to include practitioners in our experiments. Although it is not too difficult to run experiments with students in a laboratory, it is much harder too find a group of practitioners. For that reason, we did not ask practitioners to come to our laboratory, but decided to go to locations where we would find many practitioners from many different companies. We used paper because we were not able to use computers on the trade fair. As we wanted to minimize the impact of different methods, we choose to conduct the experiments with students and practitioners on paper. Despite the fact that we had the intention to run all experiments with

both students and practitioners, the difficulty to find enough practitioners, (i.e., get access to trade fairs) made us decide to conduct only the second experiment with practitioners.

Future research could help to get a better understanding of the impact of TCO information on judgment and decision-making. The impact of a social context on judgment and decision-making is an interesting topic for a new study. In this study individual judgment and decision-making has been investigated. However, humans do not live in an isolated world, they live and work in setting where they interact with each others (Katz and Kahn 1978); future research could therefore research the use of inaccurate costing information in relation to reflective thinking in a social context. Additionally we could model the imperfectness of TCO numbers, which is central in this study, differently. While we modeled inaccuracy as minimum and maximum TCO numbers, we might also give explicit ranges for TCO numbers (i.e., costs are between €2.00 and €3.00). Investigating the impact of ranges on judgment and decision-making is relevant while human decision makers do deviate from normative theory by attaching not enough or too much weight to a range (Fischer 1995).

REFERENCES

- Alba, J. W. and J. W. Hutchinson (1987). "Dimensions of consumer expertise." *Journal of Consumer Research* 13(4): 411-454.
- Antonioni, D. (1994). "The effects of feedback accountability on upward appraisal ratings." *Personnel Psychology* 47(2): 349-356.
- Banker, R. D., H. Chang and M. J. Pizzini (2004). "The balanced scorecard: judgmental effects of performance measures linked to strategy." *The Accounting Review* 79(1): 1-23.
- Baron, J. (1981). "Reflective thinking as a goal of education." *Intelligence* 5(4): 291-309.
- Bettman, J. R., M. F. Luce and J. W. Payne (1998). "Constructive consumer choice processes." *Journal of Consumer Research* 25(3): 187-217.
- Bonner, S. E. (1994). "A model of the effects of audit task complexity." *Accounting, Organizations and Society* 19(3): 213-234.
- Boroah, V. K. (2001). *Logit and Probit: Ordered and Multinomial Models*. Thousand Oaks, CA, Sage.
- Briers, M., P. Lockett and C. Chow (1997). "Data fixation and the use of traditional versus activity-based costing systems." *Abacus* 33(1): 1-20.
- Cardinaels, E., F. Roodhooft and L. Warlop (2004). "Customer profitability analysis reports for resource allocation: the role of complex marketing environments." *Abacus* 40(2): 238-258.
- Carnevale, P. J. D. (1985). *Accountability of group representatives and intergroup relations. Advances in Group Processes: Theory and Research*. E. J. Lawler. Greenwich, Connecticut, JAI Press Inc. 2: 227-248.

- Carr, L. P. and C. D. Ittner (1992). "Measuring the cost of ownership." *Journal of Cost Management* 6(3): 42-51.
- Chambers, R. J. (1996). "Ends, ways, means and conceptual frameworks." *Abacus* 32(2): 119-132.
- Chang, C. J., J. L. Y. Ho and W. M. Liao (1997). "The effects of justification, task complexity and experience/training on problem-solving performance." *Behavioral Research in Accounting* 9(Supplement Conference): 98-116.
- Chapman, C. S. (1997). "Reflections on a contingent view of accounting." *Accounting, Organizations and Society* 22(2): 189-205.
- Cronbach, L. J. (1987). "Statistical test for moderator variables: flaws in analyses recently proposed." *Psychological Bulletin* 102(3): 414-417.
- Curley, S. P., F. J. Yates and R. A. Abrams (1986). "Psychological sources of ambiguity avoidance." *Organizational Behavior and Human Decision Processes* 38(2): 230-256.
- Dearman, D. T. and M. D. Shields (2001). "Cost knowledge and cost based judgment performance," *Journal of Management Accounting Research* 13: 1-18.
- Degraeve, Z., F. Roodhooft and B. v. Doveren (2005). "The use of total cost of ownership for strategic procurement : a company-wide management information system." *The journal of the Operational Research Society* 56(1): 51-59.
- Dewey, J. (1933). *How We Think. A Restatement of the Relation of Reflective Thinking to the Educative Process*, D. C. Heath and Company.
- Doney, P. M. and G. M. Armstrong (1996). "Effects of accountability on symbolic information search and information analysis by organizational buyers." *Academy of Marketing Science* 24(1): 57-65.
- Ellram, L. M. (1995). "Total cost of ownership. An analysis approach for purchasing." *International Journal of Physical Distribution & Logistics Management* 25(8): 4-23.
- Fischer, G. W. (1995). "Range sensitivity of attribute weights in multiattribute value models." *Organizational Behavior and Human Decision Processes* 62(3): 252-266.
- Fox, F. V. and B. M. Staw (1979). "The trapped administrator: effects of job insecurity and policy resistance upon commitment to a course of action." *Administrative Science Quarterly* 24(3): 449-471.
- Galbraith, J. R. (1973). *Designing Complex Organizations*. New York, Addison-Wesley Pub. Co
- Gupta, M. and R. R. King (1997). "An experimental investigation of the effect of cost information and feedback on product cost decisions." *Contemporary Accounting Research* 14(1): 99-127.
- Isselin, E. R. (1988). "The effects of information load and information diversity on decision quality in a structured decision task." *Accounting, Organizations and Society* 13(2): 147-164.
- Ittner, C. D., D. F. Larcker and M. W. Meyer (2003). "Subjectivity and the weighting of performance measures: Evidence from balanced scorecard." *The Accounting Review* 78(3): 725-758.
- Jaccard, J. (2001). *Interaction Effects in Logistic Regression*. Thousand Oaks, CA, Sage.

- Jaccard, J., R. Turrisi and C. K. Wan (1990). *Interaction Effects in Multiple Regression*. Thousand Oaks, CA, Sage.
- Kadous, K., L. Koonce and K. L. Towry (2005). "Quantification and persuasion in managerial judgement." *Contemporary Accounting Research* 22(3): 643-686.
- Kahneman, D. and A. Tversky (1979). "Prospect theory: an analysis of decision under risk." *Econometrica* 47(2): 263-291.
- Katz, D. and R. L. Kahn (1978). *The Social Psychology of Organizations*, John Wiley & Sons, Inc.
- Keeney, R. L. (1977). "The art of assessing multiattribute utility functions." *Organizational Behavior and Performance* 19(2): 267-310.
- Koriat, A., S. Lichtenstein and B. Fischhoff (1980). "Reasons for confidence." *Journal of Experimental Psychology: Human Learning and Memory* 6(2): 107-118.
- Langer, E. J. (1978). *Rethinking the role of thought in social interaction. New Directions in Attribution Research*. J. H. Harvey. Hillsdale, NJ, Erlbaum. 2: 35-58.
- Langer, E. J. (1989). "Minding matters: the consequences of mindlessness-mindfulness." *Advances in Experimental Social Psychology* 22: 137-173.
- Lillis, A. M. (2002). "Managing multiple dimensions of manufacturing performance - an exploratory study." *Accounting, Organizations and Society* 27(6): 497-529.
- Lipe, M. G. and S. Salterio (2000). "The balanced scorecard: judgmental effects of common and unique performance measures." *The Accounting Review* 75(3): 283-298.
- Lipe, M. G. and S. Salterio (2002). "A note on the judgmental effects of the balanced scorecard's information organization." *Accounting, Organizations, and Society* 27(6): 513-540.
- Marchant, G., J. Robinson, U. Anderson and M. Schaedewold (1991). "Analogical transfer and expertise in legal reasoning." *Organizational Behavior and Human Decision Processes* 48(2): 272-290.
- Menard, S. (1995). *Applied Logistic Regression Analysis*. Thousand Oaks, CA, Sage.
- Nelson, M. W., R. Libby and S. E. Bonner (1995). "Knowledge structure and the estimation of conditional probabilities in audit planning." *The Accounting Review* 70(1): 27-47.
- Pampel, F. C. (2000). *Logistic Regression - A primer*. Thousand Oaks, CA, Sage.
- Payne, J. W., J. R. Bettman and E. J. Johnson (1993). *The Adaptive Decision Maker*. Cambridge, Cambridge University Press.
- Pryor, J. B., F. X. Gibbons, R. A. Wicklund, R. H. Fazio and R. Hood (1977). "Self-focused attention and self-report validity." *Journal of Personality* 45(4): 513-527.
- Raiffa, H. (1968). *Decision Analysis. Introductory Lectures on Choices under Uncertainty*. Reading, Massachusetts, Addison-Wesley.
- Reck, J. L. (2001). "The usefulness of financial and nonfinancial performance information in resource allocation decisions." *Journal of accounting and public policy* 20(1): 45-71.
- Schiff, A. D. and L. R. Hoffman (1996). "An exploration of the use of financial and nonfinancial measures of performance by executives in a service organization." *Behavioral Research in Accounting* 8: 134-153.

- Schwartz, J., G. Chapman, N. Brewer and G. Bergus (2004). "The effects of accountability on bias in physician decision making: going from bad to worse." *Psychonomic Bulletin & Review* 11(1): 173-178.
- Shelton, S. W. (1999). "The effect of experience on the use of irrelevant evidence in auditor judgment." *The Accounting Review* 74(2): 217-224.
- Sherlock, P. (1991). *Rethinking Business to Business Marketing*. New York, The Free Press.
- Shields, M. D. (1980). "Some effects of information load on search patterns used to analyze performance reports." *Accounting, Organizations and Society* 5(4): 429-442.
- Shields, M. D. (1983). "Effects of information supply and demand on judgment accuracy: evidence from corporate managers." *The Accounting Review* 58(2): 284-303.
- Simon, H. A. (1979). "Rational decision-making in business organizations." *Economic Sciences* 69(4): 493-513.
- Snyder, M. and D. Kendzierski (1982). "Acting on one's attitudes: procedures for linking attitude and behavior." *Journal of Experimental Social Psychology* 18(2): 165-183.
- Stocks, M. H. and A. Harrell (1995). "The impact of an increase in accounting information level on the judgment of quality of individuals and groups." *Accounting, Organizations and Society* 20(7/8): 685-700.
- Tetlock, P. E. (1983). "Accountability and Complexity of Thought." *Journal of Personality and Social Psychology* 45(1): 74-83.
- Tetlock, P. E. (1985). "Accountability: the Neglected Social Context of Judgment and Choice." *Research in Organizational Behavior* 7: 297-332.
- Tetlock, P. E. (1992). "The Impact of Accountability on Judgment and Choice: Toward a Social Contingency Model. " in: M. P. Zanna (ed.). *Advances in Experimental Social Psychology*. 25: 331-376.
- Vera-Muñoz, S. C. (1998). "The effects of accounting knowledge and context on the omission of opportunity costs in resource allocation decisions." *The Accounting Review* 73(1): 47-72.
- Vera-Muñoz, S. C., W. R. Kinney, Jr. and S. E. Bonner (2001). "The effects of domain experience and task presentation format on accountants' information relevance assurance." *The Accounting Review* 76(3): 405-429.

Table 1
Total number of respondents and the choices per condition

First Experiment

<i>Reflection</i>		Minimum cost estimation			Maximum cost estimation		
		No reflection	Reflection	Total	No reflection	Reflection	Total
<i>Uptime C</i>	96%	15 (4) ¹	14 (1)	29 (5)	35 (7) ²	35 (4) ³	70 (11)
	99%	14 (8)	14 (12)	28 (20)	37 (27)	38 (24)	75 (51)
Total		29	28	57	72	73	145

¹ Between parentheses: number of respondents choosing Brand C, from A, B or C

² Two participants choose Brand B in this condition (not included in the numbers in the table).

³ One participant choose Brand B in this condition (not included in the numbers in the table).

Second Experiment

<i>Accountability</i>	No accountability				High accountability¹				
	No TCO		TCO		No TCO		TCO		
<i>TCO info</i>									
<i>Experience</i>	Students	Practitioners	Students	Practitioners	Students	Practitioners	Students	Practitioners	
<i>Uptime B</i>	96.0%	39 (7) ²	20 (3)	35 (11)	16 (7)	15 (2)	19 (3)	11 (4)	12 (4)
	99.5%	34 (14)	16 (12)	33 (30)	18 (12)	17 (10)	13 (9)	15 (14)	14 (10)
Total	73	36	68	34	32	32	26	26	

¹ Participants who did not provide contact information (37 students, and 5 practitioners) are excluded.

² Between parentheses: number of respondents choosing Brand B, from A or B

Table 2
 Logistic regression results first experiment (H1 and H2)

<i>Predictor¹</i>	Minimum TCO <i>N</i> = 57	Maximum TCO <i>N</i> = 145
<i>Constant</i>	-.373 (.349)	-.155 (.578)
<i>Uptime C</i>	.015 (.102)	.028 (.000)
<i>Reflection</i>	-.051 (.946)	-.554 (.184)
<i>Uptime C ´ Reflection</i>	.036 (.043)	.002 (.806)

Values of coefficients (and *p*-values for two-tailed Wald test)

Table 3
Logistic regression results second experiment, for conditions with TCO info

<i>Choice B</i> is dependent variable	Regression for testing H3 ¹	Regression for testing H5 ²
<i>Constant</i>	.897 (.004)	.509 (.032)
<i>Uptime B</i>	.032 (.000)	.022 (.000)
<i>Accountability</i>		.004 (.992)
<i>Experience</i>	-.696 (.091)	
<i>Uptime B</i> × <i>Accountability</i>		.001 (.950)
<i>Uptime B</i> × <i>Experience</i>	-.019 (.020)	
<i>N</i>	154	154

Values of coefficients (and *p*-values for two-tailed Wald test)

¹ A specification including *Accountability* as an independent variable was also estimated. Because the Chi-square difference was not significant ($\chi^2_{df=4} = .572$, $p = .966$), we estimated the coefficient for *Uptime B* × *Experience* using the empirical specification provided in the table above.

² A specification including *Experience* as an independent variable was also estimated. Because the Chi-square difference was not significant ($\chi^2_{df=4} = 7.817$, $p = .099$), we estimated the coefficient for *Uptime B* × *Accountability* using the empirical specification provided in the table above.

Table 4

Logistic regression results second experiment, N = 327 (for H4)

<i>Choice B</i> is dependent variable	Regression for testing H4 ¹
<i>Constant</i>	-.880 (.000)
<i>Uptime B</i>	.015 (.001)
<i>Experience</i>	.481 (.208)
<i>TCO info</i>	1.682 (.000)
<i>Uptime B</i> × <i>Experience</i>	.012 (.122)
<i>Uptime B</i> × <i>TCO info</i>	.017 (.033)
<i>Experience</i> × <i>TCO info</i>	-1.120 (.045)
<i>Uptime B</i> × <i>TCO info</i> × <i>Experience</i>	-.031 (.006)
<i>N</i>	327

Values of coefficients (and p-values for two-tailed Wald test)

¹ A specification including *Accountability* as an independent variable was also estimated. Because the Chi-square difference was not significant ($\chi^2_{df=8} = 2.288$, $p = .971$), we estimated the coefficient for *Uptime B* × *TCO info* × *Experience* using the empirical specification provided in the table above.

Figure 1

Quantitative information provided in first experiment (*Uptime C = 99%*)

Labor cost per operator (€/hour)	€ 30,00
Electricity price (€/Kwh)	€ 0,13
Chemicals (€/liter)	€ 10,00

	Brand A Cost per hour		Brand B Cost per hour		Brand C Cost per hour	
Life time (hours)	2.900		2.700		2.800	
Uptime (%)	99,0%	€ 1,01*	96,0%	€ 6,25*	99,0%	€ 1,41*
Inspection time (minutes/hour)	9,0	€ 4,50	11,0	€ 5,50	10,0	€ 5,00
Purchasing price (€): Component A	€ 2.800	€ 0,97	€ 2.300	€ 0,85	€ 2.100	€ 0,75
Component B	€ 1.700	€ 0,59	€ 1.400	€ 0,52	€ 1.300	€ 0,46
Component C	€ 2.000	€ 0,69	€ 1.600	€ 0,59	€ 1.600	€ 0,57
Component D	€ 1.300	€ 0,45	€ 1.100	€ 0,41	€ 1.000	€ 0,36
Energy (kwh/hour)	7,0	€ 0,91	6,0	€ 0,78	5,0	€ 0,65
Chemicals (liter/hour)	0,09	€ 0,90	0,09	€ 0,90	0,08	€ 0,80
Total cost per hour		€ 10,01		€ 15,80		€ 10,01

* Minimum downtime cost estimation

Figure 2

Text introducing the task in the first experiment (No refecton)

“You are the manager in a company. One machine has to be renewed. You can choose between three brands, all three meet all specifications.

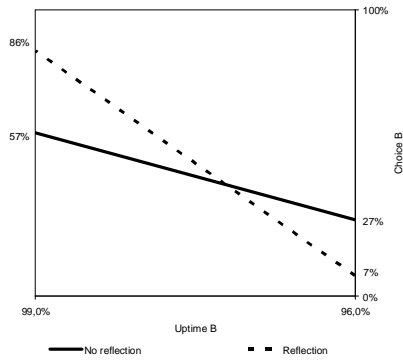
In the table below information is provided about:

- The lifetime of a new machine;
- The uptime percentage of the machine;
- The time an employee needs to check the machine;
- The purchasing price of the four component new machine;
- Energy consumption per hour;
- The amount of chemicals necessary.

The ‘uptime (%)’ is the percentage of the time the machine is available for production. 100% – ‘uptime (%)’ = ‘downtime (%)’. Downtime is caused by machine failure, maintenance, etc. that cause rescheduling of production, not meeting delivery times, etc. Some downtime cost, like failure cost, maintenance cost, etc. can be calculated objectively. Objectively calculable costs are included in the table below. Other downtime cost, like cost of not reaching times of delivery or additional transports are highly uncertain and therefore not included in the table below.”

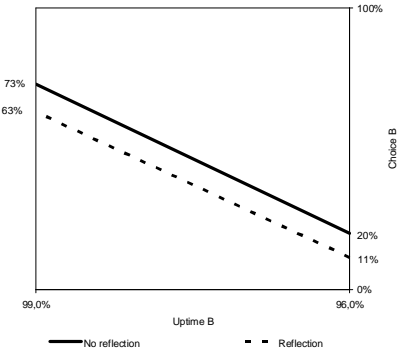
Figure 3
Interaction of *Reflection* by *Uptime C*, first experiment

Figure 3a
Minimum TCO, N = 57



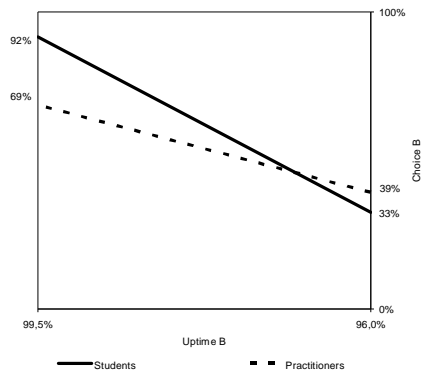
Percentage of participants choosing Brand C

Figure 3b
Maximum TCO, N = 145



Percentage of participants choosing Brand C

Figure 4
Interaction of *Experience* by *Uptime B*, N = 154

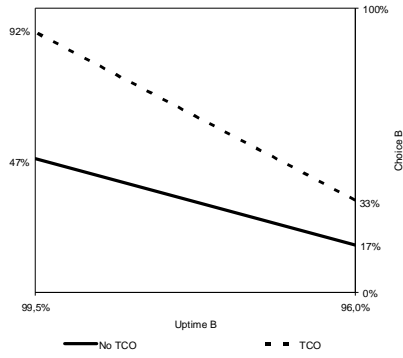


Percentage of participants choosing Brand B in second experiment, for conditions with TCO info

Figure 5

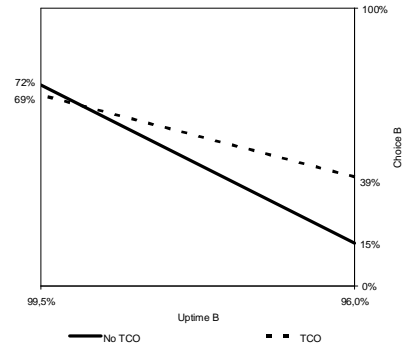
Interaction of *TCO info* by *Uptime B*, for two levels of *Experience*

Figure 5a
Students, N = 199



Percentage of students choosing Brand B

Figure 5b
Practitioners, N = 128



Percentage of practitioners choosing Brand B