

Measuring customer satisfaction: Why, what and how

Janet McColl-Kennedy & Ursula Schneider

To cite this article: Janet McColl-Kennedy & Ursula Schneider (2000) Measuring customer satisfaction: Why, what and how, Total Quality Management, 11:7, 883-896, DOI: [10.1080/09544120050135434](https://doi.org/10.1080/09544120050135434)

To link to this article: <https://doi.org/10.1080/09544120050135434>



Published online: 25 Aug 2010.



Submit your article to this journal [↗](#)



Article views: 399



Citing articles: 25 View citing articles [↗](#)



Measuring customer satisfaction: why, what and how

JANET MCCOLL-KENNEDY¹ & URSULA SCHNEIDER²

¹University of Queensland, Brisbane, Australia & ²Karl-Franzens-Universität, Graz, Austria

ABSTRACT *This paper seeks to challenge researchers and business organizations to think about the measures they are using in their attempt to measure customer satisfaction and any subsequent decision-making and actions that may result. Specifically, the paper endeavours to raise awareness of the difficulties involved in measuring customer satisfaction and of using these measures for decision-making. The assumption associated with the measurement instrument and the methods of survey, together with the advantages and disadvantages of standardized vs customized instruments are explored. Next, the partially contradictory objectives of research and business and the frequent necessity of making trade-offs are discussed. In conclusion, the paper offers suggestions regarding what we can do in terms of customer satisfaction measurement. Firstly, we should see the procedure of measurement of customer satisfaction as no neutral act but as an intervention which affects subsequent interaction with our customers. Secondly, we should always remember that as organizations we are trying to nurture relations with our customers, not merely to measure and document what we have found in our research. Thirdly, we should be prudent in our use of measures and use these as yardsticks in a learning process. Finally, we should remember that we need standardized and repeated measures for statistical analysis but that this may not be valued by business organizations.*

Introduction

This paper will not elaborate on the technical side of measuring customer satisfaction. Its purpose is rather to position the measuring endeavour within a broader epistemological and strategic framework and to establish criteria for the controlling of intangibles in a business context. In the first section two different management paradigms will consider customer satisfaction. This section will question standardized questionnaire-type surveys that are so common in service industries. Such surveys are burdensome, often without consequences and rather backward-oriented. In the second section, major measuring requirements in a business context will be discussed and their validity using Polanyi's concept of tacit knowledge. The paper is thus aimed at second-order or double-loop learning (Argyris & Schön, 1996) with regard to customer satisfaction. Second-order learning focuses on the feasibility of frames and premises that support the further elaboration of issues, while first-order learning leaves those frames and premises unquestioned in order to progress within them. While the latter seems more important to the practical field at first sight, the former can contribute

Correspondence: Associate Professor Janet McColl-Kennedy, Graduate School of Management, The University of Queensland, Brisbane QLD 4072, Australia

advice to withdraw from unpromising endeavours. As Pascale put it: “nothing is less productive than to make more efficient what should not be done at all” (Pascale, 1991). The conclusion will summarize reservations against customer surveys and codified measures while still arguing that in a learning context they are indispensable.

Why measure customer satisfaction?

In the ‘new economy’ (Arthur, 1996; Drucker, 1993; Schneider, 1996) knowledge is a resource as well as, increasingly, a product: with tangible goods becoming globally standardized and best practices travelling fast, companies gain competitive advantages through constant innovation, better targeting of customers and additional services. Those strategies cannot be applied to the arm’s length type of customer relations. The higher the innovative and service component, the more the customer becomes part of the performance equation. Customer relations then constitute an important asset that should be monitored just like physical assets. Most emerging approaches to the measurement of intellectual capital agree on the importance of customer capital, as expressed in sales, satisfaction and reputation (Edvinsson & Malone, 1997; Kaplan & Norton, 1996; Schneider, 1996; Sveiby, 1997). Accordingly, those approaches distinguish between reference customers (reputation), new customers or first trial customers (new sales) and repeated customers (satisfaction, sales). Independently of approaches to the measurement of intellectual capital, marketing literature has suggested a wide array of industry-specific models to monitor customer satisfaction (for an overview see Bearden *et al.*, 1996, also Hayes, 1992).

We can therefore conclude that management and marketing theorists as well as practitioners agree on the importance of customer relations for a business’s success. In order for that vision not to remain pure rhetoric, it is important to put into operation the concept of customer relations so that it can be monitored and managed. Most pundits would agree on that. But that is where unanimity stops. When confronted with the task to define the vision of ‘understanding customer relations’, management theorists will supply different answers depending on contradicting predispositions of different paradigms.

Different epistemological approaches and their consequences on measuring customer relations

We shall now discuss two management concepts and argue that measurement is more important to one of them. We shall keep our discussion short as this is not the context to discuss deeply different ways of knowing. Classical management theory is grounded in the hypothesis of rational decision-making. It is assumed that decision-making improves with the quantity of information available to a decision-maker. Decision-makers are supposed to use their own and third party knowledge according to purpose, without preferences. It is equally assumed that decision-makers can learn from surrogates, that is, acquire the qualifying context together with codified data. Based on those assumptions, management theory advises managers to accumulate data, to extend human information processing capacity by hardware and software and to mine data and texts available from operating procedures for hidden patterns that could improve future decision-making. Classical theory-based approaches hold if there are linear causal relationships that can be generalized for different contexts, that is, if people behave in a predictable manner that is stable over time. Behaviour need not be deterministic, as long as we can find a distribution that delivers good approximations, such as normal distribution that is assumed in many data mining procedures.

If, on the other hand, we let go of some assumptions of classical theory, as they seem unrealistic, we have to deal with causal loops (interdependencies), microdiversity that is not

necessarily normally distributed, with path dependency and partial irreversibility of processes. CAS (Complex Adaptive Systems) theory accounts for those premises and can, in addition, allow for behavioural specificities, such as worse decision-making with increasing knowledge, subjective preferences for sources and presentational forms of information, conscious and unconscious manipulation of data, and others. Partial path dependency (current decisions of a consumer are influenced but not determined by his/her experiences with past decisions) and partial irreversibility (money spent on one item cannot be spent on another) contribute to non-average outcome (Wollin, 1999). Non-average outcomes are no problem to statisticians as long as they do not affect macrostability, that is, the behaviour of an overall system in observation. This is the case if they smooth each other out and follow a predictable pattern. To assume smoothing out and predictability is not, however, suitable if numbers are small (such as with few dominant customers, quite common in business-to-business relations), if elements of an entity are isolated or only loosely coupled in terms of space, time or common domains (such as for culturally different customers or groups that are not exposed to the Internet and television), if there is irreversibility (such as in buying products that require learning and would lead to sunk costs, if abandoned) and if we account for purposeful breaking of patterns as acts of innovation (cf. Wollin, 1999). In the latter cases seemingly random aberrations in a system (such as street kids wearing rugs) can become triggers around which new patterns emerge (in the example above, wearing rug-type cloths could turn into fashion).

Even if we maintain the rather narrow assumptions of classical theory we run into some problems of empirical research in social sciences that are as well known, and constantly ignored. Those are representations on the one hand (are customers ready to be interviewed on the phone a valid or biased selection of all customers in question?) and measurement procedures as well as interpretation of results on the other: Which indicators do we use, how stable are those indicators over time and with regard to measurement intervention and how do they interrelate? How can we observe without bias in a world of unlimited stimuli? Does it make sense to generalize all customers for all categories of products and services under all types of situational conditions. Of course, those questions are not new. We do not claim that they are. What is amazing, however, is that they are not taken seriously. In economics we usually argue as follows. Let us assume my grandma has four tyres. She can thus be defined to be a bus. As she is defined to be a bus, she will run on petrol. Therefore let us feed her petrol. This metaphor of course exaggerates, usually we do not die from conclusions drawn from economic theory. Still, we would like to see more discussions of the question what value of information we can gain from measurement. Usually this question is treated as technical. The expectation is that humankind will develop ever better methods to deal with them. But, what if they were undecidable in principle and what if we did harm in the meantime, like to poor grandma in the metaphor? Then, the ultimate argument, supplied by traditional science, that although what we have is unsatisfactory, we do not have an alternative would not be very valid anymore.

All our studies are necessarily burdened with severe caveats as to the samples (access to data), reductions and isolations included in the method. The original authors diligently dedicate a section of their papers to those caveats but they seem to be lost in the process of diffusion of knowledge within the scientific community and, even more so, in the practical field. Feeble hypotheses thus gain the force of powerful truths and are acted upon. This may even produce empirical evidence—in the form of self-fulfilling prophecies. To conclude, it is not so much the theoretical weaknesses of our models that measure ‘reality’ that constitute the problem, but rather our tendency to forget those weaknesses and treat results as true images of reality.

From a practical point of view the collection of ever more data produces the following problems:

- Human beings, as the sources of many of those data (especially on attitudes, opinions, future plans) will become increasingly unwilling to bear the transaction costs of providing them. Customers have started to feed back that they feel molested by questionnaires (especially as they only get standardized reactions to very idiosyncratic responses). Some companies have made it a policy no longer to react to the myriads of questionnaires received from PhD students. From a game theory perspective, free riding on the information produced by surveys is the alternative with the highest pay-off, while co-operation does not produce an easy-to-quantify benefit, if any at all.
- The more sophisticated the procedures to collect and process data, the higher the danger that data collecting and processing become ends to themselves. This constitutes 'controlling-bias' (Schneider, 1998).

We can conclude this paragraph by stating that measuring customer satisfaction inserts itself into the classical paradigm of management and is exposed—even within this simplifying paradigm—to a number of difficulties. As this issue is about the measurement of customer satisfaction, we will concentrate on measurement in the following section. But a warning seems appropriate that the whole measurement endeavour, which has become so popular under the headline of intellectual capital, might end up like the efforts to measure organizational concepts in contingency theory. After millions had been spent, researchers came home not with the Holy Grail but with a broken teapot (Kieser-Kubicek, 1992, p. 4).

We therefore formulate the following cautious proposition: paradigm shifts in management theory have led us to account for system thinking and more realistic assumptions on behaviour. Although idiosyncrasies are probably exaggerated in societies that develop common and habitual patterns of expectations through a number of media (cf. Luhmann, 1996), we have to distinguish systems with average outcomes from systems with non-average outcomes. While we can apply measuring according to standardized statistical procedures to the former, they are meaningless to the latter. Table 1 contrasts two basic epistemological paradigms in a black-and-white manner to enlighten their differences and implications for research.

What do we measure?

Roughly spoken, the chain of argument in management theory is the following. Finding out about customer preferences will allow one to provide customized products and superior service to current customers which will entail further sales as well as a boost in image so that new customers can be gained. Customer feedback helps continuously to improve performance. In particular, it can inspire employees to increase their efforts (Fig. 1).

A corresponding research design will usually contain three major constructs, namely service quality as expressed by several indicators, customer satisfaction, another construct that can only be measured by using indicators and a third construct, namely success, again to be defined by indicators. On all three categories of indicators there is no general consensus but rather competing ideas, resulting in competing theories and consultancy products.

Depending on the product and channel of distribution, several models have been developed for the constructs of product and satisfaction. Characteristics and prices of products, speed of delivery, friendliness and competence of personnel involved, and time of recovery are indicators used frequently for the first construct, while satisfaction is measured in emotional/attitudinal dimensions (such as feeling esteem) and in action-oriented dimensions (such as readiness to repurchase; see Bearden *et al.*, 1996). New technologies make it

Table 1. *Epistemological approaches and measurement*

Approaches	Understanding ‘reality’
Classical science (positivistic, Popperian)	<p><i>Basic assumptions:</i> Phenomena are given a ‘whole’ can best be understood by dividing it into isolated parts and by adding the knowledge on those parts</p> <p><i>Rules:</i></p> <ul style="list-style-type: none"> • describe the phenomena to be studied as accurately as possible • formalize (the more maths the more scientific) • design methods of objectivated perception (measurement) to exclude subjective bias • synthesize your findings
Interpretative, constructivist (constructionist) approaches	<p><i>Basic assumptions:</i> Phenomena are not given, but socially constructed A ‘whole’ can best be understood by experiencing it as such, thus by intuition and empathy</p> <p><i>Rules:</i></p> <ul style="list-style-type: none"> • Do not focus on isolated elements but on their interrelations • Invest effort in the construction of phenomena, experiment with different constructions • Recognize pattern rather than achieving accuracy in isolated details (fuzzy rather than predicative logic)

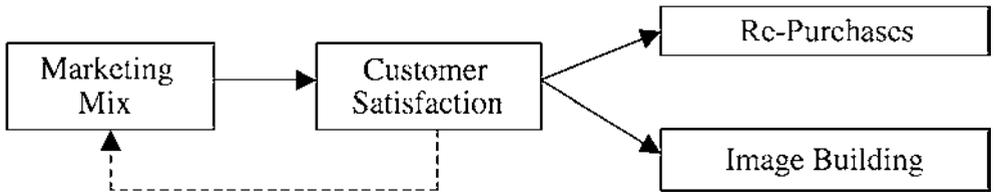


Figure 1. *Basic research design to measure customer satisfaction.*

a lot easier for all parties involved to do research on a continuous basis. “Driving the customer focus is a new breed of technology, including database tools that let companies gather information about their customers like never before, sales force applications that let them deliver service and Web technologies that let them establish more personalized relationships with customers . . .” (*InformationWeek*, compiled by Creemers, 1999). But new technologies do not eliminate the critical aspects of the basic research design. We are insecure about the relationship between constructs. Satisfaction, for instance, may not always lead to returns. In turn, returns do not automatically mean success (some customers simply are not profitable as better costing would show; see *Information Week*, compiled by Creemers, 1999). Moreover, we are insecure about the relationship between indicators and constructs. Figure 2 shows a general research design.

Thirdly, we cannot fully control the situation in which a survey takes place. To take verbal reactions as a proxy for the attitudes and (planned) actions of respondents may be wrong, as polls before elections have shown. Therefore, despite all the effort invested in valid testing instruments construct validity will remain an unsolved issue. Without entering into detail we could sum up that in order to obtain higher construct validity, frequent and very sophisticated testing is needed. To maintain a positive relationship of costs and benefits while

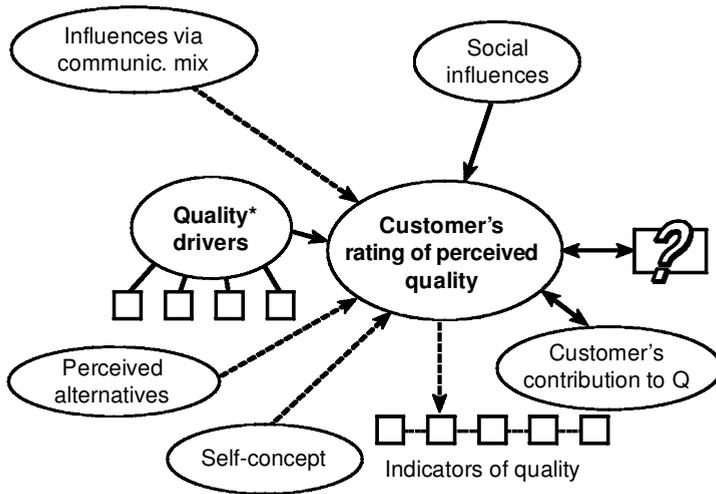


Figure 2. Enlarged model of influence factors on customer satisfaction.

measuring intangibles, on the other hand, asks for simple procedures. This refers us to the second part of our paper: Will measuring intangibles pay off?

The model presented above could be amplified by an additional construct relating employee satisfaction to product/service quality, as has been suggested by various sources. As said before, there is a trade-off between different research criteria: the more factors we include, the less tractable the model becomes, especially if we account for polynomic functions, or even 'worse' for exponential functions. Translated back into a marketing context the model works best if customers, or groups of target customers, behave in a similar manner. This can be assumed if they are either influenced by the same type of cultural formation as well as commercial advertising and/or if they depend on their mutual behaviour. Both assumptions were plausible in a mass production society; the latter continues to be so as best purchasing practices are diffused by literature, conferences and consultants. On the other hand, the Internet provides business as well as end-customers with opportunities to innovate, so that small numbers and sudden shifts in behaviour cannot be excluded.

Which model should we underlie our idea of monitoring intangible assets under the considerations given above? This depends on the advice required for future action. If, for instance, the Swiss in general and on average are 97% satisfied with their medical treatment (cf. Bruhn & Grund, 1999), this can mean different things: those reached on the phone and ready to devote time, who may be biased toward the more harmony-seeking part of a population, either claim to be satisfied *vis-à-vis* a research authority or are really satisfied. Can we know for sure that Swiss doctors are better than US doctors where the average satisfaction rate is much lower? Maybe. But, we could also find an explanation referring to the different organization of healthcare in both countries or in relation to a more indirect culture in Switzerland that would not express discontent right away (cf. Hall & Hall, 1990). As we see, the advice that can be gained from such expensive endeavours as national customer satisfaction indices is not very specific. As long as different groups can be held responsible there is not much hope for consequential action and if there is no consequential action, what should the effort be good for? On the other hand, let us assume people are asked more specifically and reveal they are not wholly satisfied with their encounters with doctors. We would then still not know whether this was due to a perceived lack in competence, to waiting

time or not enough time devoted or to the doctor not giving any explanations. The more detailed the study would be, the more probable that it really delivers information but also that only a biased minority will be ready to answer a phone interview and that the whole study will become more cost-intensive. We would concede that, in principle, a national index may trigger action by the legislator, by professional federations, by the educational system, by doctors themselves or by patients seeking treatment elsewhere. But given the pluri-responsibility, the biased and vague type of information and insecurity as to the quality of higher scores in other countries, this is highly improbable. Feedback on customer satisfaction, as we see it, makes more sense at an industry or company level. At an overall level it raises some serious questions with regard to costs, liability and general accessibility, which will be discussed later.

Criteria for measurement in a business context

Efficiency and manageability

In a business—as compared to a pure research context—efficiency becomes an important criterion beyond validity, reliability and objectivity. As a general rule, measurement makes sense as long as its cost are outweighed by the benefits generated by the ‘information added’ through the gathering and processing of data. The problem is, companies can hardly determine the costs of measurement and, even less so, its benefits. In general, better information should lead to better decisions, and thus to better success. Apart from decisions not always been implemented as planned, behavioural theory shows us that there is no linear relationship between the quantity of information and the quality of decision-making (Dörner, 1989; Schneider, 1990)

Furthermore, due to the information paradox, companies must bear the costs to develop a measurement system before they can evaluate its benefits. They can alleviate the paradox by letting competitors be the pioneers and then imitate them (benchmarking). But this would mean to lag behind. Despite there being no precise answer to the requirement of efficiency, we can nevertheless formulate some heuristics: the measurement of intangibles should be organized in a way that ‘minimizes’ its costs and ‘maximizes’¹ its use.

To minimize costs, simple and standardized procedures are recommended. They should mainly rely on data gathered for other purposes (synergy). The latter should be gathered and documented as a by-product of operations so that no additional staff are needed. To maximize use, however, requires other recommendations. Management must ensure that information reaches the ‘right’ decision-makers in the ‘right’ time, is interpreted ‘correctly’ and translated into action. This is usually not considered an issue of measurement, but of the further use of measures taken.

If we bear the context of use in mind while designing new instruments to measure intangibles, we must consider some empirical evidence that refers to behaviours other than purely ‘rational’² in the context of business organizations.

Organizational context

Following the broader discourse on measuring not only customer satisfaction but also other intangibles, we assume in this paragraph that measures are generated internally and provided for reasons of internal rather than external reporting. We shall discuss some possibly harmful implications of the simple fact that measures are compiled by those to be measured (although it will not be the same persons within an organization who do and control). Observations of

how (top) managers decide supply a very strong argument for translating narratives into financial codes. It is Peter Drucker's observation that managers only understand the language of numbers: what is not measured will consequently not be managed (Drucker, 1993, p. 44).

On the other hand, Brown has found that managers cannot digest more than a few chunks of quantitative information. Five to seven codes that should be defined clearly and simply are acceptable, more sophisticated measures will be rejected (Brown, 1997, pp. 15–20).

Organizational behaviour literature supplies the following observation. Companies do not establish departments and divisions because they grow, but grow in numbers of people employed because they establish departments and divisions. As mentioned before, a new function, 'controlling of intangibles', would elicit its specialists to prove their indispensability and importance, probably by developing ever more sophisticated measurement systems. Owing to the results mentioned above, those reports would be widely ignored and therefore wasteful (Schreyögg, 1996, p. 180).

A fourth category of studies we should consider here is related to the coupling of single measures with appraisal systems and incentive schemes. Although it seems quite reasonable at first sight, the use of codes as 'carrots' implies some dangers: appraisees will tend to manipulate the measurements (as to the timing of measurement or the exposure of the behaviour in question only at the moment of measurement). Alternatively, they will tend to focus on the code 'to the letter' and thus possibly counteract its substantial meaning or its interaction with other criteria which are difficult to put into operation. We could talk here of the tendency of the quantifiably suppressing quality. Measures used within appraisal and incentive systems are submitted to the same contradictory requirements as measures for intangibles in general: to be valid they should be differentiated, sophisticated and only used as a set of interdependent factors; to be manageable, however, they must be 'simple and stupid'.

Another dialogue within the emerging community of people interested in measuring intangibles centres around the issue of standardized versus customized instruments. Standardized measures seem suited for external reporting. They allow for comparability. They can be developed by public or private experts and acquired either as a public good or at much lower costs than customized instruments.

Reliability and objectivity can be more easily guaranteed. Standardized measures produce the portability of all knowledge related to their use as an externality. But standards could also be defined as the minimal common denominator, as the outcome of compromise. They might be defined too vaguely to inspire action or be too formalized to be more than an exercise in documentation and an argument to be used in marketing.

Standardized measures may entail the danger of form overruling content, certification of a quality becoming more important than its generation. The criticism of the ISO certification procedure as a bureaucratic, backward-oriented, innovation-preventing routine points to this danger and should be kept in mind if we talk about a European Customer Satisfaction Index.

Customized instruments, on the other hand, could be designed in a way that reflects special aspects of a business (its strategic uniqueness) and thus be more valid than standardized measures. They could thus inspire bench-breaking rather than benchmarking, innovation rather than collusion. Their disadvantages are that development costs must be internalized, that they are less appropriate for external reporting and possibly more open to self-delusion.

To capture the advantages of both procedures we will probably have to develop measures that are standardized for special groups (defined by industry, size, stage within industry cycle, etc.).

A third dialogue focuses on the question of financial versus non-financial measures. Non-financial measures can be qualitative indicators (so-called 'soft facts') which are believed to translate into financial measures in future periods. A narrow focus on financial measures has been criticized as being backward-oriented, to miss tracing key success factors and supporting unhealthy short-term orientation (Edvinsson & Malone, 1997, 8ff). To express intangibles by non-financial measures is thus regarded as a necessary counterweight against 'bean-counters' with a lack in entrepreneurial attitudes.

This argument has its merits. Without qualitative measures of human skills or customer relations their long-term building might always fall short of the expectation of immediate financial returns. On the other hand, there is no reason to 'maximize' R&D expenditure, training hours or customer contact and satisfaction *per se* if they do not result in higher earnings (cash flows) or higher company value. At the end of the day, qualitative measures need to be correlated with financial results, while time lags between cause and effects should be accounted for.

To conclude, we can state requirements for the measurement of intangibles in general, and customer satisfaction in particular. Table 2 summarizes our argument.

We can state that research and business-context criteria are partially contradictory and require trade-offs. Efficiency may counteract validity as well as reliability, manageability may require concessions as to representability, motivation for action may ask for inappropriate

Table 2. Criteria to measure intangibles

	Criteria	Definition	Example	Ease of fulfilling
Classical research criteria	Validity	A measure captures what we want to know, what it pretends to measure	Absenteeism as a measure of employee dissatisfaction	Double problem of construct validity and indicator validity; never 100%
	Reliability	Stability of measurement instruments; different measurements produce same results	Customer survey used by different affiliates at different points in time	' <i>Ceteris is rarely paribus</i> ' complex, real-time situations are no laboratories
	Objectivity	Distant and interest-free perspective of an observer	Intelligence tests designed and validated by 'scientific' procedure	To be interest-free is logically impossible but can be brought into line by sound methodology
Business context	Efficiency	Benefits of measure outweigh costs of their gathering, procession and diffusion	Skills measured as formal qualifications versus assessment in a 6-month trainee period	Easier for singular measures drawn from existing accounts and statistics; increasingly easy for secondary research (Internet); often in contradiction to validity and strategic relevance
	Usability	Measures are paid attention, interpreted in an intended way and translated into action	Managers give more feedback after a survey showing a corresponding deficit	Simplicity, plausibility, understandability, ease to access
	Strategic relevance	Measures are important to monitor strategy and inspire the development of new strategy	Innovativeness measured as new products-to-sales ratio	Often driven out by the easy to quantify and a 'squirrel' mentality of accumulating details

Table 3. Trade-offs in measuring Intellectual Capital (IC)

Criterion	Mark, degree	Consequences
Cost of establishment and administration	↓	Standardized measures, by-product of traditional accounting/controlling procedures
Usability of information, value added by information	↑	Simple, few and clear measures, correlation to financial measures
Validity	↑	Sophisticated procedures, rather customized than standardized; qualitative, non-financial measures

generalizations or exaggerations of measures. Trade-offs must be guided by the purpose of measurement, which can only be decided by an organization's strategists (Table 3).

Cost, liability, accessibility

Many projects in the realm of measuring intangibles are shared endeavours between the practical field and research institutions (cf. OECD, 1999, Technical Meeting, June): public money is provided to kick-off the basic idea of developing new measures and to learn about the viability of tools. Does this constitute a necessary contribution to fundamental research or a market-distorting subsidy? What about liability for information? In most cases it is impossible to establish a clear causal link between a certain piece of information and harm done to its recipients. Therefore, we do not have liability legislation for research results (imagine a world in which researchers were as liable as plumbers or pharmacists), but liability would become an issue if information was sold by private providers. It is also an issue for boards in relation to their investors. This constitutes one of the reasons for their reluctance to accept mandatory external reporting on insecure intangibles. Knowledge is an enabling not an enforcing factor. Owing to its more or less tacit components it is open to interpretation that makes its regulation difficult. Still, we will need some rules to prevent us from fraud and insider trading. Accessibility of information is closely linked to the question of who bears the cost to generate it. If funding is public, we should rather assume that a public good has been produced. Information, as a common good, does not usually constitute competitive advantage. Furthermore, evaluation procedures within bureaucratic and scientific communities are generally inappropriate to deal with cost-efficiency. Therefore, the idea of establishing markets to allocate costs and benefits of information generation is appealing. Companies, groups of companies, could arrange and pay for external measurement or internalize the endeavour. They will do so if they perceive higher benefit than costs. But, as said before, costs and benefits are insecure, spread over time and influenced by the fact, that '*ceteris is never paribus*' in real-life management. This adds up to the trivial observation that whoever develops measurement tools will and can only do so if there is a consensual belief that tools will pay off, if not now than in the future when all current problems will be solved: Do we capture the right information rightly, is this information used properly? Again, we are confronted with a question of quasi-religious nature which cannot be decided scientifically.

A whole variety of research methods

We have used surveys as a blueprint for our critical discussion of measurement methods because they are most common in customer satisfaction research (cf. Bearden *et al.*, 1996).

Meanwhile, we are well aware that various techniques have been developed to render verbal reactions to eliciting questions more valid: the critical incident technique and various checks to elicit (in)consistent answers to similar questions can detect obvious lies but not a person's cultural and social 'distortions', or his or her self-deception. Interviews over the phone, as used for national customer satisfaction indices, will even miss non-verbal clues and need to be restricted to unambiguous simple statements. Those statements deliver information that is not very valid economically because it is too general to create much value. Our main argument against aggregate information from questionnaire-type instruments stems, though, from its underlying assumptions of macrostability and average outcomes. It could well be that the 'post X', net-oriented generation produces non-average outcomes more often than not. It has also been said that this generation neither takes the burden to argue nor to develop opinions (OE1, 18 August 1999, results of youth study, cf. Sennett, 1998). They will vote with their feet, or keyboards, rather than get involved. Furthermore, any kind of interview constitutes unavoidably an intervention into an ongoing relationship. We understand the price of any non-standardized good, such as a service, to be an expression of an ongoing relationship rather than a fixed datum. To ask for a customer's satisfaction and preferences changes this relationship and could well be used to direct his/her attention towards those dimensions a company (or an industry) is good at. Measurement changes its function completely if understood as a marketing, i.e. communication, tool. All the shortcomings discussed above will weigh much less then, as we do not measure any existing reality but a reality that is co-created by our endeavour.

Validity of measures: A knowledge perspective

If we want to discuss validity in a broader epistemological context we can refer to Polanyi's concept of tacit knowledge (Polanyi, 1967) or Boisot's distinction between codified and unmodified knowledge (Boisot, 1984).

Tacit knowledge (or implicit knowledge in the notation by Nonaka & Takeuchi, 1995) is not the opposite of explicit knowledge as many knowledge management theorists wrongly perceive (Davenport & Prusak, 1998; Edvinsson & Malone, 1997; Schmitz & Zucker, 1996), but its qualifying context; so to say the hidden part of the iceberg. To make explicit means to divest the tip of the iceberg of its foot for the sake of economizing the process of diffusion of knowledge. Explicit knowledge can be passed on as a structure of 'essentials', but without its qualifying context. The ultimate explicit form is codification. To understand the difference, let us consider a face-to-face situation of a salesperson and a customer. This situation is completely uncoded. It contains the complexity and richness of all the clues present in the environment and the person of the customer. Most certainly the salesperson will not be able to deal with this complexity, and miss many of the clues. If he/she is given a form to capture part of those clues, which are then codified, his/her colleagues will receive quite a different type of information. Customer X might be described as '1-3-2-3', where 1 stands for female, 3 for age group over 50, 2 for first time customer and 3 for her overall ranking of satisfaction on a five-point Likert scale.

We can see quite clearly that recipients of such codes can hardly reconstitute the complexity and information-richness of the original situation of measurement. Codes do not function like holograms. They are valid only if recipients/users have a shared understanding (pointing to a degree of homogeneity which is harmful to innovation of any business) and if they apply to a context that is very similar to their original context. This, of course, is a general problem of modern society. We trade off richness of information against its diffusability and

work on the grounds of the assumption that the basic structure, transported by the code, is sufficient for any further use of the knowledge³ in question.

As mentioned earlier, from the perspective of interpretative theories and complex-adaptive theory this assumption is questionable. This is no argument against measurement, but it is an argument against confusing the map with the territory and against allocating the majority of attention, actions taken and financial means to the restricted information we can gain from such measurement.

Conclusion

What conclusions can be drawn from our analysis? On the one hand, we must rely on reduction, isolation and codification to know and to pass on and use knowledge in a society based on the division of labour and knowledge. This is a strong point in favour of learning about intangibles by trying to measure them in a traditional manner.

On the other hand, we have seen that measures, as forcibly isolated codes, tend to be misunderstood, underused or misused in business organizations. They might be wasteful or even do harm by leading to inadequate decision-making. In our introduction we promised advice on how to avoid the traps of a new and inappropriate management fad. Our contribution so far has been to raise awareness of the difficulties involved, as means for research are a scarce resource in the academic as well as the practical field and all parties have to be very careful about dedicating funds to ambiguous purposes. We suggest more discussion of underlying assumptions so that both fields can be prevented from running into dead ends. In the following we shall formulate some hints with appropriate caution:

What then can we do?

Firstly, we should see that the procedure of measurement in the case of customer satisfaction is no neutral act, but an intervention. Whenever we design a method to learn about customer relations, we should do it with the alterations in mind that we would like to happen. We thus do not just want to know how we performed in the past, we want to create the future. As we cannot not influence our customers,⁴ while we survey them, we should rather influence them according to our vision and business strategy.

Secondly, we should never forget the core task to *create* intangibles over the effort to control them. If educational reform leads to teachers being distracted from their core task, namely to inspire young people to learn, by the need to document and legitimate each step they take (as is perceived by teachers in the UK),⁵ the quality of the core task may even deteriorate. Managing customer relations must therefore first and foremost mean to design and cherish those relations, not just to measure and document them.

Thirdly, we should be prudent and use our measures as sticks and carrots in incentive systems designed according to a simplistic stimulus–response–based theory of behaviour. If we consider all the caveats that are especially related to validity, we cannot take responsibility for such use. However, measures can be useful yardsticks within learning processes.

Finally, we should keep in mind that the starting point of all statistical procedures is the endless repetition of the same event. In a simultaneous setting this endless repetition must be substituted by the sameness of behaviour of a large number of elements. In order to make those methods applicable, we need standardization and homogenization. This seems to be quite the opposite of what is recommended to businesses and individuals in a new economy?

Let us finish with a half joking provocation: What do we get when we calculate averages? Could it be mediocrity?

Notes

1. As real decision-makers do not have complete information on all possible alternatives, we cannot use the notions minimize, maximize or optimize in their strict mathematical sense. They are meant here as 'to decrease as much as possible under the conditions given'.
2. Rationality, if looked at more closely, is a disputable concept. Here we refer to purposeful rationality as defined in the (neo)classical model of decision-making.
3. Information and knowledge are used interchangeably in this paper. This—although happening quite frequently—is not appropriate in other contexts of argumentation. For the distinction see Schneider (1996).
4. A variation of a famous dictum by Watzlawick *et al.* (1967).
5. Personal statement of teachers at the European Forum, Alpbach, 1997.

References

- ARGYRIS, C. & SCHÖN, D. (1996) *Organizational Learning II, Theory, Method and Practice* (Reading, MA, Addison-Wesley).
- ARTHUR, W.B. (1996) Increasing returns and the new world of business, *Harvard Business Review*, July–August, 1996, pp. 100–109.
- BEARDEN, W.O., NETEMEYER, R.G. & MOBLEY, M.F. (1996) *Handbook of Marketing Sales* (Newbury Park, SAGE Publications).
- BOISOT, M. (1994) *Information and Organizations. The Manager as Anthropologist* (New York, Harper Collins).
- BROWN, M.G. (1997) *Kennzahlen. Harte und weiche Faktoren Erkennen, Messen und Bewerten* (München, Wien).
- BRUHN, M. & GRUND, M.A. (1999) Theory, development and implementation of national customer satisfaction indices—the Swiss index of customer satisfaction, *Conference, Customer Satisfaction: Theory and Measurement*, Vienna, May.
- CHIA, R. (1998) From complexity science to complex thinking: organization as simple location. In: *Organization*, Vol. 5 (Newbury Park, SAGE Publications), pp 341–369.
- CREEMERS, R. (1999) Industry watcher, 27 January and 30 March 1999, robc@inter.NL.net.
- DAVENPORT, TH.D. & PRUSAK, L. (1998) *Working Knowledge* (Boston, MA, Harvard Business School Press).
- DÖRNER, D. (1989) *Die Logik des Miölingens* (Hamburg, Reinbek).
- DRUCKER, P. (1993) *Postcapitalist Society* (New York, Harper Business).
- EDYINSON, L. & MALONE, D. (1997) *Intellectual Capital* (New York, Harper Business).
- FEYERABEND, P. (1976) *Wider den Methodenzwang* (Frankfurt/Main, Suhrkamp).
- HALL, E.T. and HALL, M.R. (1990) *Understanding Cultural Differences*, (Yarmouth ME, Intercultural Press).
- HAYES, B. (1992) *Measuring Customer Satisfaction* (Milwaukee, Wisconsin, ASQC Quality Press).
- HORIBE, F. (1999) *Managing Knowledge Workers. New Skills and Attitudes to Unlock the Intellectual Capital in your Organization* (Toronto, John Wiley).
- KAPLAN, R.S. & NORTON, D.P. (1996) *The Balanced ScoreCard* (Boston, MA, Harvard Business School Press).
- KIESER, A. & KUBICEK, H. (1992) *Organization*, 3rd edition (Berlin, de Gruyter).
- LASKE *et al.* (1983) *Mitbestimmung in Tirol*, Band 2 (Innsbruck, Arbeiterkammer Tirol) (Ed.).
- LUHMANN, N. (1996) *Social Systems* (Stanford CA, Stanford University Press).
- MATURANA, H. & VARELA, F. (1990) *Der Baum der Erkenntnis* (München, Goldmann).
- NONAKA, I. & TAKEUCHI, H. (1995) *The Knowledge Creating Company* (Oxford, Oxford University Press).
- OECD (1999) <http://www.oecd.org//dsti/sti/industry/indcomp/act/ams-conf/symposium.htm>, BORNEMANN, M., KNAPP, A., SCHNEIDER, U. & SIXL, K. (1999) Holistic measurement of intellectual capital, *International Symposium on Measuring and Reporting Intellectual Capital: Experience, Issues and Prospects in Amsterdam* (Amsterdam).
- PASCALÉ, R. (1991) *Managing on the Edge*, How the smartest companies use conflict to stay ahead (New York, Simon & Schuster).
- PICOT, A. (1994) Einleitungsreferat zum 4, *Internationalen Symposium für Informationswirtschaft: Mehrwert von Information—Professionalisierung der Informationsarbeit*, Graz.
- POLANYI, M. (1967) *The Tacit Dimension* (London, Routledge & Kegan Paul).
- SCHMITZ, CHR. & ZUCKER, B. (1996) *Wissen gewinnt. Knowledge-Flow-Management* (Düsseldorf and München, Metropolitan).
- SCHNEIDER, U. (1996) *Wissensmanagement. Die Aktualisierung des intellektuellen Kapitals* (Frankfurt, FAZ).
- SCHNEIDER, U. (1998) The Austrian approach to the measurement of intellectual potential, *National Business Conference*, Hamilton, Canada, January.
- SCHNEIDER, U. (2000) *Wissensmanagement. The Process of Globalization* (Graz, Skriptum).

- SCHNEIDER, U. (1990) *Kulturbewusstes Informationsmanagement* (München and Wien, R. Oldenbourg).
- SCHREYÖGG, G. (1996) *Organisation. Grundlagen moderner Organisationsgestaltung* (Wiesbaden, Gabler).
- SENNETT, R. (1998) *The corrosion of character: The personal consequences of work in the New Capitalism*, (New York, Norton).
- SPENCER-BROWN, G. (1997) *Gesetze der Form* (Lübeck, Bohmeier).
- SVEIBY, K. (1997) *The New Organizational Wealth. Managing & Measuring Knowledge-based Assets* (San Francisco, Berrett-Koehler).
- WATZLAWICK, P., BEAVIN, J.H. & JACKSON, D.D. (1967) *Pragmatics of Human Communication* (New York, Menschliche Kommunikation, Berlin 1974).
- WEICK (1985) *Der Prozeß des Organisierens* (Frankfurt).
- WOLLIN, A.S. (1999) Deterministic chaos or indeterministic order: the underlying simplicity in theories of complexity, *International Workshop on Self-organisation, Evolutionary Economics and Innovation*, UQ, Brisbane, 12–15 July (submitted).
- WOLLIN, A.S. (2000) Simplicity in theories of complexity: Defining, knowing and doing, in FOSTER, J. and METCALFE, J. (Eds) (2000) *Frontiers of Evolutionary Economics: Competition, self-organisation and innovation policy* (Cheltenham: Edward Elgar).