Journal of Applied Statistics
Vol. 33, No. 4, 405–415, May 2006

Routledge
Taylor & Francis Group

On the Planning and Design of Sample Surveys

RON S. KENETT

KPA Ltd, Raanana, Israel and University of Torino, Torino, Italy

ABSTRACT *Surveys rely on structured questions used to map out reality, using sample observations from a population frame, into data that can be statistically analyzed. This paper focuses on the planning and design of surveys, making a distinction between individual surveys, household surveys and establishment surveys. Knowledge from cognitive science is used to provide guidelines on questionnaire design. Non-standard, but simple, statistical methods are described for analyzing survey results. The paper is based on experience gained by conducting over 150 customer satisfaction surveys in Europe, America and the Far East.*

KEY WORDS: Questionnaire design, cognitive science, individual surveys, household surveys, establishment surveys, control charts analysis of survey data

Introduction to Sample Surveys

Statistical analysis is a science that relies on a transformation of reality into dimensions that lend themselves to quantitative analysis. Self-administered surveys use structured questioning to map out perceptions, using samples from a population frame, into data that can be statistically analyzed. In drawing a sample, several sampling schemes can be applied. They range from probability samples such as cluster, stratified, systematic or simple random sampling, to non-probably samples such as quota, convenience, judgment or snowball sampling.

The survey process consists of four main stages: (1) Planning; (2) Collecting; (3) Analyzing; and (4) Presenting. Modern surveys are conducted in a wide variety of techniques including phone interviews, self-report paper questionnaires, email questionnaires, internet-based surveys, SMS-based surveys, face-to-face interviews, and videoconferencing (Kaplan *et al*., 2003).

In this paper we focus on the planning and analysis of surveys and begin with a discussion of various aspects of questionnaire design. We point out differences between employee surveys, customer satisfaction surveys, household surveys and establishment surveys. The distinction between these types of surveys strongly affects the planning and design stage. In many cases, practitioners neglect to account for such differences, creating problems in the data that even sophisticated data analysis cannot overcome. We begin with a general discussion on questionnaire design.

Correspondence Address: Ron S. Kenett, University of Torino, Torino, Italy. Email: ron@kpa.co.il

0266-4763 Print/1360-0532 Online/06/040405–11
DOI: 10.1080/02664760500448974

Questionnaire Design

Self-reports are the primary source of data in customer satisfaction surveys, life-style surveys, employee surveys such as 360 degree surveys, etc. Market researchers rely on the answers that customers provide to learn about individuals' satisfaction levels, loyalty and attitudes towards products, services and organizations. Human Resource managers use employee surveys to assess organizational climate and conduct performance appraisal evaluations. Subjective reports however are a potentially unsafe source of data, and changes in question wording, question format, or question context can result in changes in the obtained results, as a few examples from social science research illustrate:

- When asked what they consider 'the most important thing for children, to prepare them for life,' 61.5% of a representative sample chose the alternative '*to think for themselves*' when this alternative was offered on a list. Yet, only 4.6% volunteered an answer that could be assigned to this category when no list was presented (Schuman & Presser, 1981).
- When asked how successful they have been in life, 34% of a representative sample reported high success when the numeric values of the rating scale ranged from −5 to +5, whereas only 13% did so when the numeric values ranged from 0 to 10 (Schwarz *et al*., 1991).
- When asked how often they experience a variety of physical symptoms, 62% of a sample of psychosomatic patients reported symptom frequencies of more than twice a month when the response scale ranged from '*twice a month or less*' to '*several times a day*.' Yet, only 39% reported frequencies of more than twice a month when the scale ranged from '*never*' to '*more than twice a month*' (Schwarz & Scheuring, 1992).
- Whether we conclude that marital satisfaction is a major or a minor contributor to general life-satisfaction depends on the order in which both questions are asked, with correlations ranging from 0.18 to 0.67 as a function of question order and introduction (Schwarz *et al*., 1991).

The underlying cognitive and communicative processes causing such effects are systematic and increasingly well-understood. Since the early 1980s, psychologists and survey methodologists developed an interdisciplinary field of research that is devoted to understanding the nature of self-reports and to improving the quality of data collection. Research in this field has addressed a wide range of topics: How do respondents make sense of the questions asked of them? What is the role of autobiographical memory in retrospective reports of behaviors and how can we increase the accuracy of these reports? What are the judgmental processes underlying the emergence of context effects in attitude measurement? Do the processes underlying self-reports of behaviors and attitudes change across the adult life span? Which techniques can we use to determine if a question 'works' as intended?

This area of research is typically referred to as 'cognitive aspects of survey methodology,' the cognitive and communicative processes investigated apply to the question-answering process in all standardized research situations.

Not surprisingly, the first task that respondents face is to understand the question asked of them (Stack & Martin, 1987; Tourangeau, 1984). The key issue is whether the respondent's understanding of the question matches what the survey designer had in mind: Is the attitude object, or the behavior, that the respondent identifies as the referent of the question

the one that the researcher intended? Does the respondent's understanding tap the same facet of the issue and the same evaluative dimension?

From a psychological point of view, question comprehension reflects the operation of two intertwined processes of semantic and contextual understanding (Clark & Schober, 1992).

Comprehending the literal meaning of a sentence involves the identification of words, the recall of lexical information from semantic memory, and the construction of a meaning of the utterance, which is constrained by its context (Anderson, 1980). Not surprisingly, textbooks urge researchers to write simple questions and to avoid unfamiliar or ambiguous terms. However, understanding the words is not sufficient to answer a question. For example, if respondents are asked, '*What have you done today?*' they are likely to understand the meaning of the words. Yet, they still need to determine what kind of activities the researcher is interested in. Should they report, for example, that they had breakfast, or not? Hence, understanding a question in a way that allows an appropriate answer requires not only an understanding of the literal meaning of the question but also involves inferences about the questioner's intention to determine the pragmatic meaning of the question.

To infer the pragmatic meaning of a question, respondents rely on the tacit assumptions that govern the conduct of conversation in everyday life. These tacit assumptions were described by Paul Grice (1975), a philosopher of language. According to Grice's analysis, conversations proceed according to a cooperativeness principle, which can be expressed in the form of four maxims.

First, a *maxim of relation* enjoins speakers to make their contribution relevant to the aims of the ongoing conversation. In research situations, this maxim licenses the use of contextual information in question interpretation and invites respondents to relate the question to the context of the ongoing exchange.

Second, a *maxim of quantity* enjoins speakers to make their contribution as informative as required, but not more informative than required. This maxim invites respondents to provide information the questioner seems interested in, rather than other information that may come to mind. Moreover, it discourages the reiteration of information that has already been provided earlier, or that 'goes without saying' (such as 'having breakfast' in the above example).

Third, a *maxim of manner* holds that the contribution should be clear rather than obscure, ambiguous, or wordy. In research situations, this maxim entails an 'interpretability presumption': research participants can assume that the researcher 'chose his wording so they can understand what he meant – and can do so quickly' (Clark & Schober, 1992). Hence, the most obvious meaning seems likely to be the correct one, and if there is no obvious meaning, respondents may consult the immediate context to determine one. As numerous studies have shown, the researcher's contributions include formal aspects of questionnaire design, such as the response alternatives provided as part of the question, and respondents draw on these features in interpreting the question.

Finally, a *maxim of quality* enjoins speakers not to say anything they believe to be false or lack adequate evidence for (Bless *et al*., 1993).

In summary, speakers should try to be informative, truthful, relevant, and clear, and listeners interpret the speakers' utterances 'on the assumption that they are trying *to* live up to these ideals' (Clark & Clark, 1977). These rules of cooperative conversational conduct are essential for understanding how respondents make sense of the questions asked of them, as the following examples illustrate (Clark & Schober, 1992). In the next sections we consider the implication of various response alternatives to questions typically posed in the context of a customer satisfaction survey.

Open versus Closed Response Formats

Suppose that respondents are asked in an open-response format, '*What have you done today?*' To give a meaningful answer, respondents have to determine which activities may be of interest to the researcher. In an attempt to be informative, respondents are likely to omit activities that the researcher is obviously aware of (e.g., '*I gave a survey interview*') or may take for granted anyway (e.g., '*I had breakfast*'), thus observing the maxim of quantity. If respondents were given a list of activities that included giving an interview and having breakfast, most respondents would endorse them. At the same time, however, such a list would reduce the likelihood that respondents would report activities that are not represented on the list (Schuman & Presser, 1981). Both of these question – form effects reflect that response alternatives can clarify the intended meaning of a question, in the present example by specifying the activities the researcher is interested in. In addition, response alternatives may remind respondents of material that they may otherwise not consider.

Frequency Scales and Reference Periods

Suppose that respondents are asked how frequently they complain to a call center service. To provide an informative answer, respondents have to determine what the researcher means by 'complain'. Does this term refer to major or to minor annoyances? To identify the intended meaning of the question, they may consult the response alternatives provided by the researcher. If the response alternatives present low-frequency categories, for example, ranging from '*less than once a year*' to '*more than once a month*,' respondents may conclude that the researcher has relatively rare events in mind. Hence, the question cannot refer to minor complaints that are likely to occur more often, so the researcher is probably interested in more severe complaints. In line with this assumption, Schwarz *et al*. (1988) observed that respondents who had to report the frequency of irritating experiences on a low-frequency scale assumed that the question referred to major annoyances, whereas respondents who had to give their report on a high-frequency scale assumed that the question referred to minor annoyances. Thus, respondents identified different experiences as the target of the question, depending on the frequency range of the response alternatives provided to them.

Similarly, Winkielman *et al*. (1998) observed that the length of the reference period can profoundly affect question interpretation. In their studies, respondents were asked how frequently they had been angry either '*last week*' or '*last year*.' Again, they inferred that the researcher is interested in more frequent and less severe episodes of anger when the question pertained to one week rather than one year, and their examples reflected this differential question interpretation. These findings have important implications for the comparison of concurrent and retrospective reports of behaviors and emotions. Winkielman *et al*.'s (1998) results suggest that discrepancies between concurrent and retrospective reports may in part be due to differential question interpretation: concurrent reports necessarily pertain to a short reference period, with one day typically being the upper limit, whereas retrospective reports cover more extended periods. Hence, the concurrent and retrospective nature of the report is inherently confounded with the length of the reference period. Accordingly, participants who provide a concurrent report may infer from the short reference period used that the researcher is interested in frequent events, whereas the long reference period used under retrospective conditions may suggest an interest in infrequent events. Hence, respondents may deliberately report on different experiences, rendering their reports incomparable.

Rating Scales

Similar considerations apply to researchers' favorite question format, the rating scale. Suppose respondents are asked, '*How successful would you say you have been in life?*' accompanied by a rating scale that ranges from '*not at all successful*' to '*extremely successful*.' To answer this question, respondents have to determine what the researcher means by 'not at all successful': Does this term refer to the absence of outstanding achievements or to the presence of explicit failures? To do so, they may draw on what is supposedly a purely formal feature of the rating scale, namely its numeric values. As mentioned earlier, Schwarz *et al*. (1991) presented the success-in-life question with an 11-point rating scale that ranged either from 0 (not at all successful) to 10 (extremely successful), or from −5 (nor at all successful) to 5 (extremely successful). The results showed a dramatic impact of the numeric values presented to respondents. Whereas 34% of the respondents endorsed a value between −5 and 0 on the −5 to +5 scale, only 13% endorsed one of the formally equivalent values between 0 and +5 on the 0–10 scale. Subsequent experiments indicated that this difference reflects differential interpretations of the term '*not at all successful*.' When this label was combined with the numeric value '0,' respondents interpreted it to reflect the absence of outstanding achievements. However, when the same label was combined with the numeric value '−5,' and the scale offered '0' as the midpoint, they interpreted it to reflect the presence of explicit failures (see Schwarz *et al*., 1998). In general, a format that ranges from negative to positive numbers conveys that the researcher has a bipolar dimension in mind, where the two poles refer to the presence of opposite attributes. In contrast, a format that uses only positive numbers conveys that the researcher has a unipolar dimension in mind, referring to different degrees of the same attribute.

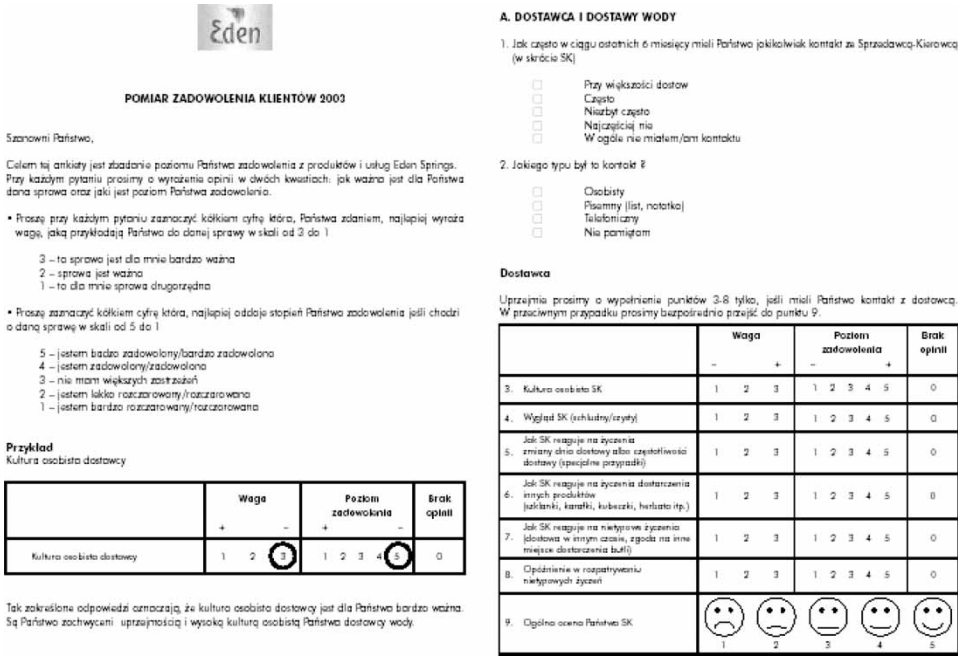
We find that, with an appropriate contextual setting, a 1–5 scale with anchoring from '*very low*' to '*very high*' is very effective in capturing customer satisfaction levels and employee satisfaction levels. Such a scale is commonly used and is therefore familiar to most respondents. We conducted several experiments using alternative 1–4 and 1–10 scales in parallel to questionnaires with 1–5 scales, which reinforced our preference for the 1–5 scale.

In written surveys we also typically include a direct evaluation of importance. In determining 'level of importance' our common practice is to refer to a 1–3 scale which permits respondents to state importance (1: not important, 2: average importance, 3: very important) in a simple and user friendly way. The combination of satisfaction level with importance allows us to classify the various topics probed by the questionnaire into four categories:

1. Topics of excellence to leverage: topics with high satisfaction levels and high importance.
2. Strengths to preserve: topics with high satisfaction levels but low importance.
3. Immediate Actions: topics with low levels of satisfaction and high importance.
4. Second priority issues: topics with low levels of importance and low importance.

An example of a questionnaire from a customer satisfaction survey we recently conducted for Eden Springs Poland, a Home Office Delivery firm specializing in mineral water distribution is presented in Figure 1.

The most important lessons that emerge from survey design research are rather general in nature. Researchers tend to view questionnaires as 'measurement devices' that elicit information from respondents. What is frequently overlooked is that questionnaires are



Eden

POMIAR ZADOWOLENIA KLIENTÓW 2003

Szanowni Państwo,

Celem tej ankiety jest zbadanie poziomu Państwa zadowolenia z produktów i usług Eden Springs. Przy każdym pytaniu prosimy o wyrażenia opinii w dwóch kwestiach: jak ważna jest dla Państwa dana sprawa oraz jaki jest poziom Państwa zadowolenia.

• Proszę przy każdym pytaniu zaznaczyć kółkiem cyfrę która, Państwa zdaniem, najlepiej wyraża wagę, jaką przykładają Państwo do danej sprawy w skali od 3 do 1

3 – to sprawa jest dla mnie bardzo ważna
2 – sprawa jest ważna
1 – to dla mnie sprawa drugorzędna

• Proszę zaznaczyć kółkiem cyfrę która, najlepiej oddaje stopień Państwa zadowolenia jeśli chodzi o daną sprawę w skali od 5 do 1

5 – jestem bardzo zadowolony/bardzo zadowolona
4 – jestem zadowolony/zadowolona
3 – nie mam większych zastrzeżeń
2 – jestem lekko rozczarowany/rozczarowana
1 – jestem bardzo rozczarowany/rozczarowana

Przykład
Kultura osobista dostawcy

	Waga + –			Poziom zadowolenia + –					Brak opinii
Kultura osobista dostawcy	1	2	(3)	1	2	3	4	(5)	0

Tak zakreślone odpowiedzi oznaczają, że kultura osobista dostawcy jest dla Państwa bardzo ważna. Są Państwo zachwyceni uprzejmością i wysoką kulturą osobistą Państwa dostawcy wody.

A. DOSTAWCA I DOSTAWY WODY

1. Jak często w ciągu ostatnich 6 miesięcy miali Państwo jakikolwiek kontakt ze Sprzedawcą-Kierowcą (w skrócie SK)

☐ Przy większości dostaw
☐ Często
☐ Niezbyt często
☐ Najczęściej nie
☐ W ogóle nie miałem/am kontaktu

2. Jakiego typu był to kontakt ?

☐ Osobisty
☐ Pisemny (list, notatka)
☐ Telefoniczny
☐ Nie pamiętam

Dostawca

Uprzejmie prosimy o wypełnienie punktów 3-8 tylko, jeśli miali Państwo kontakt z dostawcą. W przeciwnym przypadku prosimy bezpośrednio przejść do punktu 9.

	Waga –		+	Poziom zadowolenia –				+	Brak opinii
3. Kultura osobista SK	1	2	3	1	2	3	4	5	0
4. Wygląd SK (schludny/czysty)	1	2	3	1	2	3	4	5	0
5. Jak SK reaguje na życzenia zmiany dnia dostawy albo częstotliwości dostawy (specjalne przypadki)	1	2	3	1	2	3	4	5	0
6. Jak SK reaguje na życzenia dostarczenia innych produktów (talerzyki, kieszki, kubeczki, herbata itp.)	1	2	3	1	2	3	4	5	0
7. Jak SK reaguje na nietypowe życzenia (dostawa w innym czasie, zgoda na inne miejsce dostarczenia butli)	1	2	3	1	2	3	4	5	0
8. Opóźnienie w rozpatrywaniu nietypowych życzeń	1	2	3	1	2	3	4	5	0
9. Ogólna ocena Państwa SK	☹ 1	☹ 2	😐 3	🙂 4	☺ 5				

Figure 1. Typical customer satisfaction survey questionnaire with satisfaction levels on a 1–5 scale and importance levels on a 1–3 scale

also a source of information that respondents draw on in order to determine their task and to arrive at a useful and informative answer. Respondents do their best to be cooperative communicators. Consistent with the assumptions that underlie the conduct of conversation in daily life, they assume that all contributions of the researcher are relevant to the goals of the ongoing exchange, and they take these contributions into account in arriving at an answer. Researchers need to be fully aware of the information that questionnaires provide, and the extent to which the questions asked determine the answers received.

Individual Surveys, Household Surveys and Establishment Survey

The main distinction between individual surveys and household and establishment surveys is in the level of information that is required to be captured. In order to capture such information, different processes apply that further distinguish between these types of surveys.

In household and establishment surveys the respondent must answer for the household or the business and not for him or herself. Moreover:

1. Questionnaires are usually mailed to a family or business and not to a person in that business.
2. It may not be known who the specific respondent should be.
3. Respondents often need to consult records in order to determine what answers should be provided for specific questions.
4. Multiple people may need to provide information for answering a single questionnaire.
5. Surveys of businesses often entail a greater response burden than do household surveys.
6. Detailed definitions and rules for how to arrive at an answer are often necessary.

Questionnaires for individual, household or establishment surveys need to be designed and evaluated accordingly. The evaluation needs to account for the difference in scope. The primary goals of a questionnaire appraisal are to document problems in question design and identify possible revisions (Forsyth *et al*., 1999). Most survey item response models (Tourangeau, 1984), assume that *question comprehension* is an early step, followed by *information retrieval*, *judgment* and *response selection*. Assumptions about information retrieval emphasize the role that respondents play as managers of information. In customer satisfaction surveys of businesses or large organizations where the respondent is asked to fill up a questionnaire reflecting an organizational perspective, retrieval involves:

- Identifying appropriate sources of information – these may include the respondents' own memory, organizational record systems, and other people within the organization.
- Selecting strategies for retrieving information from identified sources.
- Assessing the match between information retrieved and information needed to select a response.

Respondents may repeat some or all or these retrieval processes until the match seems sufficient.

In the judgment step, the respondent's task is to synthesize or integrate information coming from a variety of sources. Depending on the sources consulted, respondents may need to synthesize information from different individuals within the organization, or they may need to synthesize information from memory with information from administrative records.

The HOD company mentioned above (Eden Springs Poland) has both individual accounts called 'Core Businesses' and major accounts called 'Key Accounts'. Core Businesses will typically have 1–3 water fountains at one location. For such customers there typically is one contact person who is responsible for all interactions with the company (ordering water, paying bills and calling customer service when needed). In Key Accounts there might be several hundred fountains dispersed in several locations. For these customers, the person getting the bill is located in a different department than the person in charge of maintenance who orders water replenishment. End users typically are a mixture of employees, at different levels of management, job functions and locations. For Key Accounts we are conducting special Key Account Surveys with the close involvement of the Key Account Managers at Eden Springs. The Key Account managers identified, within each Key Account, an appropriate Decision Maker who is responsible for contracts and payment issues, a Daily Contact who handles orders and maintenance, and representative End Users from the customer's various locations. Three types of questionnaires were designed and the survey involves distributing the questionnaires, through the Key Account manager, to specific individuals in each company. Basically there are three surveys conducted simultaneously. One survey of Decision Makers, a second survey of Daily Contacts and a third survey of End Users. Following the design of a questionnaire, a pilot run is carried out with typical respondents in order to validate the clarity of the questionnaire. The pilot study participants are debriefed for feedback on the questionnaire design and content. The final questionnaire is updated to account for the feedback and comments. Once the questionnaires are distributed, a follow up procedure is initiated to track response rates. In some cases, focused follow up activities are initiated to increase response rates in specific strata. At some point in time the survey is closed and the data analysis begins. The next section describes various aspects of customer satisfaction survey data analysis.

Questionnaire Analysis

Survey analysis typically involves a descriptive study of demographic parameters and satisfaction levels from a number of items. In many cases, surveys are conducted as a census of all customers or all employees, and sampling is due to non-response not to proactive sampling.

Possible bias in the returned questionnaires' demographics versus the population who received questionnaires can be evaluated using the M-Test (Fuchs & Kenett, 1980; Kenett, 1991; Kenett & Zacks, 1998). In simple terms the M-test consists of comparing the number of expected returns, by demographic strata, conditioned on the total number of returns. Standardizing the difference between actual and expected returns and using a Bonferroni bound determines if there are strata with under(over) representation.

Another effective questionnaire analysis technique involves a non-standard application of control charts to the analysis of satisfaction surveys data as described in Kenett & Zacks (1998) and Kenett (1998, 2002).

Control charts were developed at Bell Laboratories to control processes and identify significant changes in proportions of events over time. The underlying concept of the Control Chart for proportions – the p chart – is that observations that characterize a binary event follow a binomial distribution. In a Customer Satisfaction Survey we might want to analyze the responses that correspond to the top evaluation rating – a '5' on a 1–5 scale. These responses correspond to very satisfied customers that give the top rating in a list of questions on satisfaction levels from specific features and dimensions of services and product characteristics. A customer responding '5' to a question on overall satisfaction or his willingness to recommend the company providing the service to others is typically a loyal customer. A response of '4' indicates mere satisfaction and implies significantly reduced loyalty levels.

Figure 2 presents proportions of '5' to 26 questions. The average proportion is 15.95%. The upper and lower control limits, UCL and LCL, respectively, indicate 3 sigma limits above and below the average proportion. Any question with a proportion above or below the control limits indicates a statistically significant difference from the overall average. Practically speaking, it signals a feature or service dimension where the proportion of '5's is significantly different from the grand average. In the control chart

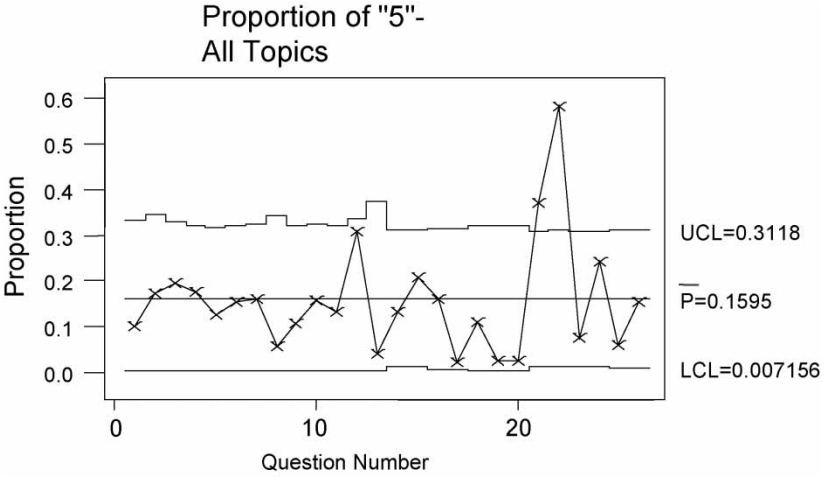


Figure 2. A control chart comparing percentage of '5' in various questions

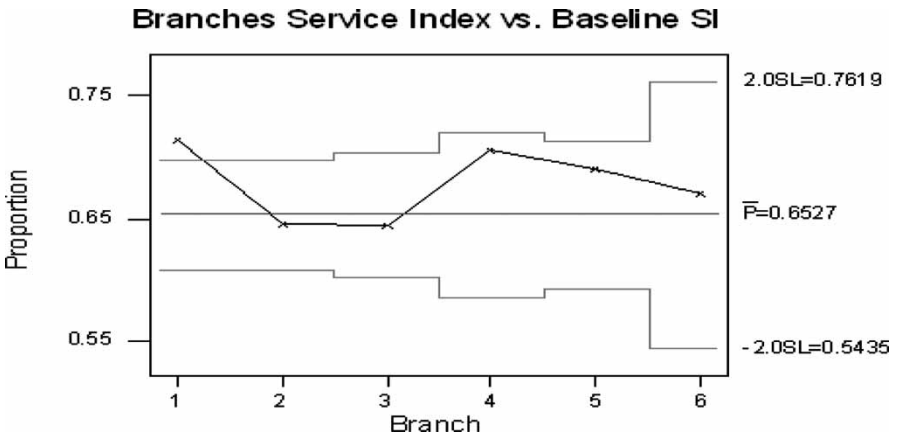


Figure 3. A control chart comparing service indices in various branches

described below, questions 21 and 22 have an unusually high proportion of '5's indicating outstanding strengths. All other questions provide responses consistent with the grand average – i.e. not worth pointing out as unusual or significantly high.

The varying number of responses, per question, is reflected by varying control limits. The more responses the narrower the control limits. The control limits delineate three regions: significantly above average, as expected and significantly below average.

Points above or below the control limits indicate proportions significantly higher, or lower, than the grand average. The probability of making such a statement when it should not be made (Type I error) is about $1/370$ – a rather low risk.

Control charts are used, for example, to classify HOD distribution branches using a Service Index that represents the percentage of '5's in a number of questions including overall satisfaction, recommendation and satisfaction from distributor. Plotting the Service Index on a control chart allows us to classify branches as outstanding (green), within expectations (yellow) and below expectations (red). In Figure 3, Branch 1 is classified as outstanding – all other branches performed within expectations.

A similar analysis can be performed for the topic of low satisfaction. Here the analysis focuses on proportions of '1's. The application of Control Charts to these proportions is identical to the example of proportions of '5's. Significantly high levels of '1's indicate severe weaknesses and opportunities for improvement.

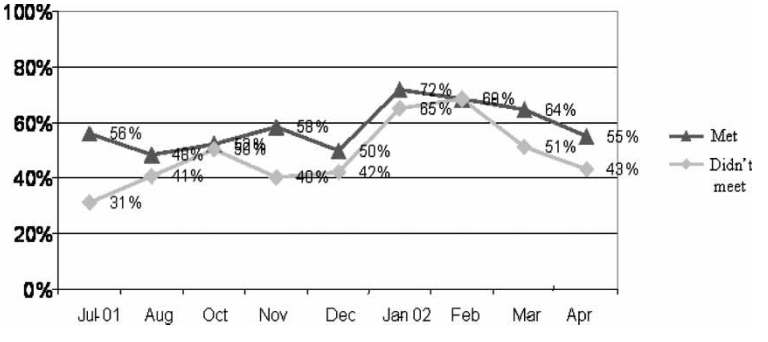


Figure 4. Overall satisfaction measured by percentage of '5', stratified by groups of customers

Tracking customer satisfaction survey data over time provides interesting insights that affect the day-to-day operation of an organization. For example, the HOD operator might want to determine if the distributors should make the effort to be in touch with their customers or not. Figure 4 presents overall satisfaction levels split by customers who were or were not in contact with the distributor as determined from operational data.

Clearly, meeting the distributor improves overall satisfaction by about 5%. If this increase represents a 5% increase in customer loyalty, one can actually determine the benefits of having the distributor stay in touch with the customer. This can be done by phone, notes or face-to-face meetings.

Summary and Conclusions

In order to design and analyze customer satisfaction, employee, household or establishment surveys effectively, three questions need to be checked:

1. Is the questionnaire properly designed?
2. Has the survey been properly conducted?
3. Has the data been properly analyzed?

The various types of surveys mentioned above require different answers. For example, an employee survey designed to assess organizational climate and provide feedback to management must both assure anonymity and provide exact details on what manager is evaluated, and this includes direct and indirect reports. Such surveys require detailed planning and coded organograms.

Developing solutions that yield positive responses to the checklist mentioned above requires knowledge and experience in statistics, cognitive psychology and business processes. Achieving such an integration is a non-trivial task. For example, running the Key Account survey at Eden Springs Poland involves close collaboration between marketing, sales, customer service and operations. The complexity of designing, conducting and analyzing surveys requires more than 'off the shelf' solutions such as the ones provided by commercially available products. Improperly designed surveys produce useless data that cannot be effectively analyzed. Organizations therefore need to develop their own approaches to specific surveys by evaluating survey methodologies used in the design, conduct and data analysis phases. This paper provides some ideas on how to do this.

Acknowledgment

This paper was supported by funding from the 'Growth' programme of the European Community and was prepared in collaboration by member organizations of the Thematic Network-Pro-ENBIS-EC contract number G6RT-CT-2001-05059.

References

Anderson, R. (1980) *Cognitive Psychology and its Implications* (San Francisco: Freeman).

Bless, H., Strack, F. & Schwarz, N. (1993) The informative functions of research procedures: Bias and the logic of conversation, *European Journal of Social Psychology*, 23, pp. 149–165.

Clark, H. H. & Clark, E. V. (1977) *Psychology and Language* (New York: Harcourt, Brace, Jovanovich).

Clark, H. H. & Schober, M. F. (1992) Asking questions and influencing answers, in: J. M. Tanur (Ed.) *Questions about Questions*, pp. 15–48 (New York: Russell Sage Foundation).

Forsyth, B., Levin, K. & Fisher, S. (1999) Test of an appraisal method for establishment survey questionnaire. *Proceedings of the Annual Meting of the American Statistical Association.*

Fuchs, C. & Kenett, R. (1980) A test for detecting outlying cells in the multinomial distribution and two-way contingency tables, *Journal of the American Statistical Association*, 75, pp. 395–398.

Grice, H. (1975) Logic and conversation, in: P. Cole & J. Morgan (Eds) *Syntax and Semantics*, *Vol 3, Speech Acts*, pp. 41–58 (New York: Academic Press).

Kaplan, O., Kenett, R. & Raanan, Y. (2003) Statistical properties of internet base market research surveys, *Proceedings of the Third Annual Conference of the European Network for Business and Industrial Statistics* (ENBIS), Barcelona, Spain.

Kenett, R. (1991) Two methods for comparing Pareto charts, *Journal of Quality Technology*, 23, pp. 27–31.

Kenett, R. (1998) Customer surveys: why and how, (in Hebrew) in: *Surveys: Some Good, Some Less*, S. Bar-Lev & C. Fuchs (Eds) (Hakibbutz Hameuchad Publishing House Ltd and Haifa University Press).

Kenett, R. (2002) *Issues in Customer Satisfaction Surveys*, DEINDE 2002, Torino, Italy.

Kenett, R. & Fainstein, R. (2003) Customer retention: the key to success in unstable economic times, (in Hebrew) in: O. Kaplan (Ed.) *Researchers Speak: Market Research and its application in the Israeli Market* (Globes Business Publishing).

Kenett, R. & Zacks, S. (1998) *Modern Industrial Statistics: Design and Control of Quality and Reliability* (San Francisco: Duxbury Press).

Kenett, R., Hacohen-Jacobs, M., Bantay, J. & Uriel, D. (1997) Is your employee a satisfied customer? Or how to incorporate employee satisfaction surveys in the human resources development strategic plan, (in Hebrew) *Proceedings of the 8th Annual Conference of the Israeli Human Resources Organization*, Zichron Ya'acov, Israel.

Kenett, R., Shade, J. & Ramahloto, M. (2003) Six sigma perspectives on stochastics for the quality movement, *Proceedings of the Third Annual Conference of the European Network for Business and Industrial Statistics* (ENBIS), Barcelona, Spain.

Schuman, H. & Presser, S. (1981) *Questions and Answers in Attitude Surveys* (New York: Academic Press).

Schwarz, N. & Scheuring, B. (1992) Frequency reports of psychosomatic symptoms: what respondents learn from response alternatives, (in German) *Zeitschrift für Klinische Psychologie*, 22, pp. 197–208.

Schwarz, N., Grayson, C. E. & Knauper, B. (1998) Formal features of rating scales and the interpretation of question meaning, *International Journal of Public Opinion Research*, 10, pp. 177–183.

Schwarz, N., Strack, F., Muller, G. & Chassein, B. (1988) The range of response alternatives may determine the meaning of the question: further evidence on informative functions of response alternatives, *Social Cognition*, 6, pp. 107–117.

Schwarz, N., Strack, F. & Mai, H. P. (1991) Assimilation and contrast effects in part-whole question sequences: a conversational logic analysis, *Public Opinion Quarterly*, 55, pp. 3–23.

Schwarz, N., Knauper, B., Rippler, H. J., Noelle-Neumann, E. & Clark, F. (1991). Rating scales: Numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55, pp. 570–582.

Strack, F. & Martin, L. (1987) Thinking, judging, and communicating: a process account of context effects in attitude surveys, in: H. Rippler, N. Schwarz & S. Sudman (Eds) *Social Information Processing and Survey Methodology*, pp. 123–148 (New York: Springer Verlag).

Tourangeau, R. (1984) Cognitive science and survey methods: a cognitive perspective, in: T. Jabine, M. Straf, J. Tanur & R. Tourangeau (Eds) *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*, pp. 73–100 (Washington, DC: National Academy Press).

Winkielman, P., Knauper, B. & Schwarz, N. (1998) Looking back at anger: reference periods change the interpretation of (emotion) frequency questions, *Journal of Personality and Social Psychology*, 75, pp. 719–728.