QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL
Qual. Reliab. Engng. Int. 2006; (in press)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/qre.767

Research

Hypothesis Generation in Quality Improvement Projects: Approaches for Exploratory Studies

Jeroen de Mast*,† and Marcus Bergman
Institute for Business and Industrial Statistics, University of Amsterdam (IBIS UvA), Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands

In quality improvement projects—such as Six Sigma projects—an exploratory phase can be discerned, during which possible causes, influence factors or variation sources are identified. In a subsequent confirmatory phase the effects of these possible causes are experimentally verified. Whereas the confirmatory phase is well understood, in both the statistical sciences and philosophy of science, the exploratory phase is poorly understood. This paper aims to provide a framework for the type of reasoning in the exploratory phase by reviewing relevant theories in philosophy of science, artificial intelligence and medical diagnosis. Furthermore, the paper provides a classification and description of approaches that could be followed for the identification of possible causes. Copyright © 2006 John Wiley & Sons, Ltd.

Received 30 June 2004; Revised 12 November 2005

KEY WORDS: discovery; problem-solving; abduction; clue generation; exploratory data analysis; diagnosis

1. INTRODUCTION

Quality improvement projects in industry aim for breakthroughs in quality and efficiency[1]. Several methodologies for quality improvement projects are popular, such as the Six Sigma method, the Shainin System and Taguchi's method (De Mast[2] discusses and compares these three methodologies). Such methodologies typically consist of a stepwise strategy to guide the project leader through an improvement project, plus a number of tools and techniques to help the project leader attain the intermediate results that the stepwise strategy prescribes.

Statistical improvement methodologies improve quality by exploiting cause-and-effect relations in the process. The 'effects' are quality characteristics, performance indicators, Critical To Quality relations (CTQs), Ys, etc.; we shall refer to them in this paper as CTQs, which is their name in the Six Sigma programme. The 'causes' are influence factors in the process, sources of variation, the causes of a problem, or simply the Xs.

De Mast[3] provides a framework for statistical quality improvement methodologies. It discerns five broad phases in improvement projects.

(1) Operationalization: definition of the problem and objectives.
(2) Exploration: generation of possible causes.
(3) Elaboration: explication and sorting of possible causes.

*Correspondence to: Jeroen de Mast, Institute for Business and Industrial Statistics, University of Amsterdam (IBIS UvA), Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands.
†E-mail: jdemast@science.uva.nl

(4) Confirmation: verification of the effects of causes.
(5) Conclusion: design of improvement actions based on the identified causes.

The core is formed by the alternation of the Exploration and the Confirmation phase. In the Exploration phase potential (conjectured) causes are identified. In the Confirmation phase the effects of the potential causes are experimentally verified. The steps in the Six Sigma's Breakthrough Cookbook can be related to these phases[3].

While the Confirmation phase—experimental verification of the effects of possible causes—has been thoroughly researched (cf. statistical theory of hypothesis testing, experimentation and statistical modelling) and is well understood, De Mast[2] concludes that the Exploration phase—the generation of ideas and possible causes—is poorly understood, and existing methodologies fall short in providing sufficient guidance.

The Shainin System gives perhaps the most elaborate guidance. The Exploration phase is called 'clue generation' in the Shainin System[4], and a number of statistical techniques are offered to the project leader, such as the multi-vari chart and pairwise comparison. Moreover, these techniques are placed in a heuristic that we shall name *progressive search through families of causes* later in this paper.

In Taguchi's method, the Exploration phase is limited to brainstorming and cause-and-effect diagrams[5] (CE diagrams). In Six Sigma's Breakthrough Cookbook potential causes are generated in step 6 (*Identify variation sources*). A vast collection of tools and techniques is offered, such as flowcharting, brainstorming, CE diagrams, run charts, control charts, multi-vari charts, etc. (Breyfogle[6], ch. 4, 5 and 15). These techniques are not, however, placed in a heuristic structure and the project leader is not given much guidance in their application.

Three examples of improvement projects that could be tackled with the procedure above could guide the reader in forming a notion of the theory.

Example 1 (Increasing flight times of paper helicopters). This example relates to a well-known classroom experiment, often used to teach statistical methods[7]. The idea is to discover how the flight time of helicopters made out of paper and other office supplies could be increased. The Exploration phase results in a list of factors (control factors such as wing length, paper thickness and tail length, but as well sources of variation such as launching height, the way in which helicopters are launched and measurement variation). The Confirmation phase aims to select by means of an experiment those factors that have a relevant effect on flight time, and to give a mathematical description of their relationship with flight time.

Example 2 (Reducing variability of moisture-percentages of coffee). To improve control of the moisture-percentage of ground and packaged coffee, variation should be decreased. During the Exploration phase, possible sources of variation could be generated. In the Confirmation phase regression analysis and analysis of variance could be used to establish which of these variation sources has the largest contribution to the total variability.

Example 3 (Solving the microphony problem). A potential problem with television tubes is that the image is distorted when the sound volume is increased (due to vibrations of certain components in the tube). Television tubes ought to be robust against this effect—called *microphony*—up to a certain extent. Over a limited period of time, the microphony performance of manufactured tubes deteriorated dramatically. In the Exploration phase the cause of this deterioration should be tracked down. The Confirmation phase should provide evidence that the cause has indeed been found.

It is the purpose of this article to present a framework for the Exploration phase in improvement projects, and to position existing tools and techniques within this framework. Both for practitioners and scientists it is important to have a crystallized formulation and demarcation of the Exploration phase and its principles. The account is intended to be prescriptive (outlining what project leaders ought to do) rather than descriptive (outlining what project leaders actually do).

The article is based on a review of the literature in relevant disciplines, such as quality management, industrial statistics, philosophy of science, artificial intelligence and medical diagnosis. Ideas in these various disciplines are combined into a coherent framework. For some instances, additional information is provided in footnotes, to keep the flow of the main text going.

DOI: 10.1002/qre

2. SOME BACKGROUND THEORY

Exploratory reasoning is studied in various disciplines under names such as problem-solving, discovery and diagnosis. This section presents a concise overview from the perspectives of philosophy of science, artificial intelligence and medical diagnosis.

2.1. Philosophy of science: discovery and abduction

Philosophy of science studies the principles of (amongst others) empirical inquiry. The objective of empirical inquiry could be described as figuring out how a system works, or—put differently—to develop an explanatory model for a system. An explanatory model enables the inquirer to predict how the system will respond to certain manipulations, and thus provides control in a certain sense. An explanatory model consists of specified causal relations among factors[‡]. In the context of improvement projects, an explanatory model specifies, in the form of transfer functions, causal relations between CTQs on the one hand, and causes, influence factors and variation sources on the other.

An explanatory model being the objective of inquiry, philosophy of science discerns its procedure into two processes: discovery and justification (Losee[12] gives an overview of the important role these concepts played in the history of philosophy of science). Justification—which corresponds to the Confirmation phase in the five-phase model described in the preceding section—is about the verification of possible explanations (or 'hypotheses' or 'conjectures'; in the context of improvement projects this could be 'possible causes'). What counts for the justification of hypotheses is not how the inquirer has thought them up, but whether they pass severe tests[§]. The confrontation of hypotheses with empirical evidence before they are accepted is the essence of scientific method, based on the rationale that empirical evidence is based on real things, which are independent of our opinions about them. Consequently, the confrontation of hypotheses with empirical evidence ensures that inquiry is self-corrective and objective (this is Peirce's error-correcting doctrine; Mayo[13] (ch. 12) provides a recent discussion).

Discovery—the Exploration phase in the five-phase model—is about the invention of possible causes or, in general, hypotheses. The thinking in this process is speculative and the activities not methodical. Compared with justification, discovery is poorly understood in philosophy of science. In fact, it has often been denied that there could be such a thing as a 'logic of discovery', given the non-methodical nature of this work[14]. Popper[8] (p. 31) described these processes as a subject of study for psychology, not logic. However, in more recent years, philosophers have regained interest in discovery, taking up ideas that were originated by Peirce around 1900. Peirce[15] identified the typical way of reasoning in discovery as *abductive*. The general form of an abduction is:

> the surprising fact C is observed;
> but if A were true, C would be a matter of course,
> hence, there is reason to suspect that A is true.

(Peirce[15] (CP 5.189); Niiniluotos[16] gives a recent discussion, plus a comparison with deductive and inductive inference). The principle is to conjecture a hypothesis (or cause) if it accounts for the observed facts. An abduction does not result in the acceptance of a hypothesis. Rather, the conclusion is that there are reasons for pursuing it, or deeming it testworthy. Recent developments in philosophical studies of discovery—under the

[‡]Explanations provide understanding of a system by showing what observed behaviour of the system should be expected as a consequence of certain causal mechanisms. The notion that explanation is based on the discovery of cause-and-effect relations is a popular view in philosophy of science (compare for example Popper[8] (p. 61) or Salmon's Statistical-Relevance model[9]) and industrial statistics (Shewhart[10], pp. 364–368); but see Kitcher[11] for an overview of alternative notions of explanation. Control based on the predictive power of explanatory models is often mentioned as one of the main functionalities of empirical science.

[§]What *severity* of a test means, is made quite precise by Mayo[13]. Her severity criterion (p. 180) is closely related to but more general than the Neyman–Pearson theory of hypothesis testing and, in particular, significance levels and p-values are closely related to severity.

DOI: 10.1002/qre

name *computational philosophy of science*—concentrate on abduction and techniques derived from artificial intelligence and the cognitive sciences[17,18].

It should be emphasized that it is the reiterated alternation of discovery and justification (or of exploration and confirmation) that is responsible for progress in scientific inquiry. Many activities are tentative and fallible and often, when insight advances, the inquirer will have to backtrack to previous assumptions. Progress is achieved by refining hypotheses on the basis of evidence, testing them again, refining them again, etc. Popper[19] depicts this process as a series of ‘conjectures and refutations’. Moreover, the activities the inquirer performs often cannot be clearly distinguished into discovery and justification. In general, inquirers collect new evidence (in particular by conducting designed experiments) specifically aimed at the confirmation of hypotheses, but sometimes the evidence collected for the generation of hypotheses is so conclusive that their justification is based on this evidence as well, without performing a confirmatory experiment. In this case, abduction becomes ‘inference to the best explanation’[16,20]. However, the principle of justification could be based on a severity criterion (Mayo[13] (ch. 9) deals with so-called ‘postdesignated tests’, i.e. tests that confirm a hypothesis on the same evidence that was used to formulate it).

Dewey[21] (ch. 6), in his famous ‘analysis of a complete act of thought’, inserts a problem definition phase before discovery activities start, and discerns a phase in between discovery and justification in which hypotheses are elaborated and made more precise, so that they can be tested. The whole procedure[¶]—putting forth hypotheses, testing them to evidence and accepting for true those hypotheses that are confirmed—is an instance of inductive inference (Peirce[15], CP 6.522–6.528; Maher[23]).

2.2. *Artificial intelligence and cognitive sciences: problem solving*

In artificial intelligence, what we call the Exploration phase is subsumed under problem solving. A crystallized formulation of the context can be found in Langley *et al.*[24] (p. 8). Human problem solving proceeds by first creating a symbolic representation of the problem (called the *problem space*). The initial state, intermediate states and the objective state are represented in the form of interrelated symbol structures. Efforts to solve the problem proceed by reorganizing this symbol structure and evaluating the result against the objective state, thus conducting a mental search for a solution through the problem space. However, this search is typically not carried on by random trial and error, it is guided in the direction of the objective state by rules of thumb, called *heuristics*. Heuristics should not be mistaken for algorithms to solve problems, there is no reason why or guarantee that they should work. The claim they make is that they are, on average, more efficient in finding a solution than random trial and error. Heuristics are typically based on information extracted from the structure of the problem at hand, contextual knowledge, experiences with earlier, comparable problems and generic approaches for solving problems.

2.3. *Medical sciences: diagnosis*

A related field is medical diagnosis[25]. The strength of medical diagnosis consists of the ability of experts to rely on know-how and experience, as well as on diagnostic heuristics, such as informal strategies, often domain-specific, that experts learn in the course of their professional life. Smith[26] translates many of these heuristics to the context of diagnosis of quality problems.

2.4. *Approaches for the Exploration phase*

The account so far shows that the activities in the Exploration phase—identification of possible causes—proceed by abductive reasoning; that it is not possible to prescribe a method or algorithm for this process, but that

[¶]20th century philosophy of science has yielded a number of variations on Dewey’s account. The ‘hypothetico-deductive’ method of the logical positivists[12] was an attempt to cast scientific method in rigid logical terms. The name refers to what was seen as the core of the method: a free invention of hypotheses, deducing testable consequences from them and subsequently testing them. Popper’s falsificationism is a variant that emphasizes that hypotheses can be refuted by empirical evidence, but never strictly confirmed. De Mast[22] relates Box’s sawtooth model and the PDCA-cycle (or Shewhart–Deming wheel) to these procedures.

DOI: 10.1002/qre

guidance ought to consist of heuristics. It is the purpose of this section to discuss a number of classes of heuristics for the generation of potential influence factors in the context of quality improvement projects.

2.4.1. *Inventory of process know-how*

The first class of heuristic principles we discuss is the *inventory of process know-how*. Persons working with a production process—operators, mechanics, engineers, etc.—will accumulate knowledge about the process, and among this knowledge will be important clues for process improvement. Typically, this process know-how is exploited in everyday problem solving and process control. However, for more persistent problems, or for systematic process optimization, this know-how often falls short and more rigorous approaches are needed, such as the five-phase approach described earlier. Within this five-phase structure, process know-how can play an important role in the Exploration phase. The know-how of persons who have worked with the process could yield a lot of ideas that are interesting to study more closely in the Confirmation phase. Thus, process know-how is used to generate potential causes.

When making an inventory of process know-how, however, one typically comes across the three problems that process know-how is often dispersed, tacit or conditioned.

Process know-how is dispersed. Different people have different pieces of the puzzle, but nobody sees the whole picture. Depending on their professional background and involvement with the process, people look to the problem under study from different perspectives, and they have different experiences. It is important for the project leader to make an inventory of all of these perspectives. This is usually accomplished by having a meeting with persons of different disciplines and backgrounds, such as operators, mechanics, internal customers, experts, etc., during which everybody brings up potential causes. These meetings could be given some structure by following the steps in a process flow, or by using an Ishikawa diagram. This approach of combining multi-disciplinary know-how could be called *knowledge pooling*.

Process know-how is tacit. Persons know intuitively what to do, but this knowledge is not available in explicit, worked-out form (tacit knowledge is knowledge that works in the background of consciousness and directs attention and action, but which is not made explicit or linguistically codified[27]). In order to help people explicate and set out in detail what they know intuitively, a *root-cause analysis* could be conducted. During a root-cause analysis, a person or group of people analyse cause-and-effect relations, typically supported by CE diagrams. Gano[28] proposed the following four principles for making root-cause analyses.

- Causes and effects are the same thing, i.e. each cause is another cause's effect, and as a consequence, *the* cause does not exist.
- Causes and effects are part of an infinite continuum of causes. Thus, in view of this and the previous principle, root-cause analysis should strive to explicate a cause-and-effect chain, rather than look for a single cause. This is the background of Ishikawa's[29] (p. 231) advice always to repeat the question 'why?' over and over.
- Each effect has two types of causes, namely action and conditional causes. An action cause is the immediate trigger that sets the mechanism in motion. Conditional causes are the necessary boundary conditions for the trigger to have its effect.
- Effects exist only if their causes exist at the same point in time and space, which gives an important clue for their identification.

Process know-how is conditioned. When a person's or a group's thinking moves on a circular track, while not coming across potential solutions that were not tried before, we say it is conditioned (Beveridge[30], p. 65). Thinking becomes conditioned because each time our thoughts take a certain course, the more likely that course is to be followed the next time. Quite generic ways of freeing our thinking from conditioning are temporary abandonment and discussion. A more methodical technique is brainstorming. Brainstorming proceeds by free association. If brainstorming is done with a group of persons, participants are asked to associate on each other's ideas. The technique is based on De Bono's[31] distinction between vertical thinking and lateral thinking. Vertical thinking is convergent and rigid, aimed at narrowing down options in an analytical, rational way.

DOI: 10.1002/qre

In contrast, lateral thinking—the basis of brainstorming—is divergent and indirect, aimed at finding new perspectives in an associative, irrational and synthetic way. The basic assumption is that random or illogical (combinations of) ideas from different domains may yield new directions that, when elaborated, may result in possible solutions. Techniques for brainstorming and lateral thinking are based on a few generic principles, such as provocation, free association and use of random perspectives.

The three approaches—knowledge pooling, root-cause analysis and brainstorming—are three extremes; in practice project leaders will probably combine elements of all three of them in a single meeting. Note that Shainin[32] rejects the use of process know-how for process improvement altogether, based on the idea that such an approach would be subjective or less effective. However, De Mast[2] argues that these claims cannot be defended in this rigid form.

2.4.2. *Deduction of ideas from scientific theories*

Potential causes could be found by consulting technical literature or experts. The derivation of ideas from literature is deductive, but their applicability to the given problem is a hypothesis that should be tested in the Confirmation phase.

2.4.3. *Exploratory data analysis*

The third class of approaches for the generation of possible causes is based on the collection and analysis of data from the running process. This type of data analysis was named *exploratory data analysis* by Tukey[33,34] and should be clearly distinguished from *confirmatory data analysis* (this distinction is the analogue in data analysis to exploratory and confirmatory reasoning in inquiry). The aim of confirmatory data analysis is to test and model the effect of a given potential cause. This is a methodical activity for which we have all the standard machinery of statistical inference, such as hypothesis testing, estimation and modelling. Tukey uses as a metaphor, the work of a judge.

Exploratory data analysis has as its purpose the *identification* of potential causes. It is about hypothesis generation, rather than hypothesis testing, and could be compared to the work of a detective (in fact, the logic of Sherlock Holmes' 'deductions' is indeed typically abductive[16]). As the project leader does not know in advance what he is looking for, graphical methods are especially powerful as they have the potential to reveal what was not expected beforehand. Displaying the data in various ways, the project leader should look for salient patterns, and then relate these to possible causes.

The first step in exploratory data analysis is the identification of salient features in the data. Shewhart[10] coined the term *assignable cause* to describe features in data that could be related to an underlying cause (as opposed to random noise, which carries no information about possible causes). The basic idea is that 'our clue to the existence of assignable causes is anything that indicates nonrandomness' (Shewhart[10], p. 26). As humans have a tendency to interpret too many patterns as non-random, Shewhart proposes a number of rules that signal deviations from non-randomness, and therefore the possible effects of causes. Examples of such rules[‖] are the various approaches in control charting and change-point analysis, such as 3σ control limits, runs rules, CUSUM (cumulative sum) charts, and their generalization, i.e. cuscore charts[35].

Salient features in the data are the fingerprints of the effects of causes; upon identification of salient features, it is up to the project leader to relate them to possible causes. The guiding principle here is *explanatory coherence*. An identified pattern in the data could inspire a sudden insight, where all of the pieces seem to fit together. Explanatory coherence is the extent to which the pieces fit together, and is based on the extent to which an idea explains a wide range of observations and experiences, is consistent with background information, and is simple (i.e. parsimonious, with only a limited number of parameters or side-assumptions)[20].

[‖] In Shewhart's view[10] (pp. 15, 16, 26), the choice of these rules is not motivated by probabilistic arguments, but is based on the experience that certain rules tend to be more successful in identifying causes than others. For example, when a project leader collects the successive measurements 8.91, 8.92, 8.92, 8.94, 8.94, 8.95, 8.95, 8.96, 9.00 he would interpret these as evidence of an assignable cause. However, the given sequence is as likely as, for example, 9.00, 8.94, 8.96, 8.92, 8.95, 8.91, 8.92, 8.94, 8.95 (in the sense that choosing randomly a sequence from all 9! enumerations of these numbers, the probability of drawing either of the given sequences is $\frac{1}{9}$!).

DOI: 10.1002/qre

For exploratory data analysis, the project leader collects CTQ measurements while the process is running (so-called *observational* data). Descriptive statistics give the project leader a feeling for the magnitude and perhaps the nature of the problem. Boyles'[36] exploratory capability analysis elaborates this idea. Next, the data could be plotted against a generic variable, for example time or a spatial parameter. A plot of the data against time is called a runs chart, but control charts are also used frequently for this purpose[37]. Assignable causes are identified by runs tests and change-point analysis techniques, which screen the data for predefined patterns. An example of a plot of the data against spatial parameters is the concentration diagram (Gryna[38], p. 22.45) or defect map[39]. In such a diagram, assignable causes are identified as clusters.

Instead of plotting the data against a generic variable, the alternative is to plot the data against variables that, given the process and the way the data were collected, are obvious candidates. If the CTQ measurements were collected as a stratified sample (with strata defined by product streams, batches of raw materials, shifts or other classifications that are obvious candidates given the process), differences among strata could be studied using box plots. Strata function as a sort of 'container factor', in the sense that they confound the effects of a multiplicity of variables (e.g. the effects of various properties of raw materials are confounded in a *batches* stratum).

Often, in addition to the CTQ a project leader also measures many other variables that are obvious candidates, or which happen to be measured with the CTQ anyway. Using scatter diagrams he could look for correlations**.

Salient features of high-dimensional data could be revealed by projecting the data onto two-dimensional planes. Matrix plots show projections onto planes defined by pairs of the original variables. Friedman and Tukey's[40] projection pursuit uses heuristics to identify planes that promise to result in revealing projections, but that are not necessarily parallel to the original variables. Principal component analysis is a heuristic that could enable the project leader to conjecture about latent factors that explain the observed variables, and thus could reveal root causes.

2.4.4. *Process and product examination*

Instead of collecting and analysing data from the running process—what we referred to as exploratory data analysis previously—the project leader could perform a close, but more qualitative examination of the process or products, looking for symptoms that could be indicative of possible causes. In its simplest form, this amounts to a detailed search for and description of symptoms (called 'cue acquisition' in medical diagnosis, and 'functional search' by Smith[26]), for example in the form of what Gryna[38] calls 'autopsies', in which defective products or processes are disassembled and carefully studied to find indications of the nature of the problem. Process and product examination could be guided by generic questions such as 'what?', 'where?', 'when?', 'how?' and 'how much?'.

Another form of idea generation by examination is manipulative abduction (or action-based reasoning[41]), which could be described as learning by probing, fiddling with and playing with the system. By experimenting in an unplanned (that is, not focused on pre-specified potential causes) and often unsystematic way, the inquirer builds feeling for and intuitive understanding of the system. By conceptualizing these tacit inferences, possible causes are identified.

Process and product examination should be guided by a sensibility for the salient and anomalous. The project leader could look for side effects that typically accompany the problem, and that could give away clues about the nature of the problem. A procedure to do this systematically is pairwise comparison[42]. In this procedure, the project leaders closely compares a *best of the best* (BOB) product to a *worst of the worst* (WOW) product, noting down differences in dimensions, weight, looks, electrical properties, etc. Repeating the comparison for

**Such an exploratory (but often very systematic) search for correlations is easily confused with hypothesis testing, and way too often p-values are used in this context. In confirmatory data analysis, p-values are used where the effect to be tested is stated prior to the data collection. It is permitted for a project leader to make scatter plots of the CTQ against all variables that were measured along and then to select the factors that appear to have the strongest correlations for further testing. It is problematic to perform this further testing on the basis of the same dataset, i.e. a hypothesis should in general not be tested on the same data that were used to generate it. Mayo[13] (ch. 9) gives an indepth discussion of this matter.

DOI: 10.1002/qre

successive pairs of BOBs and WOWs, the project leader hopes to find differences between BOBs and WOWs that recur in each pair.

2.4.5. *Suggestions from analogous problems*

Similar problems that have been solved in the past could inspire the solution of current problems. The underlying idea is that similar problems have similar backgrounds or structures.

One form this idea could take is the use of standard categories in which causes are sought. For example, causes of industrial problems are often sought in the standard categories Manpower, Machines, Materials, Methods and Measurements (the five Ms; Ishikawa[43], pp. 230–231). Offering standard categories does not serve the purpose of categorizing potential causes, but guides a systematic search for them, i.e. it broadens the focus of people's attention. People might be focusing exclusively on man and machine related causes for a problem so by asking them to shift their focus to one or more of the other categories broadens the perspective. The idea to look for causes in standard categories goes back to Aristotle, who offered—as a heuristic for finding arguments in debates—ten categories in which to find possible predicates (essence, quantity, quality, relation, place, time, position, state, activity and passivity). Smith[26] offers a highly detailed taxonomy of causes in quality problems.

Another way in which experiences with earlier problems are deployed in solving current problems, is the *Theory of Inventive Problem Solving*, usually abbreviated as TRIZ (after the Russian Теория Решения Изобретательских Задач). TRIZ was developed in the Soviet Union from the 1950s onward by Genrich Altshuller and his group[44,45]. They studied millions of patents, trying to identify patterns that they could use to advize problem solvers. Altshuller claims that around 95% of problems have been solved by somebody else before, be it in a different context. When brought down to their bare essence, there are only 40 essentially different 'innovative principles' underlying the millions of patents Altshuller studied. Altshuller recommends problem solvers to make an abstract analysis of the problem under study, laying open the essential contradiction that is the core of the problem. Formulating the problem as a contradiction between two out of 39 'generic parameters' (such as *temperature* and *amount of substance*), the problem solver can consult a 'contradiction matrix', which shows by means of which innovative principles similar contradictions have been resolved in the past (e.g. conflicts between temperature and amount of substance have been resolved in the past by, among others, inventive principle no. 30: flexible shells and thin films).

2.4.6. *Progressive search through families of causes*

The core idea of progressive search was well stated by Arthur Conan Doyle's character Sherlock Holmes: 'When you have eliminated all which is impossible, then whatever remains, however improbable, must be the truth'[††]. The problem space is divided in a number of classes (or families of causes). By observation (often data collected deliberately for this purpose), complete classes can be eliminated at once, and thus, the project leader zooms in on the classes that contain the important causes. The approach goes under a variety of names; for example, Smith[26] refers to it as 'tree search', Shainin[32] calls it 'eliminate and zoom in'. In fact, progressive search through families of causes is the core of the Shainin System, which advocates it as superior to the other approaches discussed in the present paper.

A simple form of progressive search is what Smith[26] refers to as a 'topographic search,' i.e. a cause is localized in place and time. Observing where the problem manifests itself, the project leader zooms in on the process prior to that point. Observing where the problem does not yet manifest itself, the project leader eliminates all causes before that point. Computer programmers often follow this strategy in tracking down a bug in their program. Instead of checking each line in the source code one by one, they temporarily insert *print* commands in the code to provide them with clues about the first manifestation of the problem. Then, they can probably eliminate large parts of the source code where the bug cannot be.

[††]This statement appears in a number of varieties in the various books, this formulation is from *The Blanched Soldier*.

DOI: 10.1002/qre

Table I. Approaches for the generation of possible causes of quality problems

Approach	Source of inspiration	Techniques
Inventory of process know-how	Tacit knowledge, technical knowledge, anecdotal evidence	Knowledge pooling, brainstorming techniques/lateral thinking, root-cause analysis
Deduction of ideas from scientific theories	Explicated technical knowledge	Literature search, consultation of experts
Exploratory data analysis	Observations, measurements	Descriptive statistics, run chart, control chart, concentration diagram, scatter plot, boxplot, matrix plot, projection pursuit, principal components analysis
Product and process examination	Observations, manipulations	Cue acquisition, symptoms analysis, autopsies, manipulative abduction, pairwise comparison
Suggestions from analogous problems	Experience in earlier problem solving projects	Standard categories of causes (5Ms), TRIZ
Progressive search through families of causes	Measurements	Topographic search, progressive search through classes of variation

Shainin[32] proposes a progressive search based on classes of variation. Using tools such as the multi-vari chart, a project leader finds out whether the dominant source of variation is temporal, stream-to-stream, unit-to-unit or within unit. If the dominant type of variation is, say, unit-to-unit, the project leader need not search for possible causes that vary within units or that exhibit stream-to-stream differences, but instead homes in on possible causes that vary from unit-to-unit.

3. EXAMPLES

Table I gives an overview of the approaches that were listed in the preceding section. This section positions some popular approaches for problem-solving in the theory that this paper develops.

3.1. Ishikawa's CE diagrams

Ishikawa[43] (ch. 3) proposes the CE diagram (also known as an Ishikawa or fishbone diagram) as a means to visualize cause-and-effect relations in the context of quality problems. A CE diagram consists of a large arrow, pointing towards a quality problem (the effect). The large arrow has a number of branch arrows, which represent the main categories of causes. Causes and subcauses are added as twigs.

Acknowledging that the CE diagram is applied in many different forms ('a good cause-and-effect diagram is one that fits the purpose'), Ishikawa[43] discerns three types of CE diagrams.

(1) Dispersion-analysis-type CE diagram. The main branch arrows represent the standard categories of sources of dispersion in quality characteristics, such as workers, materials, inspection and tools. The diagram helps identify causes, deeper causes and root causes by reiterating the question 'why does this dispersion occur?'.
(2) Production-process-classification-type CE diagram. The main branch arrows are steps in the production process, while twigs added to main branches represent causes thought to be acting in that particular process step (causes that are active in several process steps are added as twigs to multiple branch arrows).
(3) Cause-enumeration-type CE matrix. Possible causes are identified by free association (without guidance of standard categories of causes or process steps) and simply listed. A CE diagram is made afterward, to sort the listed causes and relate them to each other.

Ishikawa argues the power of CE diagrams from their ability to stimulate learning by pooling the knowledge of different people, their ability to focus discussions (preventing participants to stray from their topic or to repeat oneself), and to guide attention from symptoms to causes.

DOI: 10.1002/qre

Trying to position the CE diagram in the theory expounded in this article, it appears that it is based on a multiplicity of principles. In the dispersion-analysis-type diagram (having *workers*, *materials*, *inspection* and *tools* as main branches) we recognize the heuristic method to use standard categories or taxonomies of causes (discussed in Section 2.4.5). The production-process-classification-type diagram exploits the same idea (namely, to broaden the focus of people's attention by systematically checking a list of categories), with standard categories of causes now formed by process steps. The reiteration of the 'why?' question plus the diagrammatic representation of cause-and-effect chains make the CE diagram a form of root-cause analysis (discussed in Section 2.4.1). The CE diagram as a tool to structure discussions and enable knowledge sharing is related to the approach of knowledge pooling (discussed in Section 2.4.1).

In addition to incorporating these three approaches for hypothesis generation (use of standard categories, root-cause analysis and knowledge pooling), the CE diagram is also a tool for book-keeping of identified potential causes, and for visualizing the results of hypothesis generation. In particular, the cause-enumeration-type diagram offers no heuristic for hypothesis generation beyond an allusion to a brainstorming-like process, and is merely a presentation device.

3.2. *Kepner and Tregoe's Problem Analysis*

Kepner and Tregoe[46] propose a number of approaches for problem solving, decision making and planning, which are widely taught and applied in business and industry. We study the Problem Analysis procedure (Kepner and Tregoe[46], ch. 2) and relate it to the theory developed in this paper.

A problem is defined as a deviation of the actual from the should-be performance. The objective of Problem Analysis is to discover the causes and define remedial actions. Problem Analysis comprises the following activities.

(1) *State the problem*. The problem is given a name, which gives a description as precisely as possible of the deviation between should-be and actual performance.
(2) *Specify the problem*. The problem solver describes symptoms and features of the problem in the four categories what (what object?, what is the deviation?), where (where is the defective object?, where on the object is the deviation?), when (when was the problem first noted?, is there a pattern in the occurrence of the problem?) and extent (how many defective objects?, how large is the deviation?, how many deviations per object?, is there a pattern in the magnitude of the deviation?). These features of the problem are contrasted to corresponding features when there is no problem (*is* versus *is not* analysis). For example, the problem solver writes down what objects do not have a deviation, where the defective object is not, when the problem was not noted, etc.
(3) *Develop possible causes*. Two approaches are recommended. The first approach is simply to list ideas based on knowledge and experience. The second approach is more methodical: to identify potential causes based on distinctions and changes. Based on the *is* versus *is not* analysis, the problem solver identifies what distinguishes objects, locations, situations and behaviour where the problem is from those where it is not. Then, the problem solver checks whether there were changes in the process that relate to these distinctions.
(4) Test possible causes against the specification. By deductive reasoning the project leader assesses for each cause whether it could explain the problem.
(5) Determine the most probable cause. Based on the extent to which each cause could explain the problem, but also the number of and plausibility of side assumptions, the most likely cause is identified.
(6) Verify assumptions, observe, experiment, or try a fix and monitor. This step comprises activities to finally confirm that the true cause was found.

As with the CE diagram, a number of principles can be discerned that underlie Problem Analysis. The procedure gives some flexibility, allowing for patterns of hypothesis generation that range in between two extremes.

(1) Problem statement → cue acquisition based on generic questions (what?, where?, when?, how much?) → combine cues with process know-how.

DOI: 10.1002/qre

(2) Problem statement → identification of BOB and WOW objects (what), locations (where), situations (when) and behaviour (extent) → identification of distinctive features between BOBs and WOWs → zoom in on causes whose behaviour (change) could be related to these distinctive features.

We conclude that Kepner and Tregoe's Problem Analysis consists of process and product examination (including a variant of pairwise comparison), augmented with the use of process know-how.

4. DISCUSSION

Improvement programmes such as Six Sigma teach people a scientific attitude towards problem solving and quality and efficiency improvement. The essence of a scientific attitude is, according to Dewey[21] (pp. 13 and 74), suspended judgment and patience to resist jumping to conclusions. This attitude is incorporated in the Six Sigma method in that it demands a good problem definition before a project leader starts to think about possible causes, and proven theories before he designs an improvement action. Suspension is also vital in the Exploration phase. Many project leaders have a tendency to act on the first two or three ideas that are raised and rush of immediately to design experiments to test them. The effectiveness of the approaches discussed in this paper depends on the project leader's patience to observe and study the symptoms of a problem thoroughly, collect data, have meetings with relevant persons, etc. An active search for and cultivation of a variety of alternative ideas opens up new directions; experiments will be conducted a bit later, but probably more effectively. Despite, for instance, Shainin's claim, it is not possible to give a method for the Exploratory phase, which is speculative and adventurous in nature. At best, one can give heuristics, which are necessarily fallible, and whose effectiveness is very context dependent. The latter implies that a project leader needs a large store of heuristics in order to be an effective problem solver in a range of situations.

REFERENCES

1. Juran JM. *Juran on Leadership for Quality—an Executive Handbook*. Free Press: New York, 1989.
2. De Mast J. A methodological comparison of three strategies for quality improvement. *International Journal of Quality and Reliability Management* 2004; **21**(2):198–213.
3. De Mast J. Quality improvement from the viewpoint of statistical method. *Quality and Reliability Engineering International* 2003; **19**(4):255–264.
4. Shainin R. Strategies for technical problem solving. *Quality Engineering* 1993; **5**(3):433–448.
5. Ross PJ. *Taguchi Techniques for Quality Engineering*. McGraw-Hill: London, 1988.
6. Breyfogle F. *Implementing Six Sigma—Smarter Solutions Using Statistical Methods*. Wiley: New York, 1999.
7. Box GEP, Liu PYT. Statistics as a catalyst to learning by scientific method, part I. *Journal of Quality Technology* 1999; **31**(1):1–15.
8. Popper KR. *The Logic of Scientific Discovery*. Hutchinson: London, 1959.
9. Salmon WC. *Scientific Explanation and the Causal Structure of the World*. Princeton University Press: Princeton, NJ, 1984.
10. Shewhart WA. *Statistical Method from the Viewpoint of Quality Control*. Reprinted by Dover Publications: New York, 1986.
11. Kitcher P. Explanation. *Routledge Encyclopedia of Philosophy*, vol. 3, Craig E (ed.). Routledge: London, 1998.
12. Losee J. *A Historical Introduction to the Philosophy of Science* (4th edn). Oxford University Press: Oxford, 2001.
13. Mayo D. *Error and the Growth of Experimental Knowledge*. University of Chicago Press: Chicago, IL, 1996.
14. Nickles T. Discovery, logic of. *Routledge Encyclopedia of Philosophy*, vol. 3, Craig E (ed.). Routledge: London, 1998.
15. Peirce CS. *Collected Papers 1–5*, Hartshorne C, Weiss P (eds.). Harvard University Press: Cambridge, MA, 1931–35.
16. Niiniluoto I. Defending abduction. *Philosophy of Science* 1999; **66**(supplemental):S436–S451.
17. Thagard P. *Conceptual Revolutions*. Princeton University Press: Princeton, NJ, 1992.
18. Magnani L, Nersessian NJ, Thagard P. *Model-Based Reasoning in Scientific Discovery*. Kluwer Academic: New York, 1999.
19. Popper KR. *Conjectures and Refutations*. Basic Books: New York, 1963.

DOI: 10.1002/qre

20. Thagard P. Rationality and science. *Handbook of Rationality*, Mele A, Rawlings P (eds.). Oxford University Press: Oxford, 2004.
21. Dewey J. *How We Think*. D. C. Heath and Co: Boston, MA, 1910 [reprinted by Dover: Toronto, 1997].
22. De Mast J. Quality improvement from the viewpoint of statistical method. *PhD Thesis*, University of Amsterdam, 2002.
23. Maher P. Inductive inference. *Routledge Encyclopedia of Philosophy*, vol. 4, Craig E (ed.). Routledge: London, 1998.
24. Langley P, Simon HA, Bradshaw GL, Zytkow JM. *Scientific Discovery: Computational Explorations of the Creative Processes*. MIT Press: Cambridge, MA, 1987; 24.
25. Elstein AS, Shulman LS, Sprafka SA. *Medical Problem Solving: An Analysis of Clinical Reasoning*. Harvard University Press: Cambridge, MA, 1978.
26. Smith GF. Determining the cause of quality problems: Lessons from diagnostic disciplines. *Quality Management Journal* 1998; **5**(2):24–41.
27. Polanyi M. *Personal Knowledge*. Routledge: London, 1958.
28. Gano DL. Effective problem solving: A new way of thinking. *ASQ Annual Quality Congress Proceedings* 2001; **55**:110–122.
29. Ishikawa K. *Introduction to Quality Control*. Quality Resources: New York, 1989; 231.
30. Beveridge WIB. *The Art of Scientific Investigation* (2nd edn). Heinemann: London, 1953; 65.
31. De Bono E. *Lateral Thinking: A Textbook of Creativity*. Ward Lock Educational: London, 1970.
32. Shainin P. Managing quality improvement. *ASQC Annual Quality Congress Proceedings* 1993; **47**:554–560.
33. Tukey JW. *Exploratory Data Analysis*. Addison-Wesley: Reading, MA, 1977.
34. Tukey JW. We need both exploratory and confirmatory. *The American Statistician* 1980; **34**(1):23–25.
35. Box GEP, Ramírez J. Cumulative score charts. *Quality and Reliability Engineering International* 1992; **8**(1):17–27.
36. Boyles RA. Exploratory capability analysis. *Journal of Quality Technology* 1996; **28**(1):91–98.
37. De Mast J, Roes CB. Robust individuals control chart for exploratory analysis. *Quality Engineering* 2004; **16**(3):407–421.
38. Gryna FM. Quality improvement. *Juran's Quality Handbook* (4th edn). McGraw-Hill: New York, 1988; 22.1–22.74.
39. Bisgaard S. Importance of graphics in problem solving and detective work. *Quality Engineering* 1996; **9**(1):157–162.
40. Friedman JH, Tukey JW. A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers*, 1974; **23**:881–889.
41. Magnani L. An abductive theory of scientific reasoning. *Proceedings of the 2002 International Workshop on Computational Models of Scientific Reasoning and Applications (CMSRA'02)*, vol. III, Arabnia HR, Younsong Mun (eds.). CSREA Press: Las Vegas, NV, 2002; 1226–1232.
42. Bhote K. *World Class Quality*. Amacom: New York, 1991.
43. Ishikawa K. *Guide to Quality Control*. Asian Productivity Organization: Tokyo, 1982.
44. Altshuller G. *And Suddenly the Inventor Appeared: TRIZ, the Theory of Inventive Problem Solving*. Technical Innovation Center: Worcester, MA, 1996.
45. Altshuller G. *The Innovation Algorithm: TRIZ, Systematic Innovation and Technical Creativity*. Technical Innovation Center: Worcester, MA, 1999.
46. Kepner CH, Tregoe BB. *The New Rational Manager*. Kepner-Tregoe: Princeton, NJ, 1997.

Authors' biographies

Jeroen de Mast works as a senior consultant and researcher at the Institute for Business and Industrial Statistics of the University of Amsterdam (IBIS UvA). He holds a PhD in Mathematics. His research interests include exploratory data analysis, measurement system analysis and the business economic and scientific context of industrial statistics. He has written several books on Six Sigma and is the Conference Chair of the 6th Annual ENBIS Meeting.

Marcus Bergman holds an MSc degree in Biometry (applied statistics) from Wageningen University, the Netherlands. He is currently working as a Lean Six Sigma consultant for Uni Network Consultancy (UNC). He is also active as an editor of the Dutch Six Sigma information platform (http://www.sixsigma.nl).

DOI: 10.1002/qre