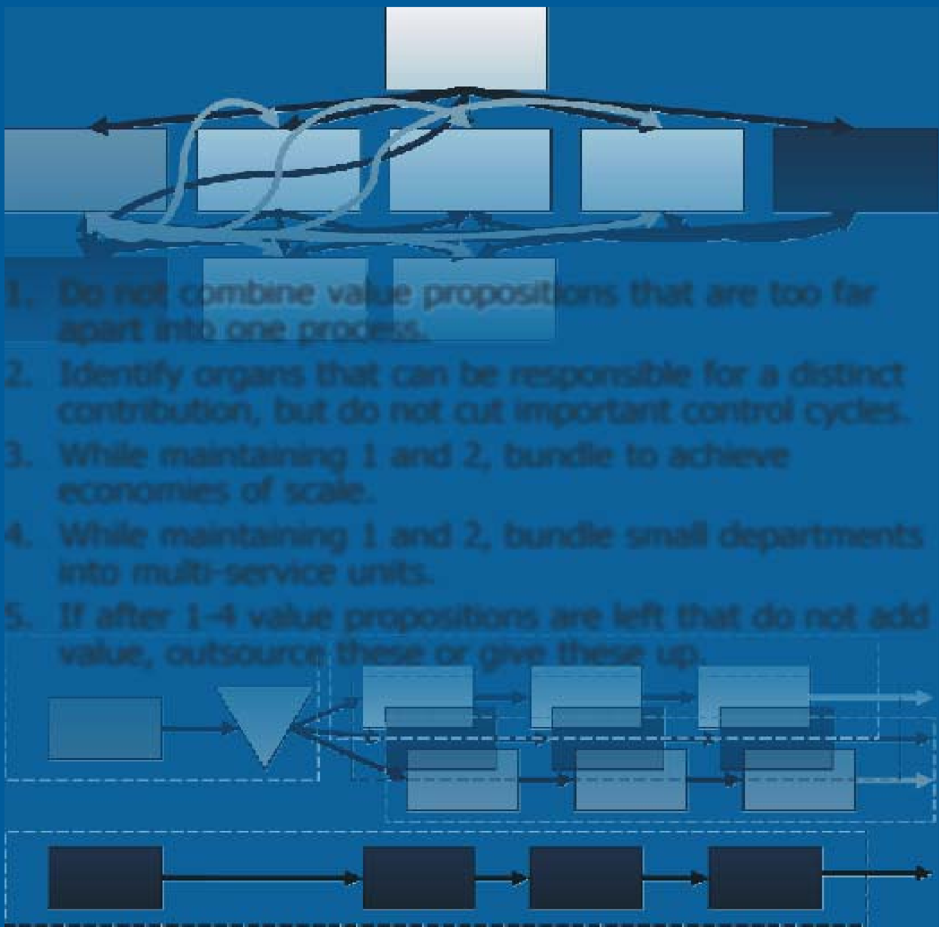
Organization Structures for Dealing with Complexity



Bart R. Meijer

Propositions regarding the thesis:

Organization Structures for Dealing with Complexity

Bart R. Meijer

1. Shops and companies offering a diversified portfolio of products and services should diversify their customer front-office as well.
2. Separation of authority and accountability causes complexity and stress.
3. "Economies of scale" are often in conflict with the principle "time is money".
4. To call every customer king does not imply to accept every king as customer.
5. Who does not honour teaching should not be rewarded research.
6. Cultivating the differences between science and design,
ignores the role of creativity for progress in science and
ignores the importance of science for successful designs.
7. Good and challenging academic education selects and improves the best students. Therefore, improving the quality of academic education in this direction will jeopardize the average course efficiency.
8. Without the influence of power, rules do not transform into rights.
9. An opera is a musical theatre play. An oratorium is a musical hear play. Performing oratoria in opera singing style deprives the listeners of their personal imagination and emotions.
10. Cyclist priority rights at roundabouts prove that ergonomics is not yet an issue in enhancing traffic safety.

These propositions are regarded as *lending themselves to opposition* and as *being defendable*. As such they have been approved by the supervisors.

Stellingen behorende bij het proefschrift:

Organization Structures for Dealing with Complexity

Bart R. Meijer

1. Winkels en bedrijven met een gediversifieerd portfolio aan producten en diensten, dienen hun contacten met klanten ook te diversifiëren.
2. Het scheiden van bevoegdheden en verantwoordelijkheden leidt tot complexiteit en stress.
3. De zogenaamde "economy of scale" is vaak in strijd met het principe "tijd is geld".
4. "De klant is koning" betekent niet "iedere koning is klant".
5. Wie het onderwijs niet eert, is het onderzoek niet weerd.
6. Het cultiveren van de verschillen tussen ontwerpen en wetenschap, gaat voorbij aan de rol van creativiteit in de vooruitgang van wetenschap en gaat voorbij aan het belang van wetenschap voor een succesvol ontwerp.
7. Goed en uitdagend academisch onderwijs selecteert de betere studenten. Onderwijs verbeteringen in deze richting brengen het gemiddelde studie-rendement in gevaar.
8. Slechts de kracht van macht maakt regels tot recht.
9. Een opera is een muzikaal theater stuk. Een oratorium is een muzikaal hoorspel. Het uitvoeren van oratoria in operastijl, ontneemt de luisteraars hun persoonlijke verbeelding en emoties.
10. Fietsersvoorrang bij rotondes toont dat ergonomie vooralsnog geen factor is bij het bevorderen van de verkeersveiligheid.

Organization Structures
for
Dealing with Complexity

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft
op gezag van de Rector Magnificus prof. dr. ir. J.T. Fokkema,
voorzitter van het College van Promoties,

in het openbaar te verdedigen op dinsdag 14 november 2006 om 15.00 uur

door

Bart Ruurd MEIJER
natuurkundig ingenieur
geboren te Velsen

Dit proefschrift is goedgekeurd door de promotoren:

Prof. ir H. Bikker

Prof. dr. T. Tomiyama

Samenstelling promotiecommissie:

Rector Magnificus	Voorzitter
Prof. ir. H. Bikker	Technische Universiteit Delft, promotor
Prof. dr. T. Tomiyama	Technische Universiteit Delft, promotor
Prof. dr. ir. G. Lodewijks	Technische Universiteit Delft
Prof. dr. ir. M. van Tooren	Technische Universiteit Delft
Prof. dr. ir. K. van der Werff	Technische Universiteit Delft
Prof. dr. ir. F.J.A.M. van Houten	Universiteit Twente
Prof. Dr. Dipl.-Ing. ETH F. Fahrni	ETH-Zürich / Universität St. Gallen

ISBN-10: 90-9020642-6

ISBN-13: 978-90-9020642-4

Acknowledgements

This research could not have been accomplished without the help and support of students, colleagues and family. First of all I need to thank my supervisors Henk Bikker and Tetsuo Tomiyama for their energy in critically going through all the concept versions of this thesis and not to forget Jenny Bikker for her support of our meetings in Heiloo, catered with a generous supply of coffee, tea and biscuits.

Also a thank you for Klaas van der Werff. With his open mind and curiosity to all aspects of Design, Manufacturing and Organization he made a great contribution to the working climate in the Production Technology and Organization group and later to the Design and Lifecylce group.

Dick Mensch proved a great source of expertise in queuing theory and statistics. Dick, thank you for your supply of articles and your effort in proof-reading the chapter and appendix on the simulation experiments.

I really enjoyed supervising students that were involved in organisation design problems in industry. The work of some of them proved very valuable for this thesis and are mentioned here: Frank van de Peppel, Leo van der Velde, Andy Schuurmans, Diederik van Rijn, Martijn Blommestijn, Marc Dijk and Joost Mackor.

Much of the conceptual work of this research was developed as conference papers to be presented at the IEEE-engineering management conferences. Jan Voûte and I usually went together, always with a number of sometimes co-authored papers and a TAO© workshop on innovation management. Jan, thank you for the inspiration and the many nightly discussions we had on business processes, management of innovation and complexity.

Finally I want to thank my wife Lilian and children Josine and Philip for their support and patience in waiting until papa's "proefschrift" was finished.

Summary

"Complexity is in the eye of the beholder" is a well known quote in the research field of complexity. In the world of managers the word complex is often a synonym for difficult, complicated, involving many factors and highly uncertain. A complex business decision requires careful preparation of the managers and workers involved. The preparation reduces the uncertainty and reveals the structure of the problems and processes to be dealt with. Thus complexity is a measure of the effort deemed necessary to resolve the uncertainty and solve the problem. Operating business processes requires dealing with a sustained stream of issues and problems in order to create the customer value. The organisational infrastructure can support as well as frustrate the efforts of the workers to deal with their part of the complexity. The goal of this thesis is to show that structure is one of the most important design variables in solving complex business management problems and to develop tools for supporting organization structure design.

Starting point for this research was the observation that the Delft School of Organization design, founded by Prof. Jan In't Veld and Prof. Pierre Malotaux, was based on a still unique doctrine about organization design:

By starting with the design of business processes and structuring these with the intent to facilitate quality management (customer value) and to optimize productivity, a structure can be found that serves as a basis for the design of a department and management structure. Thus an implicit and natural match is accomplished between the quality control and management requirements from the market and the management and control capabilities of managers and directors.

This match between process management requirements and management capabilities prevents many of the induced uncertainties and unnecessary coordination that could frustrate workers and managers if this match is not accomplished. Since most (re)organization processes start with reduction of head

counts and the (re)distribution of management power without even looking at business processes, a mismatch is not unlikely.

Complexity in the operations of business processes is driven by three factors: diversity, uncertainty and interrelations. Uncertainty is the most important of all three. Diversity can amplify uncertainty resulting in more complexity, but without the uncertainty, complexity disappears. Interrelations can reduce uncertainty if the nature of the coupling between different drivers is known. Without this knowledge interraltions seemingly increase uncertainty. In probability theory, interrelations are modelled as conditional probabilities. Shannon's complexity measure is a useful quantitative measure of complexity that supports this business operations oriented notion of complexity.

Structure design decisions are concerned with grouping or splitting functions into distinct sets called organs or organelles. With business process design as a starting point and by qualitatively looking at the sources of uncertainty, this thesis proposes the following structure design rules to support structure design decisions:

1. *Do not combine value propositions that are too far apart into one process.*
2. *Identify organs that can be responsible for a distinct contribution, but do not cut important control cycles.*
3. *While maintaining 1 and 2, try to create economies of scale (efficiency).*
4. *While maintaining 1 and 2, try to achieve some flexibility and reduce the vulnerability associated with small departments and product oriented structures by merging these (support) functions into larger multi-service units.*
5. *If after 1-4 value propositions are left that do not add value to the company, outsource them or give up these lines of business.*

The first rule prevents uncertainties induced by orders from different markets competing for the same resources. Diversity from two or more stochastic sources mixed into one source may even become a bigger source of complexity than the sum of these sources. The second rule prevents uncertainties induced by spreading the responsibility for controls over different managers. The third rule supports efficiency but may require investments in advanced planning and control to prevent reintroducing the uncertainties avoided by rules 1 and 2. The fourth rule reduces uncertainty over the availability of (human) resources. The fifth rule promotes a strategic focus. Strategic focus or limiting yourself to doing what you are good at is in itself an effective means to reduce complexity.

Quantitatively, business processes can be modelled as sets of states and state-transitions. The probabilities of the states as well as of the state-transitions can be used as input to calculate Shannon's complexity. With information theory it was demonstrated how structuring or sorting these sets can influence this measure. The conditional probabilities of transitions between sets, expressed as conditional information then represent the coordination effort between these sets. Thus by looking at the symmetries in conditional information, a set of theories is derived that supports splitting or combining decisions or prescribes a directive from one set to the other. If business processes are modelled as queue-server systems, a simulation study demonstrates that mixing job-streams is causing coupled leadtime behaviour between these streams, which can cause a big uncertainty over the leadtime of individual jobs for at least one of the streams. As a rule of thumb, job-streams that have logistic parameters more than a factor 4 to 8 apart are better served with a multiple queue-server system. Having dedicated business processes for both markets produces better leadtime results with less total capacity than a single process would need for the same performance. This result provides additional support for the first structure rule.

Shannon's complexity measure demonstrates how the choice of aggregates and the clustering of sets of states can reduce or increase the complexity of the issues relevant at that level. However this type of reasoning does not resolve the uncertainty. In addition to the design rules and the structuring theorems, tools and models have been described and developed that support finding the information needed to resolve the uncertainty. To name a few: discounted free cashflow methods as a financial reference for value creation, the customer value mix to support finding the factors of a definition of quality that balances technical requirements with business economics. Also an inter-human or inter-department communication model was developed to support the identification of sources of misunderstanding.

Compared to the processes of operations, the processes of Research and Development or Innovation are not as easily structured for quality and productivity. The interfaces, transitions and coordination between different parts of development work are under continuous influence of the same development work. This means a more dynamic and self-organized structure is needed. Inspired by the practices of set-based concurrent engineering and the working principles of genetic algorithms an evolutionary development organisation is proposed as a means to reduce the risk and impact of reversing early development decisions. Set-based or evolutionary development processes rely on creating and maintaining a high level of redundancy in the concepts and concurrency in the resources which can be recombined at any stage in the development. The concurrency maintains the speed, the redundancy reduces the risk and prevents the impact of reversing early decisions.
In contrast, reversing early design decisions in the context of work breakdown structure based development processes, typically throws a development schedule months back compared to the original plan.

Samenvatting

"Complexiteit is een kwestie van perspectief" is een bekend statement in het werkveld van complexiteitsonderzoek. In de wereld van de managers wordt het woord "complex" vaak gebruikt als synoniem voor moeilijk, ingewikkeld, afhankelijk van veel factoren en hoogst onzeker. Een complex management besluit vraagt zorgvuldige voorbereiding. Die voorbereiding reduceert de onzekerheid door het ontdekken van factoren en verbanden. In feite is complexiteit dus een maat voor de hoeveelheid inspanning die geleverd moet worden om inzicht te krijgen en een beslissing te nemen. Bedrijfsprocessen kenmerken zich door een continue stroom van dergelijke meer en minder complexe beslissingen die moeten bijdragen aan het realiseren van klantwaarde, waarvoor de klant ook bereid is te betalen. De organisatie infrastructuur kan dit proces ondersteunen maar ook frustreren. Het doel van dit onderzoek is te tonen dat structuur of organisatiestructuur één van de belangrijkste ontwerp variabelen is bij het oplossen van bedrijfskundige problemen. Ook is een bijdrage geleverd aan de ontwikkeling van gereedschappen ter ondersteuning van het ontwerpen van organisatiestructuren.

Het uitgangspunt van dit onderzoek was de nog altijd unieke organisatieontwerpdoctrine van de Delftse School voor Organisatieontwerp zoals opgericht door Prof. Jan In't Veld en Prof. Pierre Malotaux: *Door eerst bedrijfsprocessen te ontwerpen en die te structureren met als belangrijkste criterium het beheersen van kwaliteit (klantwaarde) en productiviteit (winstgevendheid), ontstaat een structuur die ook als basis kan dienen voor de afdelings- en personele structuur. Aldus wordt als vanzelfsprekend een match verkregen tussen de besturings- en beheersingseisen vanuit de markt en de besturings- en beheersingsmogelijkheden van directeuren en managers.*

Deze match tussen de eisen van procesmanagement en de management mogelijkheden van managers, voorkomt veel van de onbewust geintroduceerde onzekerheid en onnodige coordinatie, die het gevolg is van het niet op elkaar

afstemmen van de bevoegdheden en verantwoordelijkheden van managers en de in de markt gewortelde eisen voor het beheersen van bedrijfsprocessen. Aangezien de meeste (re)organisatie processen beginnen met aantallen ontslagen en herverdelen van de management macht, zonder ook maar te letten op bedrijfsprocessen, is een mismatch eerder waarschijnlijk dan uitzonderlijk.

Complexiteit in bedrijfsprocessen wordt hoofdzakelijk bepaald door drie factoren: diversiteit, onzekerheid en afhankelijkheden. Onzekerheid is de belangrijkste. Diversiteit kan onzekerheid zelfs versterken, maar zonder onzekerheid is diversiteit geen probleem en verdwijnt daarmee ook de complexiteit. Afhankelijkheden kunnen onzekerheid reduceren indien de aard van de interactie tussen verschillende factoren bekend is. Echter, zonder deze kennis wordt de onzekerheid schijnbaar groter. In waarschijnlijkheids theorie worden afhankelijkheden gemodelleerd als voorwaardelijke kansen. Shannon's maat voor complexiteit is een geschikte kwantitatieve maat die ook goed aansluit bij de bedrijfskundige betekenis van het begrip complexiteit.

Met het ontwerpen van bedrijfsprocessen als beginpunt en door vooral kwalitatief bronnen van onzekerheid te bestuderen zijn de volgende vijf "vuistregels" voor organisatiestructuurontwerp tot stand gekomen:

1. *Bedien geen waardeproposities die te ver uiteenliggen met slechts één bedrijfsproces.*
2. *Breng deelfuncties die samen één duidelijke functie vervullen samen in één orgaan en pas daarbij op geen belangrijke regelkringen te verdelen over meer organen.*
3. *Met inachtneming van regels 1 en 2, probeer functies te bundelen voor het realiseren van schaalgrootte.*
4. *Met inachtneming van regels 1 en 2, voorkom te kleine en kwetsbare afdelingen door kleine (ondersteunings) functies te bundelen tot grotere en flexibele multi-service eenheden.*
5. *Indien na het toepassen van regels 1 tot en met 4 er activiteiten overblijven die economisch niet rendabel zijn (=geen bedrijfswaarde creëren), probeer deze in te kopen bij beter toegeruste organisaties of stoot deze activiteiten af.*

De eerste regel voorkomt de onzekerheden die het gevolg zijn van de concurrentie tussen orders afkomstig uit verschillende markten voor dezelfde bedrijfsmiddelen. Diversiteit gevoed door verschillende markten, gemengd in één bedrijfsproces kan zelfs leiden tot een hogere complexiteit dan de som van de afzonderlijke markten. De tweede regel voorkomt de onzekerheden die voortkomen uit het spreiden van verantwoordelijkheid voor de beheersing over meer afdelingen en managers. De derde regel ondersteunt productiviteit en efficiency, maar vergt vaak investeringen in geavanceerde planningstechnieken om de voordelen van het toepassen van regel 1 en 2 te handhaven. Het voordeel is echter wel dat deze investering nu afgewogen kan worden tegen de voordelen van bijvoorbeeld gescheiden bedrijfsprocessen met een veel eenvoudiger beheersing. De vierde regel voorkomt de onzekerheid over de beschikbaarheid van bepaalde specialistisch medewerkers. Een multi-service afdeling kan werken aan opleidingen en een vervang beleid. Een kleine single service afdeling mist bij ziekte van een medewerker wellicht alle capaciteit om haar taak uit te voeren. De vijfde regel moet vooral een strategisch focus ondersteunen. Beperk de bedrijfsactiviteiten tot die activiteiten waarin men excelleert en winst maakt. Die strategie draagt belangrijk bij aan het reduceren van de complexiteit van zowel de markt als van de interne processen.

Kwantitatief kunnen bedrijfsprocessen gemodelleerd worden als een verzameling toestanden en toestandsovergangen. De kansen van optreden van die toestanden en de kansen van de overgangen kunnen gebruikt worden in het berekenen van Shannon's complexiteit. Met behulp van informatietheorie is getoond hoe het opleggen van een structuur aan die toestanden of het anders bundelen van toestanden invloed heeft op de complexiteit van dat perspectief. In feite worden met het anders ordenen van de toestanden en transities ook de functies anders geordend. De voorwaardelijke overgangen tussen verschillende verzamelingen, vertaald naar voorwaardelijke complexiteit is een maat voor de afstemmings-behoefte tussen twee toestandsverzamelingen. Door te kijken naar de waarde en de symmetrie van de voorwaardelijke complexiteit van de overgangen zijn een aantal regels afgeleid waarmee bepaald kan worden of de verzamelingen gebundeld moeten worden of gescheiden kunnen blijven, dan wel dat er spraken is van een éénzijdige afhankelijkheid (directive).

Door bedrijfsprocessen te modelleren als wachtrij-server modellen is een simulatie gebouwd waarmee de invloed van gemengde orderstromen kan worden bestudeerd. Zoals te verwachten was worden de doorlooptijdprestaties van verschillende stromen vrijwel identiek, indien deze orderstromen gemengd worden. Met name de onzekerheid over de maximale doorlooptijd van individuele orders neemt fors toe. De uiteindelijk gerealiseerde maximum doorlooptijd kan voor specifieke klanten onaanvaardbaar zijn. Als vuistregel kan gehanteerd worden dat orderstromen die in logistieke eigenschappen (ordergrootte en orderfrequentie) meer dan een factor 4 tot 8 verschillen, beter kunnen worden afgehandeld door specifiek gedimensioneerde bedrijfsprocessen. Deze oplossing zal resulteren in betere doorlooptijdprestaties bij een lagere totale capaciteit. Dit resultaat is ook een bevestiging van het belang van de eerste structuurontwerpregel.

Met Shannon's complexiteitsmaat is getoond hoe de keuze van een aggregaat en het bundelen van toestanden een reductie of een toename in de complexiteit van dat perspectief kan betekenen. Echter, deze redenering verandert niets aan de bronnen van onzekerheid die de aanleiding waren voor complexiteit. Daarom zijn naast de ontwerpregels ook diverse gereedschappen beschreven en ten dele ook ontwikkeld die kunnen helpen bij het vinden van de informatie die noodzakelijk is om de onzekerheid te verminderen. Om er enkele te noemen: de discounted free cashflow methode als financiele referentie voor het realiseren van bedrijfswaarde, de klantwaardemix voor het vinden van de producteigenschappen die klantwaarde bepalen en ter onderbouwing van een definitie van kwaliteit die zowel de technische specificaties als de bedrijfseconomische kant omvat. Ook is een model voor mens-tot-mens of afdeling-tot-afdeling communicatie ontwikkeld dat helpt bij het nader identificeren van bronnen van misverstanden of onbegrip.

Vergeleken met productie- of distributieprocessen zijn de processen van onderzoek en ontwikkeling of innovatie veel minder eenvoudig te structureren op basis van kwaliteit en productiviteit. De interfaces, overgangen en coordinatie tussen verschillende delen van ontwikkeling veranderen vaak als gevolg van diezelfde ontwikkelingsprocessen. Dit vergt een meer dynamische en zelf-organiserende aanpak van de bedrijfsprocessen. Geinspireerd door het voorbeeld van set-based concurrent engineering en door de werkingsprincipes van genetische algorithmen wordt een evolutionaire ontwikkelingsorganisatie voorgesteld. Set-based en evolutionaire ontwikkelingsprocessen zijn gebaseerd op het creëren en handhaven van redundantie in de concepten en een grote mate van gelijktijdigheid in de ontwikkeling ten behoeve van de snelheid. Met name de redundantie van de concepten maakt het mogelijk om ook in een later stadium concepten te herconfigureren zonder veel tijd te verliezen op de planning. Meer traditionele "work-breakdown-structure" gebaseerde methoden lopen vaak het risico dat in een laat stadium, vroeggenomen conceptbeslissingen moeten worden teruggedraaid. Dergelijke beslissingen leiden vrijwel altijd tot maanden vertraging ten opzichte van de oorspronkelijke planning.

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION TO THE PROBLEM FIELD

"Complexity is in the eye of the beholder" is a well known quote in the research field of complexity. In the world of managers the word complex is often a synonym for difficult, complicated, involving many factors and highly uncertain. A complex business decision requires careful preparation and close attention of the managers and workers involved. This preparation often reduces the uncertainty or reveals the structure of the problems and processes to be dealt with. A complex problem becomes less complex for those involved in solving it. This is the eye of the beholder. An experienced eye will perceive a different level of complexity than an inexperienced. However we cannot say, less complexity for the experienced, more for the inexperienced. Inexperienced observers may overlook, thus underestimate complexity. The goal of this thesis is to show that structure is one of the most important and often forgotten design variables in solving complex technical as well as business management problems. In this thesis a design procedure and rules for structure decisions will be presented. This chapter will outline the reasoning of this thesis.

1.1 Changing boundaries as cause and cure for complexity

Through the 70's, 80's and 90's, lowering trade barriers, free flow of capital, cheap transportation of men and goods and internet have brought global competition to almost any part of the world. Geographical boundaries that separated markets have changed or disappeared completely. CEOs of large multinational companies shared a religion with their shareholders and engaged in a game of mergers and takeovers to gain size and power as the sole strategy to survive. In every trade, it was believed, there will be space only for four to five world players and the rest is doomed.
In the same era, vertical integration was abandoned too. Especially since the 90's, competition and the need to speed up innovation forced companies to reconsider the business they were in. Focus on core competencies became the new credo [Prahalad 1990]. This was a clear reduction of complexity for the CEOs. The company only needs to excel at their core-competencies. The other assets needed for the business were considered commodities that could be sourced one way or the other. The number of markets or trades expanded along the lines of focused competencies.

Global competition has increased the pace of development. Economic product life is becoming shorter and shorter. Electric appliances for home use sell for only six months before a new line of slightly different appliances is introduced. Midsize cars are restyled on a yearly basis and receive a major make-over every three years. Computer systems have a sell by date of less than half a year and seem completely out of date within three. At the same time, sustainability, product liability and other physical product life related issues have stretched the time horizon designers need to take into account considerably.

The unprecedented increase in wealth in western societies through the 90's has induced a great level of individualism with consumers. The need for distinction has brought fashion to almost any product. The mass customization movement makes us believe that product variety will grow to almost infinitely to match all our very own individual needs.

All the trends above are related to changes in the structure of the marketplace, preceded or followed by changes to the business processes and companies that deliver to those markets. Geographically expanded markets come along with increased diversity. The need for local product variations did not disappear, nor did the diversity in trade fares, tax rates, accountancy practices and business culture. Managers are faced with continuous change and their perception of complexity changes along with it. Complexity management has become an important topic on the agenda of both management scientists and managers [Craig 1997][Wiendahl 1994].

Doing business in bigger markets also means dealing with increased uncertainty; who and where are my customers and what do they need? Who are my competitors? In addition, the increased pace of innovation puts more time pressure on the business processes and on the changes these processes may have to go through. Managers need to evaluate more options and deal with more influences, some of which are entangled in some complicated way. Along this line of thought we can define complexity as follows:

Definition 1.1 Complexity.

> ***Complexity is a quantitative notion of a problem, proportional to the effort deemed necessary to respond adequately to that problem in a specific context.***

Complexity ranks problems by the effort or amount of work needed to solve them. In line with linking complexity to effort and capacity, Malotaux considered the number of variables and even more important the number of relations between these variables as the main drivers of the complexity of decision making processes [Malotaux 1978]. Koolhaas identified the number of elements, the number of relations and time dependence as the three basic dimensions of complexity in management. This notion of complexity is closely related to the engineering science notion of complexity [Koolhaas 1980] and the numbers drive the problem solving effort.

Nam Suh implicitly relates complexity to the effort needed to engineer a system to perform in a certain context. Suh's complexity measure is related to the probability that the engineered system will perform as expected. This probability is modelled as the overlap between the so-called system range, the range of acceptable performances as specified by the functional requirements, and the so-called design range, the range of expected performances by the engineered system as realised by the choice of construction principles and their design parameters. In accordance with Shannon's measure of complexity, a low probability for the design to perform adequately, corresponds to a high complexity. Suh's axiomatic design methodology intends to systematize the choice of appropriate functional requirements, design parameters and the process of optimizing these, such that the effort in solving these issues becomes minimal [Suh 2001,2005].

Complexity is not an absolute notion. The effort to solve a problem, depends on skills and experience of the problem owner as well as on the size of the problem relative to the problem owner. In business processes, the size or complexity of problems is mainly driven by a combination of three factors: uncertainty, diversity and interrelations. Uncertainty is the main factor which can be amplified by diversity. Without uncertainty over the possible response, there is not really a problem and complexity is zero. Diversity alone, can be handled without difficulty with appropriate resources[1] as long as the influence of interrelations is small. The presence of interrelations can also induce uncertainty over the systems behaviour, especially if there is a lack of knowledge. Suh's axiomatic design method aims at preventing these interrelations or sequencing them in such a way that there exists a straightforward process to solve the design problems.

1.1.1 Timing and time-span

An important factor in perceived complexity is timing and time-span. Especially in research and development taking decisions is one thing, deciding when these decisions should be implemented is another. Launching new products too soon is almost as costly as launching too late. Discussing future products with today's sales personnel can lead to serious misunderstandings about customer needs. Explicit awareness of time-spans, time scales and timing prevents uncertainties introduced through misunderstandings that would occur without this awareness. Financial managers and R&D engineers often operate with a different mindset about time [Lambert 1996][Meijer 2000a].

1.1.2 Complexity handling capacity

Choosing the right perspective and mapping reality onto a suitable model helps to reduce the perceived complexity. If effective and efficient management of business processes is the goal, setting system boundaries as well as setting time boundaries through limiting time-span and timing is a powerful means to segment big and

1. Computers with automated data terminals are very effective in supporting processes that need to handle high levels of combinatorics.

complex problems into many smaller (less complex) ones that can be treated concurrently or sequentially, each at their own pace.

The complexity perceived by individuals as well as companies changes over time. The time to respond to these changes is limited. In a buyers market where a lot of companies are competing for their market share, most products loose half their sales value within six months. Being first on the market and being successful from day one now implies first time right. From day one the manufacturing cost need to be at a level low enough to maintain a positive margin even after six months of price erosion. As a result, companies are working together in supply chains where each company can focus on their core competence. Each supply chain partner is responsible for its own research and development to maintain their competitive edge. Generally speaking, higher quality at lower cost price is the result. At the same time, supply chains are also the vehicle to regain control over the distribution channels, thus limiting the competition and the complexity caused by competition.

Boswijk introduced the term *complexity handling capability* (CHV[1]) of organizations. In his view the complexity acting upon organizations has to be matched by the complexity handling capacity of that organization and he described many concepts to make this match [Boswijk 1993]. Despite the double interpretation of the Dutch term "vermogen[2]", Boswijk only considered "vermogen" as capability, not as power or capacity. The influence of deadlines, nor the influence of organization structures on the performance that could be achieved when releasing this capability were discussed in his work.

The problems of today involve more aspects, the number of options to evaluate is bigger and decisions often need to be made in shorter time. Adding more manpower to the problem solving process is not a solution. The increased complexity often calls for more coordination and more (skilled) workers call for even more coordination.

1. CHV stands for "Complexiteits Hanterings Vermogen" in the Dutch language.
2. "Vermogen" in the Dutch language means both *capability* as well as *power.*

The well known phenomenon of diminishing returns may be expected. Moreover, the training needed for introducing extra workers may more than kill the capacity added to the process. Adding people may reduce rather than increase the problem solving capacity [Brooks 1995].

Changing the processes and the structure of this problem solving process is a better option. The optimal span theory of Glickstein argues that for a given average inter connectivity between members of a group, there exists an optimal group size (span). Smaller groups have less capacity, larger groups are less efficient in coordination. Organizations that grow under pressure of performing more complex tasks have to split into parts and organize hierarchies to maintain their effective capacity. [Glickstein 1996, p53-60]. In a later paper, Glickstein proposes to combine the law of diminished returns that Brooks used, with his optimal span theory to calculate the theoretical efficiency of an organization. This model is capable of reproducing Brooks empirical finding that beyond a certain number of agents, project leadtime increases rather than decreases [Glickstein 2003].

1.1.3 Organizational boundaries as cause and cure

As argued in previous sections, changes in complexity are met with strategic and organizational responses that mainly affect boundaries between interacting systems. The goal of these organizational responses is to either increase the complexity handling capacity of the company or to change the complexity the company is facing. Without adequate strategic or organizational responses, managers will usually become exposed to the problems caused by complexity and they will become bottlenecks in the problem solving network.

Adequate organizational responses to complexity require understanding. Understanding the causes of complexity is a cognitive process, strongly influenced by knowledge, experience, modelling capabilities and analytical skills. In science as well as in business practice, models are used to analyse and control the complexity of systems. Especially technical sciences and engineering disciplines have a long standing tradition of using models to analyse and predict physical properties and behaviour of materials, structures and other phenomena we try to explore and exploit. The choice of scope or applicability of the model and its internal structure are important design variables. These variables largely determine reliability, robustness and efficiency of the model [Meijer 1998].

1.2 Is there a limit to complexity?

If markets expand like the universe, if product and process diversity increase autonomously, then complexity like entropy will increase continuously. But is it real, this idea of ever growing complexity? The assumption is that we cannot control the expansion of markets and the growth of product diversity. A closer look reveals that companies as well as customers make choices that at least for a while limit complexity to what they can handle. We can limit our scope of the world to a smaller steady state system that is well insulated and well protected by (human made) constraints.

Marketing often promotes to increase product diversity as a means to increase marketshare by offering more choice. But too much diversity could also harm business. Cannibalism is a well recognized problem. Distribution constraints as well as customer constraints, may easily be overlooked. Procter & Gamble have learned in the 1990's that offering distinct diapers for boys and girls did not give them more shelf space in the supermarket. In stead of an increased turnover through more diversity, a decrease in turnover resulted from stock keeping problems in the supermarkets[1].

Unlimited choice is not always appreciated by the customer, because it assumes explicit awareness of all the qualities offered. Marketing and sales may respond by

1. This result follows from queuing theory with finite buffers. Operating at a load of 0.99 while halving the shelve space from 24 packs to 12, increases the stockout probability from 4.5% to 8.2%. In practice, because of shelf space discretization problems the loss may be even worse.

combining options into package deals, lifting the burden of having to evaluate the price of every other option and reducing product diversity[1].

“One stop shopping” is a recognized efficiency target for a customer. However it is unlikely that this customer will appreciate the “shopping experience” if buying daily needs becomes a quest. Moreover some supermarkets used to change their layout regularly in order to seduce customers to buy things outside their regular shopping's. This is a deadly sin in the eyes of the efficient customer. What we see today in supermarkets in the Netherlands is dedicated shelf space for weekly “specials” in addition to fixed product specific shelf space; easy for the customer as long as supermarket logistics can keep up with keeping both product locations filled.

In advertisement campaigns, marketers are often selling us role models for their products (“the choice of a winner”). This policy is not aimed at offering unlimited choice. The aim is to sell us the one thing that winners apparently have. We like to be associated with our winners. Thus marketing and sales have means to reduce the diversity of their offers.

In markets where product diversity is high and where the transparency of distinct product offers is limited, intermediates and value added resellers find a new business to help customers with their product choice. From the customer’s perspective, the complexity of finding the right product is exchanged for the (reduced) complexity of finding a trustworthy intermediate. This is a system boundary change.

Both the Procter and Gamble case as well as the customer responses to unlimited choice indicate that product diversity may only work if the diversified products are competitive in their own niche and if there are no logistic constraints either in the distribution or with the customers that prohibit sales and delivery of the products or services.

1. E.g. compare the product offers of Asian and European middclass car manufacturers on the European market.

Maybe there is a natural bound on complexity, beyond which the efforts associated with dealing with complexity become unacceptable. As a result new ways of dealing with complexity are being developed that cause a drop to a more comprehensive level (figure 1.1). Complexity defined as effort, is often associated with cost and risk.

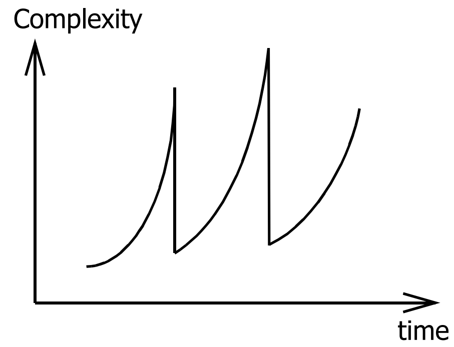


Figure 1.1: Complexity over time

A high risk is often rewarded with a high margin. High risk and high cost of existing solutions offer an incentive for innovations and rationalizations of the existing product offerings. New services may enter the marketplace hiding a lot of complexity from customers against a reasonable fee. A product innovation with a price/ performance breakthrough may cause a drop in complexity; a diversified product is replaced by a new less diversified concept serving all customers. Low end, less sophisticated, "back to basics" products may (re)appear into the market if the more sophisticated products have become rather expensive. Surprisingly, some of these "back to basics" products are manufactured and sold by new and lean organizations, the existing diversified organizations cannot compete with any more[1].

1.3 Complexity, complex systems and order

In section 1.1 complexity has been defined as a quantitative notion proportional to effort. Other researchers have studied systems and the existence or order or regularity to capture complexity. Their notion of complexity may be different, but not

1. E.g. low-cost airline companies with direct sales infrastructure (Easyjet, Ryanair).

contradictory to regarding complexity as a measure proportional to effort. If complexity is measured along a scale of order, then low order means high complexity, corresponding to a high effort to comprehend and predict. Complex system models and their transitions from apparent chaos to order are often explained in terms of structures and mechanisms that require minimum effort or minimum energy to survive in a continuously changing environment. In this section some of these complex system theories are introduced briefly, because there are many analogies between these complex systems and the systems managers face and interact with.

Stacey studied the concept of chaos and complex adaptive systems in the context of creative processes within organizations. Most organizations have a dual nature. Superficially they present themselves as real legitimate hierarchical systems with clear boundaries. This system is bureaucratic, stable and not responsive to changes in the environment. The complementary or shadow part of this organization is the informal part which is self-organizing to a large degree, has no clear boundaries and which is largely responsible for the capacity of the legitimate system to change. However this dual layered complex adaptive system can just as easy degenerate into responses that expose corruption, vicious personal striving, harmful politics as into responses that are desirable and essential for learning and adaptation [Stacey 1996].

Prigogine became convinced that the irreversibility or order at macro level is the manifestation of randomness at micro level. However, this irreversibility may result in different emergent orders. Order emerges at far from equilibrium states if there is also a dissipative mechanism that forces the system to evolve in only one direction. This is the irreversibility. However there is no theory to predict which of the possible structures is preferred. The random fluctuations together with the dissipative mechanism are determent for the evolution about to happen. At equilibrium universal laws of physics rule, but far from equilibrium specific mechanisms determine order that can be described in terms of probability distributions of ensembles. A collection of unstable interacting elements can as a group show stable statistical properties that at macro level are being observed as order [Prigogine 1997].

Several scientist at the Santa Fe Institute are devoted to unravelling models for the origins of life that could help us to understand order and complexity in many situations in life, including economy. Well known Santa Fe scientists like Stuart Kaufmann and John Holland work along similar lines of thought as Ilya Prigogine, however their models are different. Prigogine's models are mostly inspired by models from physics. The models of John Holland and Stuart Kaufman mainly originate from biology and genetics.

John Holland is well known as the father of genetic algorithms, a very powerful mathematical problem solving formalism that mimics the breeding processes, fitness and survival in biology [Holland 1975]. Another generalization or class of mechanisms that can exhibit emergence, developed by Holland is named constrained generating procedures (CGP's). The CGP's unify models as cellular automata, neural networks and others [Holland 1998]. In this thesis the structure of genetic algorithms has been used to propose a concept for an evolutionary organization for solving complex problems. The CGP formalism has not been used explicitly in this thesis although it is recognized that the possibility of recursively using the formalism has great potential in generating complex behaviours.

1.4 Complexity drivers and Shannon's measure of complexity

Although there are many similarities and relations among the theories described above, there is not (yet) a universal notion of complexity and complex systems that is acceptable to all branches of science. From Horgan's 1995 article in Scientific American, one may even doubt if the quest for a universal theorem on complexity will ever end. Even within branches of science the discussion on the nature of complexity has not finished yet. The statement "the complexity is in the eye of the beholder", points out that the understanding of the concept of complexity is very much driven by the perspective of an individual human being and scientists are no exception to that rule. These perspectives capture complexity as a figure, complex systems as well as mechanisms that explain how complex behaviours may evolve or emerge. Yet most complexity definitions involve concepts such as combinatorics (elements and relations), structure and uncertainty [Horgan 1995].

If we look at these commonalities, we can conclude that complexity drives effort or energy to accomplish a particular goal and has a number of aspects that can appear isolated or in combination. These aspects are:

i the number of objects or entities and their relations.
ii the structure that emerges from these objects and their relations.
iii the nature of these relations (static, dynamic, stochastic)
iv uncertainty over i, ii and iii.

The drivers of complexity related to business processes, as mentioned in section 1.1, are diversity, interrelations and uncertainty. Diversity and interrelations are represented by aspects i and ii. Uncertainty is introduced, either through stochastic relations or through uncertainty over objects and the nature of their relations (aspects iii and iv). The complex systems and order perspective mainly comes forward in combinations of ii and iii and sometimes iv[1].

Shannon's Entropy is a quantitative measure for the amount uncertainty or the amount of information that is needed to either solve a problem in all of its aspects and interrelations or completely describe a system and its state [Shannon 1948].

1. E.g. Prigogine's dissipative structures that can cause microscopic behaviours to "freeze" into notably different behaviours at macro level, each with their own distinguishable entities, relations and modes of interaction

Shannon's information entropy is defined as follows:

Definition 1.2 Shannon's information entropy.

Given a state space S with n states and a probability distribution P for S, where $\sum_{i=1}^{n} p_i = 1$***, the information entropy H is defined as:***

$$H(P) = \sum_{i=1}^{n} p_i \cdot \log_2(p_i)$$

One should note however that the choice of entities and corresponding states are made by the researcher. How many entities to distinguish, how to choose state variables or how to map continuous variables onto a finite number of discrete states, depends on the intent of the researcher.

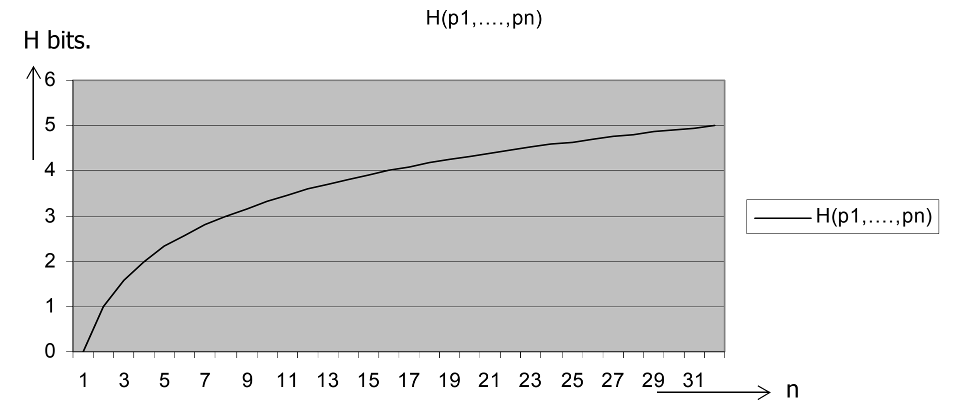


Figure 1.2: Entropy for uniform probability versus the number of states

If imposing structure means reducing the number of states to be considered simultaneously, then the entropy per set of nodes in the structure is also reduced. However the total entropy is not reduced. At best the total entropy is the same, if there are no interrelations between states. This means the probabilities are independent. Imposing a more fine grained structure generally increases entropy.

Maximum entropy corresponds to maximum uncertainty, which corresponds to a uniform probability distribution of the state space (figure 1.2).

For a more detailed introduction to Shannon's entropy see appendix A.

1.4.1 Uncertainty and Probability

Shannon's entropy measure is based on probability distributions. Applying Shannon's entropy function as a measure for complexity implies that probability is assumed as a suitable mathematical model for uncertainty.

Smithson has published a taxonomy of different types of ignorance [Smithson 1989]. In this taxonomy probability is indeed a model for uncertainty, but probability is not the only concept of uncertainty. Note that in this taxonomy, the distortion category relates to systematic errors, whereas incompleteness refers to the stochastic category of errors (figure 1.3). In case of vagueness or ambiguity, variants of

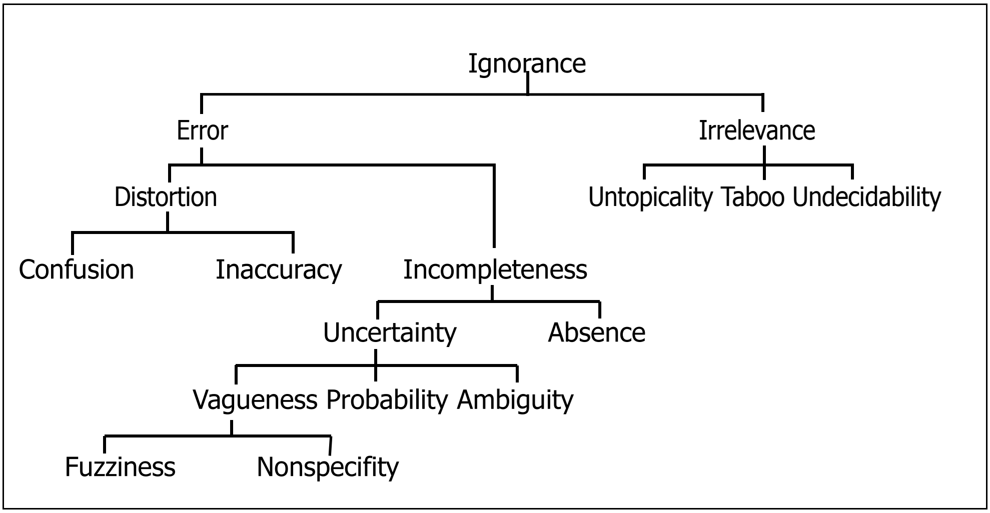


Figure 1.3: Smithson's taxonomy of different types of uncertainty

probability theory, such as evidence theory [Dempster 1968, Schafer 1976] and possibility theory or fuzzy sets [Zadeh 1978] may be used. However these methods also use normalization techniques that make them not fundamentally different from probability theory. Uncertainty is distinguished from 'absence'. Absence is sometimes referred to as incompleteness. More precisely, Parsons and Hunter put forward that uncertainty is generally considered to be a subjective measure of the certainty of something that can be treated numerically in a number of ways. Absence is the occurrence of missing facts, which is usually dealt with by essentially symbolic or logical methods. This dichotomy has lead to two categories of completely different formalisms for dealing with incompleteness: a symbolic category related to absence and a numerical category related to uncertainty. It is only recently acknowledged that all formalisms have their use for solving different problems. In fact the symbolic formalisms are important for defining the sets and states over which the numerical formalisms can be applied. Also hybrid methods are being developed [Parsons 1998, Smets 1999].

The discussion on uncertainty or absence leading to incompleteness can be related to the discussion on open versus closed systems. Recognition of absence as being different from uncertainty makes the issue of open- or closed- systems more explicit. Probability theory by definition considers closed world descriptions since by axiom the exhaustive sum of probabilities over all mutually exclusive as well as joint possibilities are constraint to sum up to 1. "Opening" the system, effectively means changing the systems boundaries and re-calibrating the probabilities within these boundaries such that their sum still equals 1. Without the option of choosing a perspective outside the world, we can only assume completeness within the bounds of our knowledge and we cannot distinguish absence from uncertainty.

Using probability theory as a representation of uncertainty is a choice that should be considered in the scope of the problem to be solved. If we are able to define the problem within a bounded problem-scope and if modelling can be done with a countable number of elements and states, then it is possible to reformulate the original uncertainty problem as a closed probability distribution problem.

For such problems, Shannon's entropy measure as a quantitative measure of complexity has relevance. Problems of management complexity, more in particular management problems related to building and managing operations are mostly related to incompleteness of information. Incompleteness of information can be dealt with through probability theory or combinations of probability theory and some symbolic formalism. Shannon's entropy measure offers the freedom of choosing the state space as well as problem scope in which the states are considered. Thus a discussion on the validity of the use of Shannon's entropy as a measure of complexity is essentially a discussion on the relevance of the states modelled and the choice of problemscope for the problem to be solved.

1.5 Problem definition, organization design and scope

Companies are facing complexity in operating their business processes. Complexity is driven by uncertainty and diversity of- and interrelations between- various aspects. In order to satisfiably serve their customers, companies need business processes operated and supported by an effective organisation design.

Starting point for this research was the observation that the Delft School of Organization design, founded by Prof. Jan In't Veld and Prof. Pierre Malotaux, is based on a still unique doctrine about organization design:

By starting with the design of business processes and structuring these with the intent to facilitate quality management (customer value) and to optimize productivity, a structure can be found that serves as a basis for the design of a department and management structure. Thus an implicit and natural match is accomplished between the quality control and management requirements from the market and the management and control capabilities of managers and directors.

This match between process management requirements and management capabilities prevents many of the induced uncertainties and unnecessary coordination that could frustrate workers and managers if this match is not accomplished. Since most (re)organization processes start with reduction of head count and (re)distribution of management power without even looking at business processes, a mismatch is not unlikely.

Yet the In't Veld and Malotaux doctrine and their supporting theories on organization design were never validated in the context of reducing complexity. Moreover structure design decisions were based on a choice of structuring principles[1] that were not directly linked to business process control requirements.

The goal of this thesis is to expand the theory of In't Veld and Malotaux with:

- *Tools to identify and resolve sources of uncertainty, diversity and dependencies to support the design of business processes.*
- *Develop structure design rules, linked to business process requirements, that accomplish a natural match between business process control requirements and management capabilities.*
- *To validate these tools and design rules for their potential to reduce complexity and to demonstrate with cases their applicability in practice.*

Although this thesis contains some excursions to service processes and applications outside the scope of industry, the main part of the thesis has been developed with industrial business processes and their organisations in mind. Applicability of the theories in this thesis to other organisations such as not-for-profit organisations, political organisations or public bodies is not ruled out but validation and the development of tools for the analysis of the value propositions of such organisations is considered outside the scope of this thesis.

1.6 Reduction of complexity in practice; Easyjet.

The low-cost no thrills airline company Easyjet has successfully concurred their place in the airline business by doing almost the opposite of the established companies. How was this possible? The answer is in how Easyjet controls complexity.

The start is ticket-sales. Easyjet does direct sales through a call-centre or internet, thus preventing a rather high fee of 15% of the ticket price that travel agents were used to receive[2]. Rather than offering x-classes and a complex system of refund and

1. grouping of functions on the basis of functional similarity, product orientation or geographical location.
2. As a result of Easyjet direct sales practice, other airline companies have succeeded in reducing travel agency fees as well.

change conditions, Easyjet offers a low flat rate for the first 60% of the seats in the aircraft. Selling these seats is almost sufficient to cover all cost of flight operations. Beyond 60% apparently there exists a special demand for that flight and ticket prices rise gradually until the last chair is being sold at approximately twice the flat rate for the cheapest chairs. Yet, this fare of twice the top rate is still more competitive than the one day return rates offered by traditional companies. There is no refund policy, but in case of being late a booking on the next flight available is offered for the difference between the fee paid and the available rates. Thus Easyjet knows exactly how much money is earned from each flight[1].

The product offered has little or no diversity. Easyjet does not offer a network with transfers but only a large set of point to point connections with a no thrills service. There is no complex catering operation. The only in-flight service offered is sales of drinks and snacks against standard prices that require as little change as possible. For passengers this is not a problem for flights to destinations that at most take 2 hours. Offering only point to point connections also means no liabilities for missed connections nor waiting for transfer passengers arriving late from other delayed flights. Handling of checked luggage is also simplified, since there is only one routing for all checked luggage of a particular flight. Yet, Easyjet customers have enough flexibility, because they are only buying one way tickets with clear conditions and Easyjet is offering three to four return flights a day on most of their destinations. At 30 September 2005, Easyjet operated 212 routes between 64 airports.

Easyjet keeps the number of different aircraft types as low as possible. At the end of 2005, Easyjet operated 55 Airbus A319, 32 Boeing 737-700 and 22 Boeing 737-300. The Boeing 737-300 series aircraft will all be phased out of service by 2007. The

1. There are no revenue leaks to other companies because there are no ticket exchange options with other companies. One has to realize that with a traditional airliner a business class ticket is a fully refundable option for air transport. These companies only earn the money if the option is called for and the passenger enters the aircraft.

Airbus A320 series aircrafts came into service from 2003 because they were larger and cheaper and allowed a shorter turnaround than the Boeing 737 was fit for.

In case of a temporary shortage of aircrafts, Easyjet has a policy of hiring on the basis of full wet lease, which means the aircraft is hired with crew and fuel, to keep the interference with parts of its own operations as small as possible.

	Product	Process
Reducing uncertainty	• *transparent fare system* • *reward advance bookings* • *single set of sales conditions*	• *direct sales / e-ticketing* • *no transfer passengers*
Reducing diversity	• *standard seating* • *standard in-flight services* • *only one-way single flight tickets*	• *standard aircraft* • *standard crew*
Reducing Interrelations	• *Point-to-point network operated from 16 base airports* • *no refund policy* • *separate sales of in-flight services*	• *no seat allocation process* • *no transfer luggage handling* • *dedicated aircraft maintenance facility* • *no revenue transactions with other airliners.*

Table 1.1: Reduction of complexity by Easyjet.

Easyjet operations are organized from 16 so-called base airports where a single type of aircraft as well as crews are stationed. From each base, every aircraft is capable of flying every service and each crew is fit for flying every service. Thus Easyjet has full flexibility on allocating aircrafts to services, staffing the aircraft and in addition maintenance and turnaround operations at each base can be standardized as well.

Easyjet preferably operates from smaller regional airfields in the neighbourhood of cities. These airfields are cheaper and the size of their operations allow faster turnaround. Shorter turnaround times, means more flying hours can be scheduled per day. As a result, the average operated aircraft utilisation with Easyjet was 11.7

so called block hours per day in 2005 whereas the total utilisation over all aircrafts owned or leased averaged at 10.7 block hours per day [Easyjet 2005][1].

Together with FLS Aerospace, a large commercial aircraft maintenance vendor, Easyjet established a joint venture named Easytech. Easytech is a dedicated line maintenance facility for Easyjet aircraft [Szurovy 1999].

1.7 Outline of this thesis

The goal of the research presented in this thesis is to develop and validate rules and tools for imposing a structure over a set of business processes. These rules and tools should result in processes that are less complex to operate and to manage. In other words the structure should support quality and reliability against lower cost.

In chapter 1 the problem field as well as the notions of complexity that will be used in the remainder of the thesis are introduced.

In chapter 2, value creation is introduced as the main driving principle for starting and sustaining businesses. A design methodology is presented for business processes design as well as a set of rules for designing a structure for these processes and it is demonstrated that this method generally supports complexity reduction in the context of Shannon's entropic measure of complexity.

In chapter 3 quantitative support for reducing complexity has been developed. A more practice oriented reader of this thesis may skip this chapter.
Shannon's complexity measure and theories from information theory have been used to demonstrate the effect of problem structure decisions on the complexity of a problem (sub)set and it is argued that the coordination need between problem (sub)sets can be modelled by the conditional information terms of the aggregated problem. By studying these terms quantitatively, it is possible to prioritize the order in which the (sub)problems need to be solved.

1. The average aircraft utilization of British Airways was 9.8 hours per day in 2005 (source http://www.bashares.com Annual reports Form 20F-2005.)

Logistic simulations and queuing theory have been used to demonstrate the effect of inhomogeneity of logistic requirements on leadtime performance. This study provides additional and more detailed logistic support for the first structure rule about splitting processes.

In this thesis, design methods have been very important as a frame of reference for this research. Yet when compared to manufacturing processes, design processes require more inter-disciplinary interactions and operate more concurrently than manufacturing processes. This nature calls for different ways of organizing and structuring. In chapter 4 the concept of knowledge logistics and an evolutionary organization model for complex design and development processes are being presented as a basis for structuring complex research and development processes.

The methods presented in the previous chapters assume that humans will comply with the structures and processes developed with these methods. Chapter 5 on boundary conditions for organization design will present a cognition and communication model as well as discuss motivation and team work theory in order to demonstrate that it is possible to create conditions under which human workers will comply with processes designed with the described methods. Furthermore, the cognition and communication model sheds additional light on the problems of creating business knowledge as is also discussed in appendix D on research methodology.

Finally in chapter 6 conclusions and proposals for future research will be presented.

1.8 Research method

Despite the existence of comprehensive textbooks on various research methods in business administration [Arbnor 1996, Saunders 2000], business administration literature is dominated by research that has been carried out with surveys as the main instrument. However, conducting a survey among companies and consultants on how they perceive uncertainty and diversity in their business practice and how this affects the organization structure is not expected to result in answers to the design questions of this thesis. First of all uncertainty and diversity are not perceived uniformly throughout a single company, let alone in an industrial sector. Secondly, in many cases the existing organization structure is the result of a historical, evolutionary process under the influence of many temporary factors and actors that are not always documented. Thirdly, the handling of complexity is performed by business processes. The organization structure is a means to facilitate the effective and efficient execution of these processes. To identify this match usually requires in depth study of companies and their organization design methods that cannot be offered by a survey.

The survey research model (figure 1.4) puts the researcher in the analysis role only. The industry creates, the market validates and researchers try to understand afterwards what happened. Conclusions from such research are often either not specific enough to support better designs and practices or the market context is no longer valid when research finally understands the mechanism.

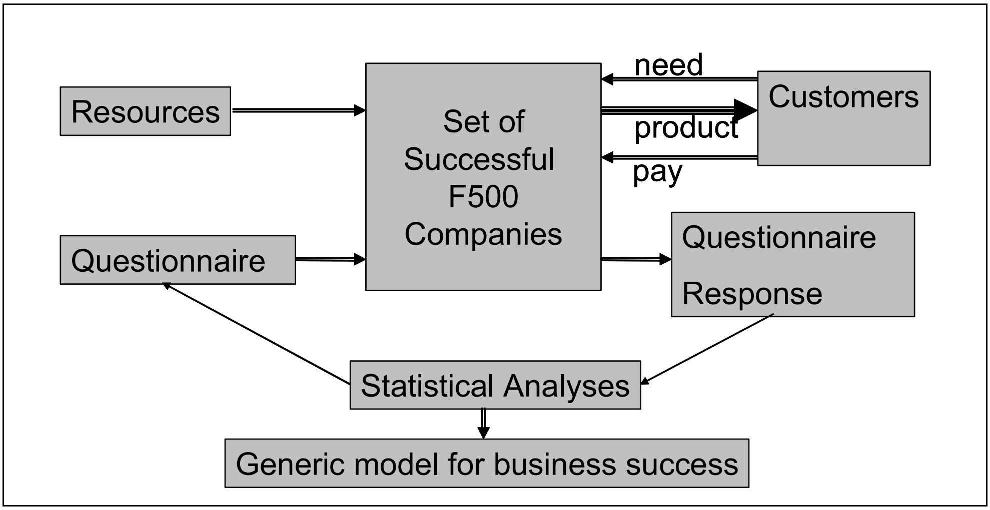


Figure 1.4: The survey research model

Henry Mintzberg noticed this problem long ago:

"Most of the contemporary literature fails to relate the description of structure with that of the functioning of an organization...... All of this is to say that the conclusions of the research often lack "context"- the type of organization and the part of it to

which they apply, as well as the relationships between the structure and the functioning of the organization. As a result, these conclusions often come across to the reader as detached from reality, devoid of real substance"
[Mintzberg 1979, p.12-13].

In other words; one needs to be a designer or an engineer with the ambition to write a prescriptive theory to be fully aware of what part of the context is essential to do the research and to correctly frame the scope of a resulting theory.

The work of Klaasjan Visscher represents a good example of the problems of studying design methods without practising design. Visscher studied the use of design methodologies in management consulting and concluded that his study did not offer one-best-way to design but four typical strategies and a meta-strategy to mix them productively. The strategies Visscher identified are far from a design process and this is no surprise. The ambition of consultants is usually constrained by what is considered achievable and "consumable" by their customers. Strategies and problem solving practices are driven by experience and skills in finding out what their customers can cope with and facilitating the process of achieving a result [Visscher 2001].
Unfortunately, the work of Visscher also illustrates typical problems of conducting surveys. One of such problems is how to select an unbiased population of consultants. Diversity in background was not a problem, but how to find best practices? The obvious answer is to interview the best consultants. Visscher selected the "best" on the basis of recommendations and awards. Often this means the best communicators in the field are selected. This is a biased population. Their work, the sustainability and degree of implementation of their advise has not been evaluated explicitly. A reference set with average performers was not reported either.

Engineers are trained to design useful theories and artifacts from scratch. The action research model is closer to engineers practice since it allows to actively participate in developing and implementing business practices. Moreover if science intends to develop theories and methods to build better businesses, then the action research model is a more promising approach. Rather than waiting for industry to develop and implement new business theories that may be analysed by academia, engineering science can create new theories pro-actively, that when put into practice may innovate industries. The theories, methods and models in this thesis are largely the result of action research, carried out by students under close supervision, supported by literature desk research and reasoning, including sources and ideas from disciplines well outside the field of management science.

The focus in this thesis has been on developing models and theories that support individual businesses to design their specific processes and structures that could effectively and efficiently face the complexity of their own choice. This calls for a validation practice similar to that of validation of design methods. Validation of the tools and methods in this thesis is done through description of some realized designs, through making explicit the logic behind the ideas and through using a quantitative meta theory on complexity based on Shannon's entropy measure. In a recent article on the validation of design methods, validation practices are compared to methods common to medicine. As in medicine, the use of various models, logic and reasoning is advocated in the validation process before double blind field tests (like surveys) may be considered as the final step in a validation process [Frey 2006].

The type of understanding and systems modelling as developed in this thesis is a necessary precondition for doing surveys and compensating for the many side effects and biases that may be present in the answers. But even then, given the influence of specific strategic choices, past performance and experience on the complexity experienced by a specific company, it is doubtful whether a survey over a number of companies can produce comparable observations that may result in statistically significant results which could not have been found using other methods.

A more elaborate version of these arguments can be found in appendix D.

2. BUSINESS PROCESS DESIGN AND STRUCTURING BUSINESS PROCESSES

Companies as well as non-profit organizations operate processes to create value for all of their stake holders. The owners are entitled to receive a return on their investment. The customers enjoy the value of products or services delivered to them in return of their pay. Suppliers receive money for their goods or services delivered. Workers receive wage for their labour capacity and finally society collects taxes for providing an economical climate in which the organization can flourish. Creating a sustainable business means creating a set of processes that can fulfil the customers needs under constraints from all other stake holders. Creating a structure for these processes is the next step that makes management of these processes effective and efficient. This chapter will present a theory for creating sustainable business processes and structuring them in order to become effective and efficient. Structure design is the key step in this theory which also distinguishes the Delft School of Organization design from other theories on organization design. The structure design process is based on a set of five generic rules for taking structure decisions. Using Shannon's entropy as a quantitative measure of complexity, it will be argued that these rules are effective in lowering complexity.

2.1 Organization Design

Organizations primarily exist for humans and they are operated by humans. Customers primarily look for organizations that can sell a product or deliver a service. Employees are often entitled to do a certain job, because they belong to an organization with enough credibility to the customer. The involvement of humans in product offers is crucial. Even in the case of machine operated or automated products or services, there are humans that exploit and service these machines. Humans as well as organizations are accepted as legal entity that can do transactions; machines are not. Automated products or services require near perfect quality standards and even then depend on human customer assistance to become accepted. Another factor is liability for errors. If humans fail, they can be addressed for their failures. If machines fail, the operator, the exploitant as well as the designer and builder have to be tracked in order to find out which human(s) can be held responsible for the failure.

One perspective on organizations is to consider them as a social construct, which emerges around management and decision making nodes from a temporal balance between the interests of various stake holders. Politics and decision making of public bodies often work this way. For such organizations the focus is on the design of the decision making process to fix the influence of all actors [Bruijn 2002]. The actual outcome, the decision, seems to be of less importance. Yet, the institutions that emerge from them are long lived because their function as well as their management structure are cast in concrete legislation and they often have a monopoly for their function. Changing these institutions is often very difficult because replacing them with new institutions will always violate the interests of many. As a result their growth seems autonomous. Changed external conditions, require modifications that can only be dealt with by additive measures, not by structural changes.

Another perspective is that organizations can be designed and redesigned in response to changing external circumstances. For companies, sustained losses, bankruptcy, mergers and takeovers are natural moments to reconsider all activities and the organization of these activities. For companies, making a redesign is easier than for public institutions. The number of stake holders is usually smaller and their interests are less diffuse. Internal consistency is important to maintain effectiveness and gain efficiency. Thus horse trading common in politics is less likely to occur.

In this chapter an organization design theory is presented that is more design oriented and that can be used for greenfield as well as for redesign studies. In paragraph 2.2 some theories for organization design will be described briefly, including Mintzberg's theory. Although Mintzberg's theory is mostly known and used for variant design from archetypical organization structures, his theory also contains an extensive list of criteria for function grouping or structure decisions. This list will be used later as a reference. In paragraph 2.3, the Delft School of Organization design will be presented. This theory is system and process oriented and allows both redesign as well as designs from scratch. Yet it lacks tooling in critical areas of the design process. The company strategy and value propositions that represent the program of requirements for the design are assumed to be known explicitly. The criteria for function grouping decisions were not generic but case based. In paragraph 2.4 sustainable value creation is introduced as the context for doing a design or a redesign. In paragraph 2.5 the design rules for structure decisions will be presented, followed by a logical validation in paragraph 2.6. For this validation, two reference theories are being used: Mintzberg's list of grouping criteria and the Business Balanced Score Card. Finally some short design cases will be presented in paragraph 2.7.

2.2 Some theories for organization design

Organization theory is an old topic. The ancient Greek philosophers Socrates and Plato already built theories on management, allocation of tasks and specialization of labour. In more recent times, Frederick Taylor introduced Scientific Management. Productivity could be raised considerably if task times in production were no longer based on never verified estimates. Taylor introduced time studies and time measurement techniques to introduce productivity targets based on real measured task time figures. Task specialization could make manual labour even more productive (the so-called, learning effect). This way of thinking however leads to organizations with departments that are centred mainly around the sharing of knowledge and tools.

Henry Mintzberg is known for his ideas on organization structures and strategic management. Mintzberg relates the organizational structure to strategy and puts the different schools of thought on these topics in perspective. The ideas of Mintzberg are being described in more depth in paragraph 2.2.1 [Mintzberg 1979, 1998].

Based on non-cooperative game theory, Os Shy presents a quantitative theory on Industrial Organization, which primarily deals with the structure of markets and the behaviour of companies in these markets. Understanding market structures and changes in market structures could support strategy formation and decisions over the internal structure.
For designing the internal structure Shy's theory, based on so-called non-cooperative game theory, is less useful. Moreover as the market structure results from behaviours of individual and independent companies, the structure itself is emergent

rather than designed. Even in understanding market structures the use of game theory still has its limitations. This argument will be put forward in paragraph 2.2.2 [Shy 1995].

Jan In't Veld who together with Pierre Malotaux, founded the Delft School of Organization Design, lists many factors to be considered in both job design as well as organization design. Their method is process and system oriented. The core of the method is the design of a so-called organ- or organelle structure for managing these processes. This structure is the result of iteratively fine-tuning the production structure and the control structure of the core process. The personnel structure is then designed on top of the organ structure [In't Veld 2002, p365-366][1]. Thus, human power and control is matched with controls that through the organ structure relate to market needs. In paragraph 2.3 this theory is described in more detail and it is argued that discrepancies between the organ structure and the personnel structure are a major source of management problems and leads to increased complexity.

2.2.1 Organization design according to Mintzberg

There are many ways to analyse and design organization structures and just by looking at existing structures, one may speculate about the intentions, culture and even history of organizations. Without knowledge on policies and customers it is not possible to value an organization design. Mintzberg relates different strategic management schools to organization design, so it may be safe to say at least that an organization serves a purpose in society and that the design of the organization should reflect the policy or strategy of the organization. This notion of contingency theory, that organizational effectiveness results from a match between situation and

1. Most references in this thesis to In't Veld are to the 8th/2002 edition of "Analysis of organization problems". Note however that the 1st edition of this book, containing most concepts and definitions used in this thesis, dates back to 1975.

structure was borrowed from Joan Woodward [Mintzberg 1979, p216]. Mintzberg recognizes no less than ten schools of thought about strategy formation.

	SCHOOL OF THOUGHT	PRINCIPLE PROCESS
Prescriptive	Design School	**Strategy formation as a process of *conception***
	Planning School	**Strategy formation as a *formal* process**
	Positioning School	**Strategy formation as an *analytical* process**
Descriptive	Entrepreneurial School	**Strategy formation as a *visionary* process**
	Cognitive School	**Strategy formation as a *mental* process**
	Learning School	**Strategy formation as an *emergent* process**
	Power School	**Strategy formation as a process of *negotiation***
	Cultural School	**Strategy formation as a *collective* process**
	Environmental School	**Strategy formation as a *reactive* process**
	Configuration School	**Strategy formation as a process of *transformation***

Table 2.1: Schools of thought on strategy formation [Mintzberg 1998]

The schools can be grouped into three main categories. The prescriptive category contains schools that are concerned with how strategies *should* be formulated. The descriptive schools are concerned with how strategies in practice *do* get formulated. The configuration school however combines or integrates the aspects and focal points of the other schools (figure 2.1). The Delft School of Organization Design could also be considered a representative of the configuration school.

Besides for strategy formation, Mintzberg is very much known for his earlier work on

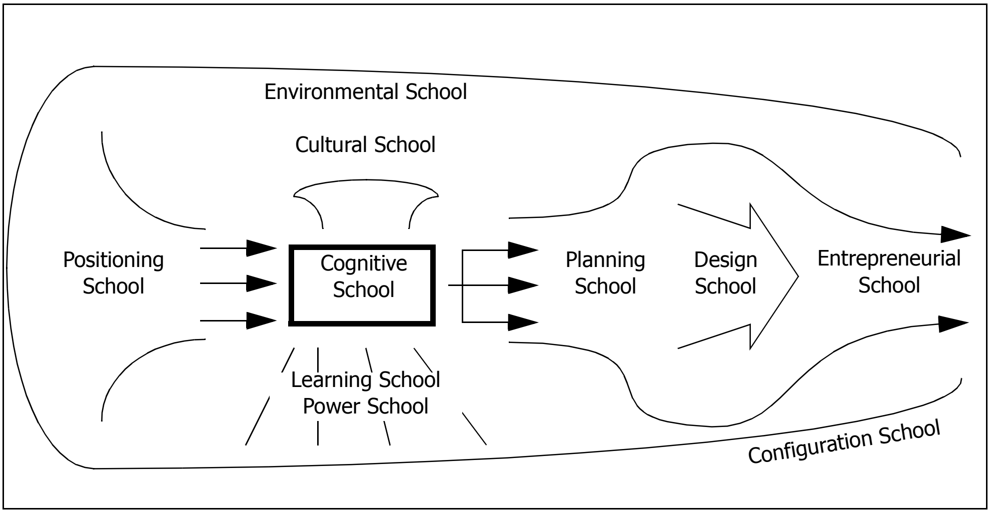


Figure 2.1: Strategy formation in process perspective [Mintzberg 1998]

structuring organizations [Mintzberg 1979]. The theory starts with a generic picture which contains the five basic parts of an organization.

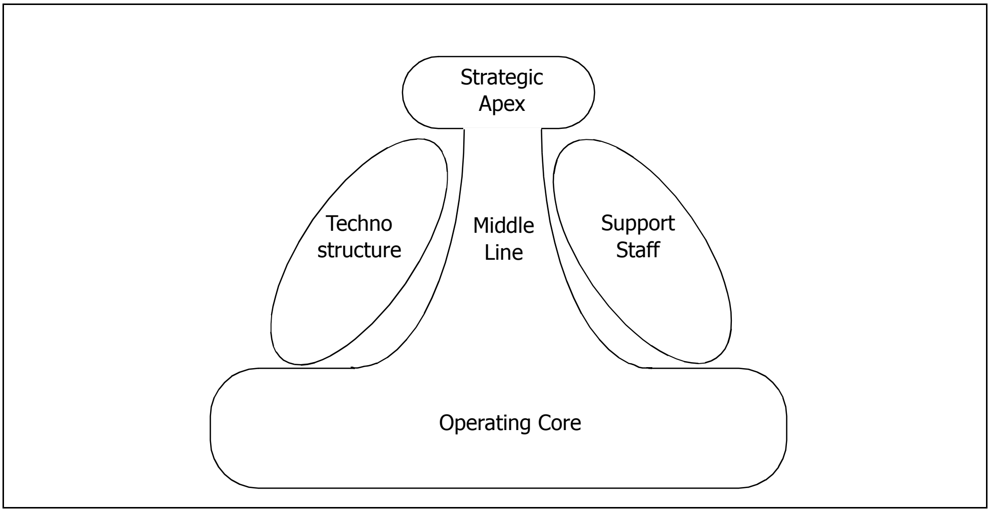


Figure 2.2: The five basic parts of organizations [Mintzberg 1979]

In this organization, the top management (strategic apex) is connected to the operating core through a "line" of authority which runs through the so-called middle management. All decisions are taken along this formal line of authority. The flanking staff units do not formally take decisions, but they support and advise the line managers and the top management who do. A total of seven archetypical organizations are being distinguished.

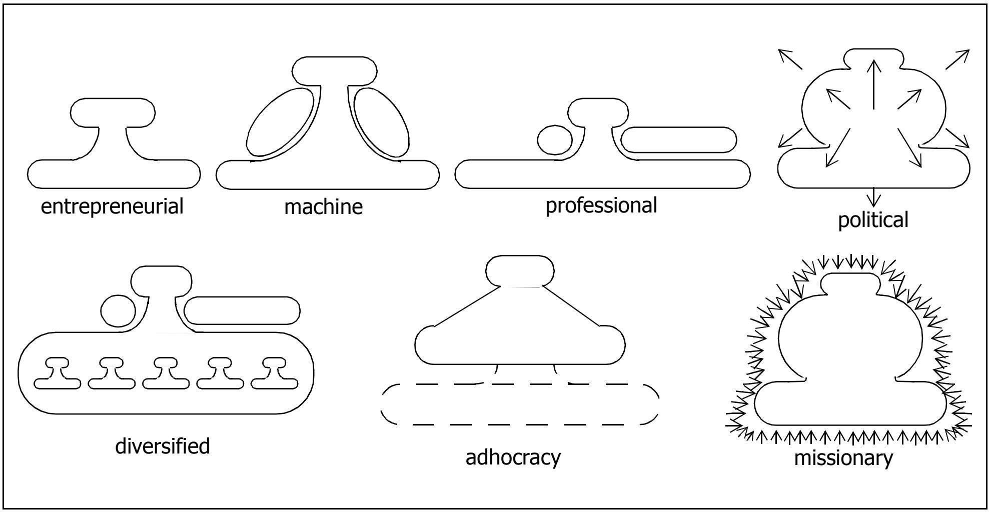


Figure 2.3: Seven configurations of structure and power [Mintzberg 1998, p307-309]

The five basic parts as well as the seven archetypical configurations are often "abused" in the sense that they are taken as reference and starting point for an organization design with desired properties. This procedure results in over complicated variant design and this method provides no guidance for organization problems deeper in the organisation. The problem is that the characteristics of table 2.2 cannot be linked uniquely or unambiguously to the external requirements and boundary conditions for the organization. Therefore the logic from archetypes to exhibited behaviours cannot be reversed for use in organization design procedures.

ORGANIZATION	CHARACTERISTIC
Entrepreneurial organization	**simple, small, flexible, usually young, not more than the boss and everyone else**
Machine organization	**highly programmed, well-oiled machine, is the off-spring of the industrial revolution, when jobs became increasingly specialized and work highly standardized**
Professional organization	**professionalism dominates, the work is mainly done by highly trained professionals, little or no line-management, some technical support**
Diversified organization	**a collection of rather independent units, coupled together by a loose administrative structure.**
Adhocracy organization	**contemporary industries have to innovate in complex ways; projects in which experts from different specialities are fused through mutual adjustment. With power based on expertise, the distinction between line- and staff diminishes**
Missionary organization	**organization dominated by a strong culture, its members are encouraged to pull together, and so there tends to be a loose division of labour, little job specialization, smaller distinction between line managers, staff groups, operating groups....**
Political Organization	**organization that tries to settle on no stable system of power, no dominant element. Conflicts tend to arise, characterized by pulling apart of the different parts**

Table 2.2: Seven configurations of structure and power.

Although Mintzberg does not explicitly advocate a process oriented design procedure for organization structures, his organization design theory is more process design oriented than most researchers acknowledge. This becomes evident when considering the extensive list of criteria for unit grouping and the effects of unit grouping (table 2.3). A process design perspective must have been tacitly present to consider the importance of mutual adjustments, work-flow inter dependencies, scale inter dependencies as well as considering time, output and work process as a basis for grouping.

THEORIES FOR GROUPING	
EFFECTS OF GROUPING	**establishes a system of common supervision among positions and units**
	requires positions to share common resources
	creates common measure of performance
	encourages mutual adjustment
BASIS FOR GROUPING	**by knowledge or skill**
	by work process or function
	by time
	by output
	by client
	by place
CRITERIA FOR GROUPING	**work-flow inter dependencies**
	process inter dependencies
	scale inter dependencies
	social inter dependencies (often factors related to safety, to prevent boredom,.....

Table 2.3: Mintzberg's theories for grouping [Mintzberg 1979, p106-124]

Furthermore Mintzberg also mentions other principles (basis) for grouping, such as by function or by market. Grouping by function is grouping by knowledge or skill, combined with work processes. Grouping by market is combining output, client and place. Structures based on grouping by function typically lack mechanisms for coordinating the work flow. Structures based on grouping by markets opt for work flow coordination at the expense of process and scale specialization. Regarding the size of units, Mintzberg forwards two hypothesis relating unit size to controls or coordination mechanisms: "*the greater the use of standardization for coordination, the larger the size of the work unit"* and "*the greater the reliance on mutual adjustment (due to inter dependencies among complex tasks), the smaller the size of the work unit".* Thus it is clear that Mintzberg considers various levels of control as important input variables for structure design and is explicitly aware of the effects of structure choices on the control of operations. This is a process design perspective as will be made more explicit in paragraph 2.3.

2.2.2 Industrial Organization according to Shy

A quick test on google-scholar (http://scholar.google.com) reveals that the term "Industrial Organization[1]" is most often interpreted as the organization of industry. Thus it refers to the structure of markets and relations between companies, customers and market regulators such as anti-trust legislation and GATT (General Agreement on Tariffs and Trade). This is also the main interpretation of Oz Shy in his book "Industrial Organization". Shy based his work on game theory. Game theory postulates rational behavior of all players in a game and defines the most likely outcome of a game as the outcome where the benefits or utility for all players is maximum. The concepts and notions of game-theory are explained in more detail in appendix B.

Given Shy's interpretation of industrial organization, game theory could provide a theoretical base for a structure theory. With their roots in probability theory, building a link between game theory and complexity theory should also be possible. In this thesis this path was not taken for two reasons:

1. *Most internal organizations are based on principles of coordination , cooperation and teamwork. Modelling this type of behaviour with game statistics is considered undesirable and inefficient.*
2. *Solving game-theoretical problems is often based on the assumption of transparancy of information. In real economics this assumption is often not realistic.*

---

1. Our research group was also named Industrial Organization since its start in 1968. But this was interpreted as Organization in Industry.

The second argument is given some more thought in the remainder of this section.

Although the use of game theory is popular among economists, there are serious limitations to this theory that affect its practical use by companies. The outcome of the game and the existence of an equilibrium solution is sensitive to the definition of the utility or payoff functions. However in reality these utility functions are not always known. Buyers as well as sellers also include intangible factors in their decisions, which are not easily accounted for in the utility functions. As a result, buyers as well as sellers may exhibit seemingly irrational behaviour which does not maximize the assumed utility.

Even if the modelling of a price game is correct, the equilibrium strategies may be doubtful. This argument is brought forward by Thomas Nagle and Reed Holden. They argue that in case of positive sum games (every one is a winner) playing is the basic strategy. However the standards by which payoff is defined may differ from player to player and this is not properly modelled in game theory. In case of a negative sum game, which often occurs in real economies, not playing (not engaging in a price war) but diplomacy (convincing others not to engage in a price war) is a better move [Nagle 1995, p117-118].

Another problem is the issue of completeness and transparency. A game, especially a zero-sum game, is a closed world model of reality. All players are known as well as all actions and pay-offs available to them. By definition, the Nash equilibrium requires that all players know all options as well as actions played by other players. In other words all players share all information. This transparency requirement is in contradiction with the axiom of economic activity. Economic activity is the result of transactions whose exchange in value is driven by differences in perceived risk. These differences are driven by differences in knowledge and skills. The more knowledgeable or skilful a buyer is, the more precise is his assessment of the price tag.

Differences in knowledge or skills create better opportunities for trades people as well as manufacturers to sell with a good margin, which may still be fair trade[1]. If transparency in information is the ideology for fair competition, then one may question the desirability of transparency both from the perspective of manufacturers, suppliers as well as from the perspective of customers. In transparent markets, the problem is planning the demand and controlling the number of suppliers and manufacturers. Non-transparency serves as an entry barrier that limits the number of suppliers and manufacturers. With a limited number of suppliers, the consequences of overcapacity may be regulated through coordination and diplomacy.

Without the entry barrier new parties may enter and overcapacity is a big risk. Overcapacity will trigger the start of a price war. At first this may seem good for the customer. But if the customers are not willing to raise their consumption to the level of the excess capacity then some suppliers will go bankrupt or move out by shutting down factories. The survivors are forced to cut manufacturing cost through product standardization and reduction of services. The end result is often a reduction in product-variety and in quality and service. This is not necessarily in the interest of the customer.

2.3 The Delft School of Organization Design

Jan In’t Veld together with Pierre Malotaux founded the Delft School of Organization Design. Their method is process and system oriented. The methodology has been developed and tested through over 400 master thesis projects in industry for over 37 years. Although this method initially targeted industry and although its applicants were usually trained as mechanical or aerospace engineers, the applicability is by no means limited to industry. The methodology has been tested successfully in areas

1. Note fair trade is also a cultural and ethical notion. In this thesis fair trade means, sharing revenues such that sustainable supplier relations are possible. Thus a knowledge differential may be exploited but should not lead to non-sustainable exploitation.

outside the engineering world among which health care, banking and insurance, courts of law and government organizations.

Problem solving

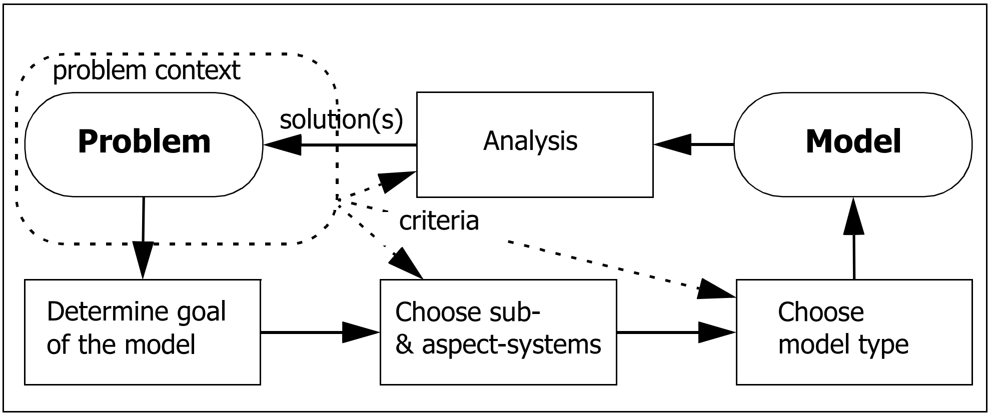


Figure 2.4: Problem solving using a model

In solving problems researchers as well as practitioners use models. For research as well as for business, the choice of goal (ambition), the perspective (reference knowledge) and scope (what factors to take into account) affect the perceived complexity of the resulting model. The goal of modelling is to develop a model of the system we try to influence that is sufficiently accurate and robust to analyse and support decision making within the context of our predefined ambition (figure 2.4).

In this problem solving framework, In't Veld defines a system as follows:

Definition 2.1 System [In't Veld 2002].

A system is a set of elements, distinguishable within a complete reality, selected by a researcher in compliance with a goal set by the same researcher. These elements have relations with each other and they may have relations with other elements in the complete reality

Elements have relations with other elements of the same system and they may have relations with other elements in the complete reality. In this definition the notions system and system-model are treated as equivalent, yet there is also a reference to the context called complete reality.

At the core of the systems modelling and design methodology is the so-called "steady state model" that represents a function-template for a unit that is capable of achieving one objective through reliable and repetitive execution of a transformation[1]. The "steady state model" has been developed into a more schematic form by Veeke [Veeke 2003].

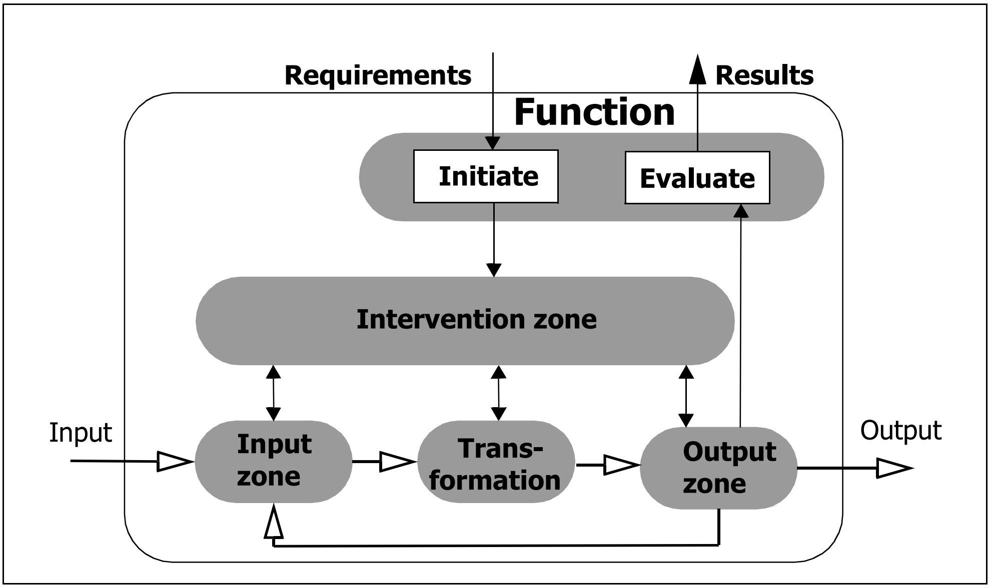


Figure 2.5: Schematic steady state model [Veeke 2003, p.23].

As In't Veld writes, his work shows similarities with that of Checkland. For one thing, Checkland too stresses that a system is a representation or mapping of reality. The applicability of this mapping is limited by the goals and intentions of the researcher that built it [In't Veld 2002] [Checkland 1993]. In the next section this interaction between goals and intentions of the researcher and the models built is discussed in more detail.

1. A transformation is a value adding activity that transforms materials and/or information into a shape suitable for the next transformation step or for the customer.

For reducing the modelling complexity, depending on the modelling goal, sub-systems of a system as well as aspect-systems may be distinguished.

Definition 2.2 Sub-system [In't Veld 2002].

> ***A sub-system is a chosen subset of the elements of a system, maintaining all the relations of those elements belonging to the sub-systems.***

Usually (sub)systems are identified on the basis of functions, that is a systems contribution to its environment.

Definition 2.3 Aspect-system [In't Veld 2002].

> ***An aspect-system contains a specific subset of the relations of the system while keeping all the elements.***

Aspect systems reduce the number of relations or influences to be considered simultaneously. Studying aspect systems often helps to identify functions that could serve as a basis for defining (sub)systems. Considering systems, aspect- and sub-systems for the purpose of analysis means setting boundaries in the problem domain. Setting these boundaries should be done with great care. The next section contains a reflection on the problems of setting boundaries or considering aggregates in problem solving.

2.3.1 A reflection on reality, system model and system boundaries

If we consider, markets and companies as sets of interacting systems, then we can observe that the boundaries between these interacting systems, both inside and outside the company are being changed regularly as a response to increases in complexity. Changing system boundaries means changing the context of problems that appear within these systems. As we will see, choosing the context of the problem is a powerful means to reduce complexity.

In problem solving, the perspective and knowledge of the problem owner have a big influence on perceived complexity. Some researchers even claim that complexity exists only in our imagination. Humans can only know reality through experience of facts about their own interactions with reality. Observation of others interacting with some reality is perceived and interpreted in notions or truths that are part of their personal experiences. Thus, both system and system model are equivalent concepts referring to a coherent image of our understanding of the logic of our observations of interactions with the world [Wittgenstein 1922].

However this equivalence is lost as soon as we consider sub-systems and aspect-systems. Although the definitions of sub- and aspect-systems seem complementary, in practice they are not. A sub-system also selects aspects since the elements in the subsystem may not have relations of a particular type (aspect), whereas an aspect-system is considered to maintain all elements with only a subset of the relations. If elements belong to a system or a sub-system, we have to assume connectivity (relations) between the elements. It does not make sense to consider an element belonging to a system if it has no relation whatsoever with the other elements of the system. Thus considering aspect-systems also means considering a subset of elements if there are elements not connected to this aspect. As a consequence, only under constraints or restrictions defined by the modelling purpose, aspect-systems may exist. The same restrictions are then responsible for the difference between system model (aspect system) and the system (our model of reality).

Choosing an aspect helps the researcher to focus on specific types of interactions between elements. Examples of aspects are control, social interaction, finance, technology, etcetera. Choosing an aspect only helps to produce less complex models if this aspect uniquely supports functions that can be identified when considering (sub-)systems. If this is not the case, then the neglect of other aspects supporting a particular function may introduce rather than reduce the uncertainty over interactions between systems. In other words, only if we can use functions to choose sub-systems and if aspects coincide with the means to realise those functions, then we can develop sets of models of interacting systems where these sub-systems can be studied in isolation of the other systems[1].

The criterion for setting boundaries and considering (sub-)systems comes from setting a goal for system modelling. Defining a goal implies a sufficiency criterion for deciding over the level of detail needed and for deciding what elements belong to the system and which ones can be left outside. Often this comes down to limiting the problem scope at a boundary where the relations with the environment exchange minimum information and can be considered quasi-static. Quasi-static is a relative notion, which can be defined as follows:

Definition 2.4 Quasi-static influences

Quasi-static influences on a (sub)system are influences that do not interact with the internal dynamics of the system significantly while maintaining the quality of the contribution of that system to its environment.

Quasi-static influences can be taken as a static input to the system or as a constraint.

Thus we can define two types of boundaries in modelling: the problem-boundary and the (sub)system-boundary.

Definition 2.5 Problem-boundary

Given our goal to influence or control a system predictably and with sufficient accuracy, the problem boundary is set at a scope where all significant influences considered outside the problem boundary, are either insignificant or can be modelled as quasi-static inputs to the system.

1. The influences of other systems are then modelled as input signals.

Definition 2.6 (Sub-)System-boundary

> ***Given our goal to influence or control a system predictably and with sufficient accuracy we can set (sub)system boundaries within the problem boundary to collect elements and their controls to study their joint interaction with and contribution to other (sub)systems to be considered within the problem boundary.***

Decomposing a large system into smaller (sub)systems is mostly done through considering functions that uniquely contribute to the desired combined functionality of the larger system. Choosing aspects may help to identify these functions. But we have to be careful. The aspects themselves need to be capable of defining and describing all relevant modes of interaction between the elements. An example from the area of assembly research can illustrate the difference. If we consider the aspect of mechanical forces that parts can exert on each other during an assembly motion, then the set of parts (elements) involved in the model are the parts that make contact during this motion. With this notion of contact it is assumed that without physical contact, defined as co-location of surfaces, force cannot be transmitted and the assembly motion is not influenced by parts not in contact with the assembly. However, this notion of force and contact cannot be scaled towards micro scale or even atom level. A more general aspect is exchange of energy. Any element that is capable of exchanging energy with another element is capable of influencing the motion of -and strains experienced by- those elements. At macro level this aspect system may be reduced to the exchange of mechanical energy through contact, but only after taking into account the modelling goal and the scale.

For business processes there are no general conservation laws that give us single aspects that have the same modelling impact that energy has for technical or physical systems. However if we define relations as the influence one element may have on other elements and vice versa then it is possible to define a list of aspects defining or describing that influence.

When defining aspect-systems for the sake of analysis, we have to verify that all other influences or interactions outside the scope of the aspect are either insignificant or quasi-static. Influences deemed insignificant means we have set a constraint or a boundary condition for that influence. The aspect-system will be valid as long as these constraints are not violated. The aspect-system will loose its validity if these constraints are violated and thus the other aspects are forgotten.

2.3.2 Bikker's Business Process (Re)design

With the problem- and system-boundaries defined and an understanding of aspect- and sub-systems, it becomes easier to understand the business process (re)design scheme that Bikker proposed [Bikker 1994] (figure 2.6). A (re)design is the result of two lines of analysis: the analysis of policies and their development and the analysis of processes and structures that will result in the design of the new situation. The analysis of policies was named, the 1st main line since it provides the policies, goals and constraints of the organization and thus defines the scope for the analysis of the processes and structures, named the 2nd main line. To match both lines of analysis several structures are being analysed or designed and evaluated against requirements specified as key performance indicators. For the analysis of both lines

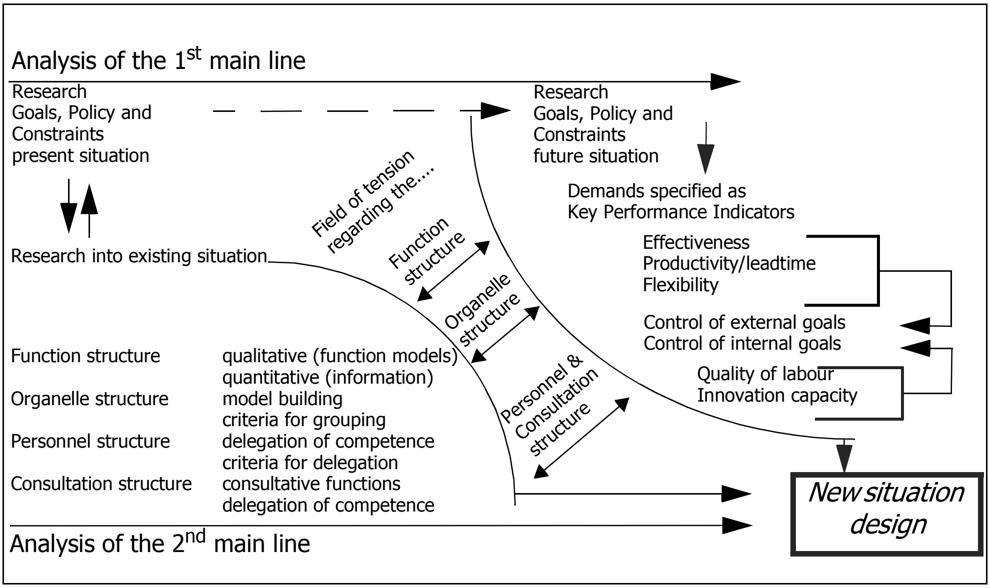


Figure 2.6: Business Process Redesign [Bikker 1994].

and for the transition between the existing- and the new-situation, the problem-solving framework of figure 2.4 may still be helpful.

Later a more aggregated version of this scheme was developed (figure 2.7). The

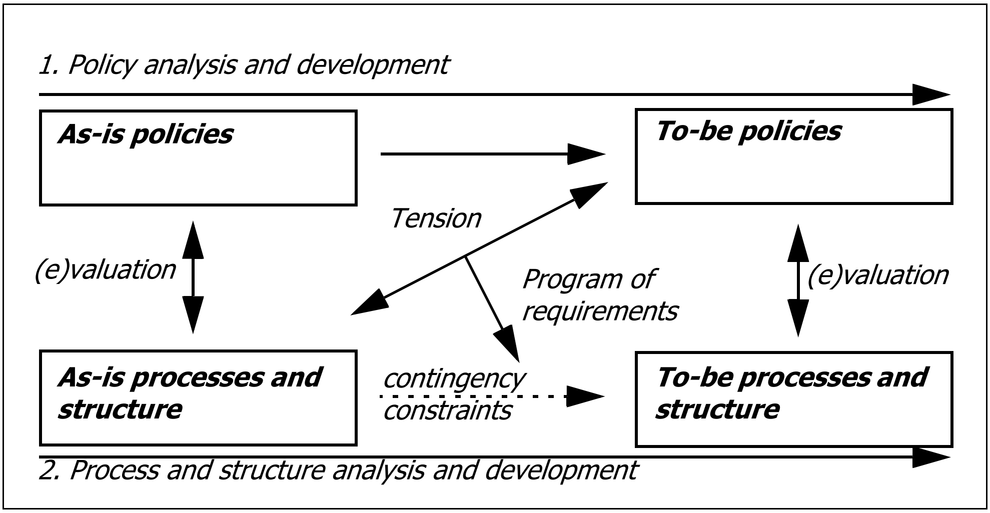


Figure 2.7: Business Process Redesign [Bikker 2001]

aggregated version revealed more explicitly the modelling of processes and structures for the analysis as well as for the (re)design. The 2nd main line of analysis is now named explicitly the analysis of processes and structures. Also it became clear that the use of key performance indicators for the evaluation of the match between the policies or requirements and the processes is important both for the present situation (As-is) as well a for the future situation (To-be). A more detailed discussion on performance indicators as well as on process design and structure analysis will follow in the coming sections of this paragraph.

2.3.3 Analysis of policies, goals and constraints

The main purpose of the organization design is to design business processes and a structure that support sustainable value creation. The basic paradigm of this method is simple as well as effective. Processes create value. The process performance is managed through monitoring and control of a set of key performance indicators (KPIs). All KPIs defined have to be supported by explicit standards (norms) and controls. Processes are structured for ease of control. Without the structure most processes will be difficult if not impossible to manage. Without effective controls, good business results are unlikely. Good fortune or luck is no replacement for a comprehensive and robust strategy implemented through well designed processes supported by an effective process structure [In't Veld 2002][Bikker 2001].

Requirements: constraints, goals and wishes.

The first step in the process design procedure is to develop a set of requirements and boundary conditions for the processes to be developed. The procedure starts with identifying all stake holders and assessing their wishes. Identifying all stake holders also sets the *system boundary* as discussed in paragraph 2.3.1. Sometimes this system boundary is also referred to as *problem scope*. All stake holders with a significant stake need to be inside. The term wishes is used to indicate that there is still freedom to acknowledge or reject these wishes. For practical use these wishes are sorted in three different categories of statements: constraints, goals and principles (figure 2.8). This is the simplified goal analysis method of In't Veld [In't Veld 2002, p.192].

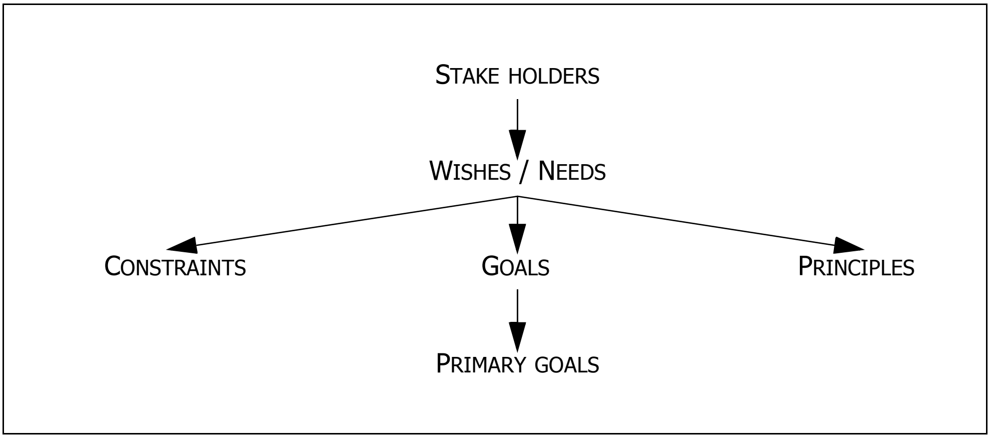


Figure 2.8: Simplified goal analysis method [In't Veld 2002].

Constraints, goals and principles play a different role in the design process. Goals specify the contribution of the process to its customer, thus the goal directs the design. Constraints put boundary conditions on the design. Constraints are very powerful in selecting alternative solutions. Principles often refer to the legacy of the company. Principles are statements about solutions and approaches that have been proved to work in the past. Re-applying principles supports learning and avoids the risks involved in developing new concepts. Yet constraints may prohibit the reuse of old principles.

Pahl and Beitz [Pahl 1996, p45] also recognize the importance of constraints. In defining design requirements, which is part of the product planning and the conceptual design phase, they distinguish three categories of specification information: objectives, constraints and guidelines. Objectives are synonym to goals and guidelines are synonym to principles.

When process requirements are listed, it is extremely useful to record the source of specific wishes or needs. Some of these wishes may be traced directly to customer value attributes where others are merely the result of internal guidelines or design decisions that were taken earlier in the process. Goals and constraints that belong to the category of customer value attributes cannot be relaxed or ignored. In marketing terms the most important constraints and goals belong to the requirement sets of order qualifiers and order winners.

Process performance and key performance indicators.

To design processes, we need to know about the transformations and about the performance criteria for the processes. Also, the constraints need to be specified in terms of performance criteria. In't Veld proposed a set of three process performance indicators and a combined indicator (performance) that is particularly useful in comparing and selecting resources (table 2.4).

Productivity is the ratio between results that are conform specifications and the sacrifices needed to produce these results. *Effectiveness* is the degree to which the results meet the customer's demands (quality, product-mix, delivery time, and delivery reliability) under the condition of chosen delivery scenarios. If effectiveness is a minimum requirement for the customer then ideally effectiveness should be at least 1.

	DEFINITION
Productivity (P)	P = **R**esults / **S**acrifices
Effectiveness (E_R)	$E_R = R_{operation} / R_{norm}$
Efficiency (E_S)	$E_S = S_{norm} / S_{operation}$
Performance (Perf)	$Perf=(R_{operation} / R_{norm})*(S_{norm}/ S_{operation}) = E_R*E_S$

Table 2.4: Performance variables [In't Veld 2002]

Efficiency is the ratio between the norm sacrifices and the real sacrifices from operations. Efficiency could exceed 1 if the sacrifices (S) are less then the norm, while achieving an effectiveness of 1 or better. This may sometimes occur when a production means is subject to statistical yield functions less than 1. This loss is usually compensated for by ordering extra products. These products may not be manufactured, if it becomes clear that the yield realised is better than expected.

Performance is defined as the product of effectiveness and efficiency. Performance is an index that relates the target for productivity to the realised productivity. In fact it can be demonstrated that the performance (Perf) equals $P_{operation} / P_{norm}$.

In this thesis the terms process performance indicator and process performance requirements[1] refer to the same concept. The major difference is the time of use. Performance indicators are measured results of processes in operation. Performance requirements are targets that are used during process design or norms for process operations planning.

Effectiveness, efficiency and productivity can be used to specify process requirements (target), to set a norm for operations (norm) and to measure the performance of operations (operation). For selecting alternative resources during

1. Sometimes the word process is left out.

design P_{norm} becomes P_{target} which is the maximum productivity possible with that resource. Under constraints set by effectiveness, performance can be used to achieve a balanced process optimization; that is to achieve maximum efficiency while maintaining effectiveness at the minimally acceptable level. In theory this should yield the same result as optimization of productivity. However, there is a mathematical down side to productivity. Productivity is the quotient of two scalars that are not dimension less. Thus, unlike other variables, productivity cannot be used in the vector algebra formalisms of signal processing, whereas the dimensionless indicators as effectiveness, efficiency and performance can.

Bikker favours an alternative set of process performance indicators, strongly related to the ones above, but some of them less formally defined [Bikker 1994].

	DEFINITION
Effectiveness	$E = R_{operation} / R_{norm}$
Productivity/leadtime	P = R / S, Order leadtime, Process leadtime.
Flexibility	Product flexibility, Volume flexibility,.......
Control	Effectiveness of control over other performance variables
Quality of labour	Indication for the well being and autonomy of the worker
Innovation capacity	Capacity to learn, adjust and improve processes

Table 2.5: Bikker's process performance variables

Effectiveness or effectivity is the most important performance indicator and its definition is conform the definition introduced by In't Veld. Effectiveness is the most important indicator because it is directly related to the drivers of customer value. Effectiveness includes notions as product quality, number of products expected, packaging and agreed delivery schedules.

Productivity is the performance indicator to be optimized in process design. Productivity is mentioned along with leadtime since leadtime is strongly correlated with productivity.

Leadtime is the time span between the moment the order was received and the

moment the result is delivered. A long leadtime often causes high levels of work in progress and high control loads. The market is not very appreciative of a long leadtime. Speed and responsiveness (short leadtime) are often rewarded with higher prices. Lack of these requires a monopoly and/or exclusivity to maintain a high price.
Flexibility refers to agility and capability to deal with variability in customer demands (in volume as well as in product specifications).
Control is an aspect that applies to all of the criteria above. It indicates if goals have been set and if and how these goals are met. However this performance variable is not needed as a process specification for two reasons:

1. *Control is the main axiom of our method; without control there is no performance achievable in any of the variables.*
2. *The performance of control is observable in the variance of effectiveness, efficiency, productivity and leadtime.*

When compared to In't Veld, Bikker dropped efficiency, because efficiency was already accounted for in the productivity indicator and because leadtime was made explicit as a strong influence on efficiency. A more political reason for dropping efficiency was that efficiency was the only indicator understood properly by financial managers. Often efficiency was abused to obscure the visibility of the other performance criteria, including effectiveness. However by dropping efficiency, the financial criteria were made more implicit and for engineers easier to ignore, which is not a good thing either.

The most important additions that Bikker made to the set defined by In't Veld were: *Quality of labour* and *Innovation capacity*. These additions brought the performance indicator set closer to the full scope of the business balanced score card (BBSC) as will be discussed in paragraph 2.6.
Quality of labour or quality of the working environment is an indicator of the well being of the employee. This includes health and safety precautions but more in particular the internal and external autonomy are made visible and valued. The idea behind this indicator is that effectiveness, productivity/leadtime and efficiency may show satisfactory, but they are in danger to degrade if the quality of labour remains low. Stressed and unmotivated workers do not produce quality and they lower

efficiency because of a higher level of sickness leave.
Innovation capacity indicates how well an organization is capable of adjusting its policies, behaviour and structure in a changing environment. Although this aspect may be difficult to quantify, it is clear that, if an organization does not keep track of strategies and their results and does no evaluation of its actions, no basis will exist for improving and innovating processes. So basically innovation capacity refers at least to the presence of secondary controls to adjust settings, but it also refers to learning as a base for redesigning processes and restructuring the organization of processes. In some (limited) sense, a low score on innovation capacity may be an early warning for problems to sustain positive scores on the primary process performance indicators.

A minimum set of process performance requirements

As long as the underlying physical variables (Results and Sacrifices) are known in real numbers, it is sufficient to specify *effectiveness, efficiency* as primary specification requirements of the processes. The attributes of effectiveness, efficiency are rooted in the customer value attributes and in the financial targets of the company related to value creation. Time, money and other capacities or resources allocated to support the execution of the process are all accounted for in the efficiency requirement. Effectiveness and efficiency indicators also support comparing different processes and process/resource combinations.
Effectiveness also includes product- and volume-flexibility requirements which when implemented may result in process flexibility capabilities. Despite the attention for flexibility both from customers and managers, the need for flexibility is the result of strategic choices of the company. Thus it follows from the set of effectiveness requirements.

To sustain those specifications we also need secondary controls and a memory to learn from. This is specified by the *capacity for learning and innovation*. In addition the *quality of labour* requirement is needed to define constraints for the organization structure. Constraints that are rooted in the quality of labour requirements are an important factor in task level decisions.

Thus the minimum set of process performance indicators uses only four requirement variables:

1. *Effectiveness*
2. *Efficiency*
3. *Capacity for learning and innovation*
4. *Quality of labour*

Note that if we can quantify the capacity for learning and innovation as well as the quality of labour as a kind of growth or decline index for effectiveness or efficiency, then the use of vector calculus for combined criteria is also supported.

2.3.4 Processes, transformations and functions

To achieve the primary goal, specified by the effectiveness indicator, it is (re)specified as a set of functions. Possibly making use of principles, a complex goal may be decomposed in a set of functions that are capable of fulfilling the goal. To built a system that can achieve the goal, the terms: function, transformation, process and system are defined in table 2.6 and pictured in figure 2.9.

FUNCTION	**The function of an element is the contribution of that element to the environment of the element.**
TRANSFORMATION	**A transformation is a recipe how this element converts specific inputs into a desirable contribution, making use of resources.**
PROCESS	**A process is a sequence of transformations that cause the input(s) of the process to be converted into a desirable output**
SYSTEM	**A system is a set of elements, distinguishable within a complete reality, selected by a researcher in compliance with a goal set by the same researcher. These elements have relations with each other and they may have relations with other elements in the complete reality [In't Veld 2002]**

Table 2.6: Function, Transformation, Process and System definitions

The term desirable in the transformation as well as the process definition intends to indicate that there is a purpose for this process. In other words, the process is intended to fulfil a number of goals. A system is a substantiation of a process. Referring to a goal in the definition of the system gives the designer or analyst the freedom to build models that are incomplete but sufficient for either the analyses or a (partial) design of processes.

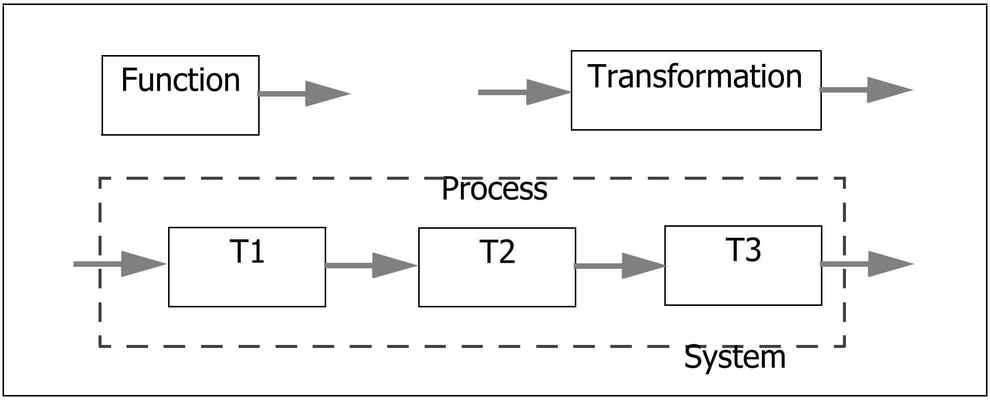


Figure 2.9: Function, Transformation, Process and System.

Designing systems as a possible substantiation of a process is a mapping that starts with the goal (function) and goes backwards towards a possible set of inputs. Backward mapping gives design freedom and it makes sure that in the end the process goals are achieved. This property of process oriented design is also mentioned by Pall as the key to the design of business processes [Pall 2000].

It is now possible to built a black-box system model of a company that uses resources, such as human labour, energy, materials and money to transform these into products and services that fulfil the need of customers. In these models different drawing conventions are applied for different types of "flows". Information refers to specifications and technical process information. Material refers to the main resources that are being used in the transformation process. What is not indicated in this model yet, are the flows of energy and the flow of logistic control. Logistic control manages the timing of committing resources and capacities for executing the transformations. Note that material is used metaphorically for every stream that is directly related to the resources that are transformed into the customer's need.

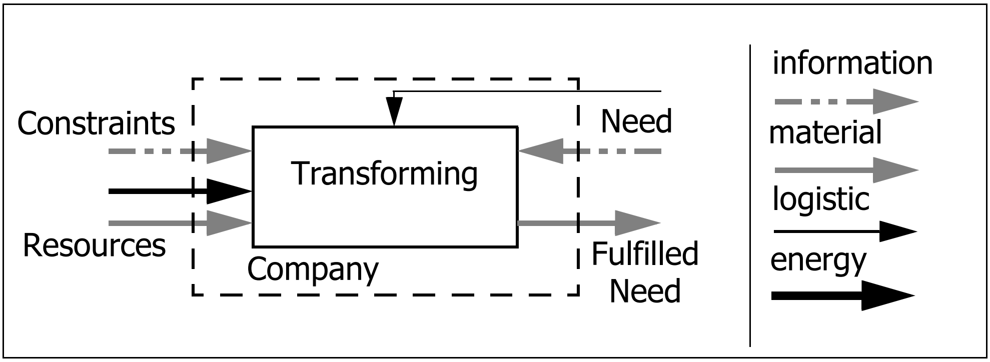


Figure 2.10: Black-box system model of a company

2.3.5 Process control

Process requirements are satisfied through the implementation of appropriate controls. In practice there exist four basic control modalities, each with their own strengths and weaknesses (table 2.7). These control modalities are identical for control of technical processes as well as for control of business processes. In practice combinations of control modalities are applied. For process design it is important that answers to the following control aspects are made explicit:

- *which control principles are implemented to achieve the correct results*
- *how compliance with constraints is realized*
- *how the influences of disturbances are minimized.*

If these questions cannot be answered for all of the process requirements and constraints then it is unlikely that the process will operate satisfactory in practice. For business as well as technical processes, luck is no replacement for robust design and control.

	STEERING	FEED FORWARD	FEEDBACK	OUTPUT ERROR
Main principle	*Inverted model controller. Prevent deviations by detailed design of forward process*	*Compensate for known disturbances causing deviations*	*Adjust deviations by re-applying the forward process*	*Adjust deviations by repair at the output*
Strength	*Predictable success and failure behaviour*	*Fast and pro active response to known disturbances.*	*Robust, even against unknown, disturbances*	*Near perfect output if not self under disturbance.*
Weakness	*High sensitivity to model failures; models often not invertible*	*Application limited to known disturbances to the forward process.*	*Slows down the system. Problems with stability, especially if disturbance rates are close to feedback rates*	*Costly and slow in case of many deviations*

Table 2.7: Control modalities and their characteristics

2.3.6 Organelle Structures

For ease of control, business processes are split into subsystems that each have a clear contribution to its environment. Each subsystem is called an organ or organelle. Together with their relations we can built the organ- or organelle[1] structure which is

1. In this thesis organ and organelle are synonyms.

the basis for building a department and a personnel structure. A (business) organ is defined as follows:

Definition 2.7 Organ(elle) [In't Veld 1981].

> ***An organ is an intentional grouping of transformations, relations and control functions that makes a distinct contribution to its environment.***

Bikker as well as In't Veld recognise three fundamental ways of grouping transformations, relations and their controls: functionally, product-flow oriented and geographical. Grouping functionally means defining organs around the sharing of common tools, knowledge and skills. Product flow oriented grouping means grouping all transformations, relations and controls needed to realise a product or a series of products. Grouping geographically means gathering transformations, relations and controls on the basis of location specific customer interactions. Note that at each level, or aggregate in the organization, one may choose a different grouping principle depending on the situation. Often, product-flow oriented grouping is the preferred principle leading to a high efficiency. In his book "Organisatie stuctuur en arbeidsplaats" In't Veld presents many variants of the above principles to organise or structure operations [In't Veld 1981].

This concludes the still unique doctrine about organization design:
By starting with the design of business processes and structuring these with the intent to facilitate quality management (customer value) and to optimize productivity, a structure can be found that serves as a basis for the design of a department and management structure. Thus an implicit and natural match is accomplished between the quality control and management requirements from the market and the management and control capabilities of managers and directors.

The match between process management requirements and management capabilities prevents many of the induced uncertainties and unnecessary coordination leadtime that could frustrate workers and managers if this match is not accomplished.

2.3.7 Reflection on the Delft School of Organization Design

The major point made by the Delft School of Organization Design is to develop organization as a means to efficiently manage the processes of operations that are responsible for the delivery of customer value. In other words, process design precedes organization design. This principle still holds today and because of the focus on effectiveness and efficiency for the process design, it may even be claimed that the theory of lean manufacturing as advocated by Womack and Jones was practised by In't Veld and Bikker long before the books of Womack and Jones were being published [Womack e.a. 1990, 1996].

Yet, especially at strategic level, the tooling for process design and organization structure design was not strongly developed. Effectiveness is a very general and abstract notion that is not easily specified unless the product and its customer value attributes are clearly defined.
Financial theories and instruments beyond cost-price calculations were often kept outside the scope of most thesis projects. Today, financial instruments are largely responsible for resource commitment decisions and constraints under which processes are operated. In paragraph 2.4 some tools will be introduced to bridge the gap between value creation and specifying process requirements.

Despite the acknowledgement that the design of organization structures on top of operations is a vital step in building companies that are both effective, efficient and agile, the organization design theories of In't Veld and Bikker are mainly descriptive and explaining theories. This means that through case based learning and associative thinking, students and other practitioners could learn how to design organizations. But the lack of a more explicit reasoning process meant that mistakes are easily made and it remained difficult to identify a starting point in the process. In paragraph 2.5 the drivers behind the basic choices are made more explicit and a set of design rules for making structure decisions is proposed. These rules now also include strategy level decisions.

Although the initial version of Bikker's business process redesign framework contains more detail on what steps to take and what structures to analyse (figure 2.6), the aggregated version (figure 2.7) is preferred by many because it shows better where to start and how to proceed. Yet the analysis of the As-is situation proved difficult in matching (partial) process and structure models to the process requirements that resulted from the policy analysis. An improved version is presented in paragraph 2.5.3.

2.4 Sustainable value creation

Companies exist to create sustainable value. Sustainable has a dual meaning here: continuity and accountability. Continuity and accountability cannot be separated. Continuity requires profits needed to sustain investments into new products and improved manufacturing processes. Profits are only granted by stake holders if a company behaves accountably for all of its actions. Without the continuity however, accountability is like an empty shell. When bankruptcy may have occurred, society pays for all damages.

Continuity requires a balance between profits and investments. Creating a customer value proposition is not such a big problem if one knows the targeted customer group well enough. Implementing this proposition, such that the resulting market position can be defended is the real problem. Usually this means that the proposition should be based on a set of assets that competitors cannot easily copy, develop or get access to. Note that assets refer to any physical or non-physical resource available to a company for developing and implementing customer value propositions. This means that detailed knowledge about a customer base is also an asset that could help to defend a market position, if this knowledge is hard to get.

Profits require the creation of (added) value. But profit alone is not enough to create a sustainable business. Modern business valuation theory regards companies as a portfolio of investments. These investments support activities that generate present and future cash flows. The free part of this cash flow (free cash flow) can be used to support new investments for sustaining the company. Without free cash flow a company will loose its power to invest and renew its economic activities.

Free cash flow is defined as follows [Grant 2002]:

NET OPERATING PROFIT		FREE CASH FLOW	
	Total operating revenues		Net operating profit
-	Cost of goods sold	-	Taxes
-	Selling, general, administrative expenses	-	Investments in fixed or working capital
-	Depreciation	+	Depreciation
Total	Net operating profit	Total	Free Cash Flow

Table 2.8: Net operating profit and Free Cash Flow

From table 2.8 it can be seen that depreciation is a non-cash expense which does not influence the free cash flow. Eventually depreciation should lead free cash to be invested in new assets (fixed capital). A positive net operating profit does not always go together with a positive free cash flow. Investments in working capital may be necessary because of growth. A negative free cash flow may eventually lead to bankruptcy if the remaining working capital is insufficient to support the level of business.

Enterprise value = Cumulative Present Value (free cash flows) + Residual Value or

more formally: $Enterprice\ Value = \sum_{t=1}^{n} \frac{E[FCF]_t}{(1+r)^t} + Residual\ Value$.

$E[FCF]_t$ is the expected free cash flow in period t, r is the risk-discount interest rate [Rijn 1999].

The residual value has two components. The period over which the enterprise value is determined is limited by n periods. After n, there is a residual value related to cash flows that still exist then, which have to be discounted to its present value. The second component in the residual value is the current value of investment holdings.

Examples of such holdings are marketable securities, investment in stocks and bonds, investments in unconsolidated subsidiaries etc.

The risk discount interest rate is usually the so-called WACC (weighted average cost of capital). If a company earns exactly its WACC for every period, then the discounted value of its projected free cash flows should equal its invested capital. If a company earns less then its WACC, then it is losing or destroying value.

Customer value is the driver of cash flow. Successfully executed customer value propositions are the basis of real profits. So it is safe to say: strategy is about building sustainable customer value propositions. In this perspective, understanding free cash flows first means understanding the driver of cash flow which is customer value; customer value for the end user as well as for all supplier-customer and business to business relations that stand in between.

2.4.1 Value creation, strategy and quality

The value of a product or service is defined as the price a customer is willing to pay, at the moment of purchase. There are several variables involved in valuation done by the customer. Kemperman has put the drivers of customer value in perspective and concluded that these drivers as well as the assessments made evolve over time.

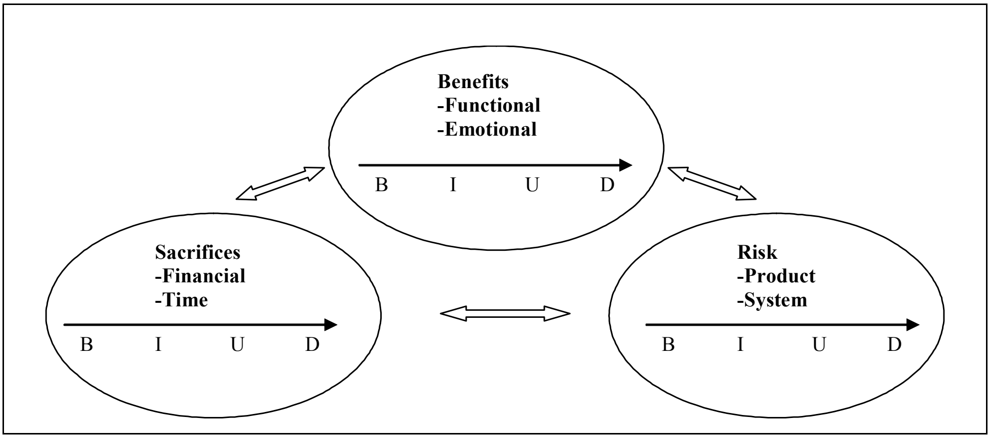


Figure 2.11: Customer value drivers.

The time aspect is indicated with the arrow and the letters refer to the lifecycle-phases of the product that are relevant to the customer: Buying, Implementing, Using and Disposing (figure 2.11) [Kemperman 1999].

Functional benefits address directly the customer needs within the customers processes of use. Emotional benefits refer to status or image a customer experiences owning such a product. This is the difference between a Rolex and a Swatch. Both brands make watches that do tell you time with enough precision for every day use. Yet, within different social environments, one or the other is preferred. Brand value, associated with exclusivity, is often reflected in the price.

The expectation of product (failure) risk is also correlated with brand value. Products of well-known brands are considered to be of higher quality than those of unknown brands. This is reflected in the price. If consumers are first-time buyers, the are likely to buy brands they know from commercials. However, if no-name brand products are priced very cheap and considerably lower than comparable products of known brands, customers may take the gamble and buy the no-name product instead. In recent years we have seen this development in the market of power tools for home do-it-yourself use. Typically no-name brands, often manufactured in China, were priced at levels only 25% or less of that of the known quality brands such as Bosch, AEG, Makita and Metabo. Interestingly, the no-name products are being sold with a warranty period even longer than customary for the known quality brands. In practice this warranty is a swap warranty. If the tool fails within the warranty period and the failure is not due to ill use or excessive wear, the tool is swapped for a new one.

Some products depend on other products or services for their benefits. The availability of these other products, today as well as over the expected life of the product is an important factor. The expectation of these conditions not being fulfilled is the customer perception of system risk. Standards and the availability of multiple suppliers are very important to reduce this risk. Changes in system risk have a big influence on the going price of a product. Digital Compact Cassette (DCC) players lost 70% of their sales value in November 1996 after Philips announced to stop the manufacturing of players and gradually phase out support for the DCC standard. A similar price development can be observed in the process of migrating from VHS to DVD for home video entertainment.

The sacrifices a customer faces are both financial as well as time. Products that are complicated in use, or that require a significant amount of time to install or to assemble before the benefits can be enjoyed (kits) are often cheaper than easy and ready for use products. Here the context of the customer is very important. Skills and abilities of the customer influence price and the fun of building a kit could be the most important benefit.

In addition to these drivers, the laws of economics influence the setting and acceptance of prices and thus customer-value. The main factors here are availability and competition. These factors can be discounted either as a time or as a financial sacrifice for the customer. The customer balances the urgency of his need with his willingness and ability to pay a higher price; 24 hour repair services often charge higher fees on Sundays and National holidays, cold drinks are always more expensive on a hot summer beach than in the super-market.

Since the application context largely drives what is an acceptable price, suppliers often set prices to attract a sufficiently large group of customers while maintaining a good margin. They try to maximize the sales value. Product differentiation could lead to an even higher market share and better economies of scale, thus increasing the solvability even more. But there is a down side to differentiation. If the distribution channels are not well separated, product differentiation can lead to market cannibalization and underselling. High value customers may seriously consider the cheaper alternative if the additional procurement cost associated with buying through the cheaper channel are not too high (these additional cost may even be negative).

More formally, suppliers have to chose their customer value proposition(s) balancing the following influences:

$$Salesvalue_{max} = Max[n_1 \bullet Value_{appl1}, n_2 \bullet Value_{appl2}, \ldots, n_N \bullet Value_{applN}]$$

$$Salesvalue_{min} =$$
$$Min[\beta_1 \sum_N n_i \bullet Price_{chann1}, \beta_2 \sum_N n_i \bullet Price_{chann2}, \ldots, \beta_N \sum_N n_i \bullet Price_{channN}]$$

where $\sum_N n_i$ *is the total number of customers and* β_i *is the fraction that will buy the product through channel i.*

From these principles of customer value a definition of quality can be derived.

Definition 2.8 Quality

> ***Good quality is a product condition where the customer value expectation at the moment of buying (t=0) is less than or equal to the customer value experience when the product is in use (t>0).***

The quality condition is also a customer retention condition.

BUYING REQUIRES:	QUALITY PERCEPTION REQUIRES:
$Salesprice_{t=0} \leq Expectation[Value]_{t=0}$	$Expectation[Value]_{t=0} \leq Experience[Value]_{t>0}$

Table 2.9: Buying and Quality perception conditions

Although this definition of quality was developed from the customer perspective, it has company wide implications. The expectations on customer value are the responsibility of marketing. It is their task to present an appealing image of the product and to make the customer think that their product is a perfect match to his needs. Thus understanding customer value drivers is not only essential for business process design, it is also a key element in the design of marketing strategies and in understanding the position of competitors. Yet the customer value experienced is created by product design, manufacturing engineering and by manufacturing, distribution and service processes.

The customer value mix is also valid in a business to business (B2B) context of Supply Chain Management (SCM). In the SCM context, procurement and sales negotiate technical quality, volume, order leadtime as well as service reliability in order to optimize their respective potential for value creation. It is important to realize that the notion of product is now extended to a full specification of the customer value mix, which includes logistic, service and packaging requirements.

The tools in this section also come close to two of the most important aspects that entrepreneurs take into account for their decisions: customer value and free cash flow. Understanding customer value offers insight for product development, advertising and marketing; cash flow or even better free cash is the fuel tank that allows the company to operate.

2.5 Design of organization structures

A process design with transformations and controls is the starting point for the design of the organ- or organelle-structure. Having discussed the importance of control in process design, the organ definition (definition 2.7) of In't Veld is sharpened a little:

Definition 2.9 Organ(elle)

> ***An organ is a cluster of functions and controls that together operate as a unit, reliably delivering a distinct contribution to the environment in which the organ is defined.***

The main difference with In't Veld's definition is the addition of the attribute reliably. This means that the organ's function is not only distinct, but the organ also contains all controls necessary to operate reliably in the context where the organ is defined.

This definition of an organ is in line with the effects of grouping discussed by Mintzberg (table 2.3). The organs and their interrelations make the organ-structure.

The organ-structure is the basis for assigning divisions, business units, factories and departments and it is the starting point for designing the personnel structure. This clustering step is most critical in the design process. A general solution that is optimal under any condition does not exist. Changes in the environment may call for changes to the processes as well as to the clustering. However changing the organ structure and the department structure associated with it, is not an operation that can be implemented overnight. Therefore the robustness of the design needs to be tested against different scenarios that could develop in the environment of the company and contingency criteria play an important role. Rather than imposing a structure or tailoring a structure, structure design is a procedure that can be followed to identify grouping decisions and their criteria in an orderly fashion.

To understand how this works, first a short example. In figure 2.12 a simple process is shown, together with some controls and two levels of organs defined. In this figure

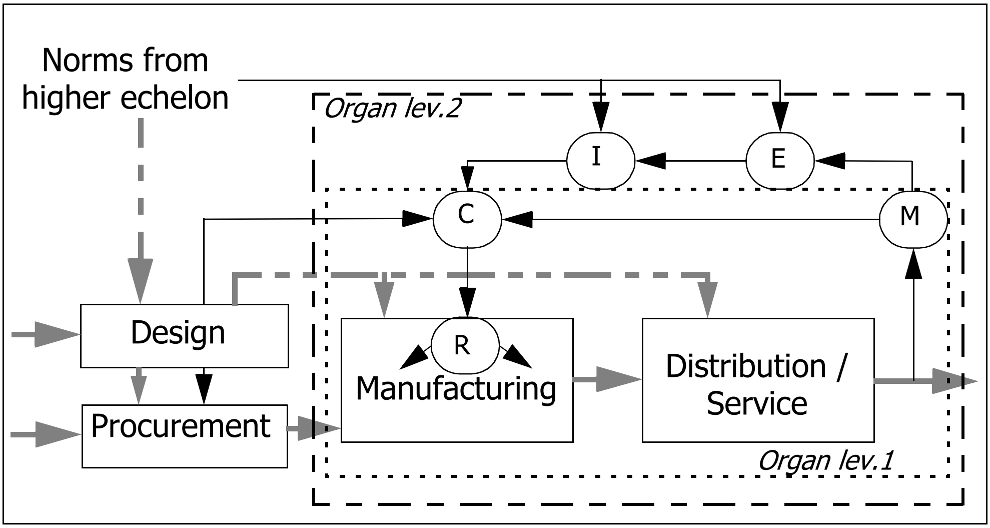


Figure 2.12: Transformations, control and organs.

M stands for measure, C stands for compare, R stands for regulate or intervene, E means evaluate and I means initiate. Two possible organ boundaries have been defined: level 1 and level 2. Placing the organ boundary at level 1 means only the manufacturing and distribution operations and a feedback control loop are bundled. To operate reliably, the environment has to be quasi static. This means the internal

feedback loop should be fast enough to cope with any disturbances that may come from the context of this organ. The organ operates in a fixed fashion, there is no adaptation in the settings of either operations or controls.
Placing the organ boundary at level 2 means the evaluation (second order feedback) is now included. This means the organ can adapt its operations to changing circumstances within limits set for this adaptation by higher echelons. The level 2 organ has autonomy over the norms for its internal controls as long as they comply with the higher echelon norms. For this reason it can be expected that the quality of labour may be better, as well as its innovation capacity.
Which organ, level 1 or level 2, is better depends on the dynamics of the context and on policies set by the company.

2.5.1 Rules for the design of organization structures

Mintzberg's theory on effects of grouping indicates that an organ supports unified control (common supervision, mutual adjustment, common measure of performance) and relies on sharing resources for executing a distinct function in its environment. If we are to limit complexity with this grouping, then it should also be clear that a unified control strategy for all of these elements is beneficial. This means the aggregated function of this organ should not be too diverse. Reversing Mintzberg's argument means we look for a set of functions that benefit from unified internal control, while at the same time supporting ease of control of the relations between elements. Function definitions that are too diverse should be split into separate more homogeneous functions that can be unified under a more abstract control principle.

Placing organ boundaries in a process design means we define an abstracted or aggregated view of this process design where these organs and their functions become modelled as black boxes and relations between these black box models. In terms of elements and relations to be considered simultaneously, this is a reduction. The reduction of elements and relations is also a reduction in complexity if the uncertainty associated with the states of the remaining elements and relations is not increased at the same time. In other words, placing an organ boundary should not result in hiding information that is critical for the functioning of elements outside the organ. If this problem occurs, then these elements should be pulled to within the organ boundaries and the organ's function should be reconsidered.

These principles have a lot in common with the principle of increasing precision with decreasing intelligence (IPDI) for intelligent machines as defined by Saridis. Saridis characterises intelligence as dealing with uncertainty. Saridis argues that for building intelligent systems we need a hierarchical control system. In this system the intelligent decisions are taken with low precision at a high (supervisory) level in the control hierarchy whereas detailed operations level decisions are taken with a high precision at a low level in the hierarchy. At a high level, there are less states to choose from (less precision) but more uncertainty. At a low level in the hierarchy, there are more states (more precision) but with less uncertainty [Saridis 1988,1989].

Thus, if setting organ boundaries is to support a reduction in complexity then we can formulate the following requirements for organs and their relations:

1. *Organ boundaries are set to limit complexity at each aggregate.*
2. *Organ boundaries allow sequential and individual analysis and improvement of these organs while maintaining the quality of the system as a whole.*
3. *The signals or information attached to relations between organs, should be quasi-static when compared to the internal dynamics of the organs.*

The 1st and 2nd requirement are function decomposition requirements. The 3rd requirement puts a constraint on the dynamics between organs. The first and second requirement should limit the complexity to be dealt with simultaneously to an acceptable level. However it should be realized that because of better[1] internal communications and controls, organs can deal with more complexity internally, than can be passed onto other organs through external relations.

Although we could test organ structures at each aggregate for their compliance with these requirements, this is not a design procedure yet. The requirements specify

1. Higher capacity

what should be the aggregated result of the process of placing organ boundaries. These requirements give no explicit clue for what aspect to consider to identify functions and what unified control principle could support placing these boundaries. For this we have to combine these requirements with the process performance requirements as discussed in paragraph 2.3.3.

Effectiveness is rooted in customer value and in the companies ambition to be competitive. At company level, effectiveness is represented by the value proposition(s). Effectiveness is also the basis for identifying functions that can be decomposed to subfunctions and organs that have to comply with the 1st, 2nd and 3rd requirement for placing organ boundaries. Too much diversity and intertwinement of value propositions can cause uncertainty and is not a good starting point for a decomposition that supports limiting complexity. Yet, sometimes this intertwinement is unavoidable and quality control considerations require that functions and their controls should not be spread over different organs. This has lead to the 1st and 2nd rule for placing organ boundaries:

1. *Do not combine value propositions that are too far apart into one process.*
2. *Identify organs that can be responsible for a distinct contribution, but do not cut important control cycles.*

The first rule limits the diversity at the highest level, creating a basis for unified control. The second rule maintains unified control and limits the complexity at each aggregate. In addition capacity for learning or innovation may be used to specify additional (control) functionality to implement this requirement.

Yet the organ structure defined this way may not be efficient. Each organ is entitled to its own private resources, which may result in very inefficient use of human labour and equipment. If the cost of coordination of sharing these resources does not exceed the cost of having multiple copies then it is efficient to look for these opportunities to share certain resources. This resulted in the third rule:

3. *While maintaining 1 and 2, try to create economies of scale (efficiency).*

The third rule leads to the introduction of various control and scheduling functions to manage shared use of (expensive) resources. Note however that this rule should be applied and evaluated only in the context of the value propositions that belong to the organs that share the resource. Ignoring this context leads to controls and scheduling rules that may starve some of the orders or unknowingly introduce a high level of functional diversity.

The first and second organ structure requirement together with *the effectiveness*, *efficiency* and *capacity for learning and innovation* requirements, partly support the *quality of labour* requirement. The balance between authorization and responsibility should have been looked at to implement these requirements. Yet, there may be organs in the systems that are too small to function socially. Usually such small departments also have trouble meeting the health and safety requirements that are also captured under the umbrella of *quality of labour*. The fourth rule proposes to merge these small functions and departments to create a small number of viable multi-service organs. In order to support unified control for this department, the logistics (service times, frequencies and priorities) of different functions should not be to far apart.

4. *While maintaining 1 and 2, try to achieve some flexibility and reduce the vulnerability associated with small departments and product oriented structures by merging these often service functions into larger multi-service units.*

Finally the initial strategy and value proposition may contain propositions that are clearly outside the scope of the resources the company can commit. Maybe the company lacks the resources to be efficient in serving customers that are outside the scope of the main profit makers. Yet these propositions cannot be simply neglected. In product-bundles these odd products or services may be highly appreciated by some customers. They grant their contracts for high margin products or services generously on the condition that you also help them out with some odd ones. This does not imply that you have to implement these odd propositions fully within the company. The customer will be happy no matter how you deliver, as long as you do and as long as you accept responsibility for these odd services. The solution is to assess the real value of the customers that consume these services and find a trustworthy supplier that is better capable of delivering the value required. One may find that some customers are not profitable at all. They consume capacity at zero or even negative margin. Such customers should not be accepted at all. This is the basis for the fifth rule:

5. *If after 1-4 value propositions are left that do not add value to the company, outsource them or give up these lines of business.*

or in other, more popular words: *To call every customer King does not imply to accept every King as customer.*

2.5.2 Organ structure design rule implications

One could rephrase the 1st rule as *do not serve different markets with one process*. Allocating one process to orders of different markets causes problems in managing order priority. These problems can only be solved with excess capacity and excess inventory. Alternatively, one of the markets may experience starvation or at least excessive leadtime. In logistics, different value propositions should be expressed as different product market combinations (PMC's). This implies that a PMC is a product representation, often implemented as a bill of materials (BOM), linked to a market that is defined as a set of customers and competitors that engage in transactions under homogeneous logistic (and value) requirements.
Thus a product may be sold to different markets, serviced through different manufacturing processes and channels. Functionally identical products may also have more than one BOM, each tailored to the channel for a particular market.

Placing functions and controls that need intense communication into different organs adds unnecessary leadtime and communication overhead to the organization, which should be avoided (rule 2). Adding additional leadtime can cause instability if it affects feedback or feed forward circuits. In "manned" organizations there is an additional problem which could be compared to degrading signal to noise ratios in long wires. Person to person communication relies on a common language and a common interpretation of information communicated. If persons belong to different departments, common interpretation of unformalized language can no longer be guaranteed. The longer the "organizational distance", the larger the difference in interpretation may become. This is especially a problem if management incentives create competition between departments rather than cooperation.

2.5.3 Modifications to Bikker's BPR scheme

The business process redesign scheme of figure 2.7 works like a design process. At first, based on the problem definition, the policy development is studied. The as-is policy and the to-be policy, supply the reference framework for valuing the processes and structure of the company. The analysis of policies makes use of the tools presented in paragraph 2.3.3 and in paragraph 2.4.1. Making a valuation of these processes and the company structure is like taking a measurement. The only difference with (most) engineering measurements is that there is no absolute scale of reference. The policy framework sets a reference relative to the competition. For a well-run company it can be expected that the as-is structure is capable of meeting the goals set by the as-is policy. These goals in turn have to meet external requirements set by the social-economical environment of the company. Problems (tension) arise usually from changes in the external requirements that call for a change of policy. The demands set by the to-be policy usually cannot be met by the existing structure. Note that these external changes are sometimes the result of internal business decisions. The demands, together with contingency constraints from continuity of service with the existing organization are the program of requirements for the new design.

However in practice the assumed match between the as-is policies and processes and structures may not exist. The policies may not be explicitly known. Not all parts of the processes, controls and their structure may be observable. This makes the analysis through the use of Bikker's BPR scheme almost impossible. Moreover policies and processes are not specified in the same language. To facilitate the analysis of the match between policies and processes, an additional step defining the so-called minimal reference model was introduced.

The minimal reference model is the "green field" model experienced researchers and consultants often make up in their mind when they know the policies and predict what to expect on the "factory" floor. The minimal reference model enables a direct one to one comparison of functions, controls and structure. If this model is compared to the observable part of the existing organization, gaps may be observed between the minimal reference model and the observed processes and practices on the factory floor. This does not automatically imply that the company is not capable of achieving its targets. The informal (non-observable) organization may be responsible for implementing some of the controls and coordination functions that were identified as missing. This is valuable information, since the informal organization relies on individuals to function. Absence of these individuals can seriously impact the companies performance if their unseen work was vital. Another reason for observed differences may be incomplete policy specifications. Through communicating the minimal reference model, experienced workers and managers of the company notice that there are vital processes and controls missing in that model. Often this means there is policy information or knowledge missing with the researcher that constructed the minimal reference model.

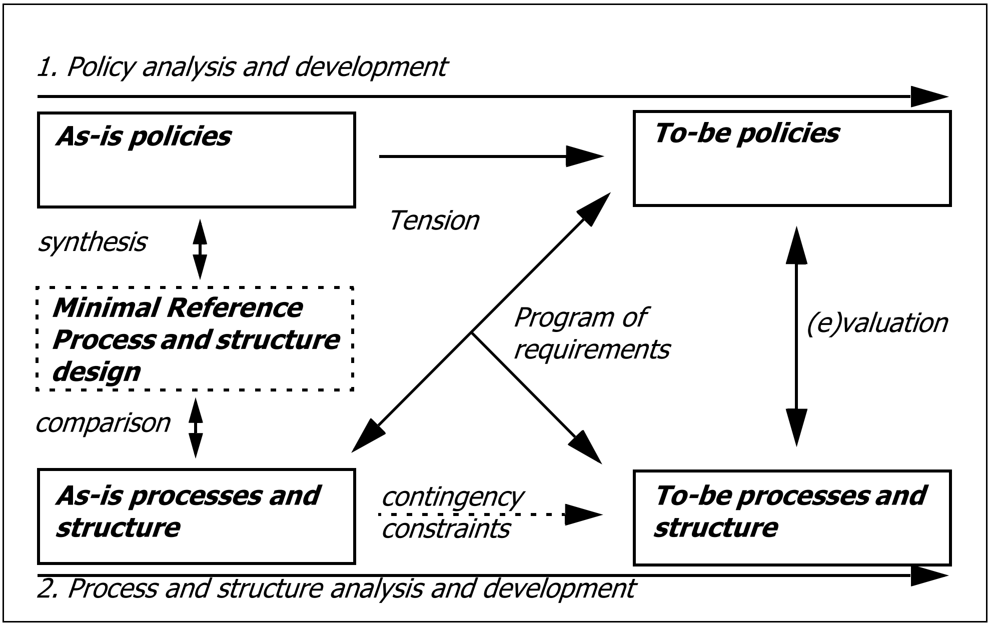


Figure 2.13: Modified Bikker's BPR model

2.6 Logical validation

In table 2.3 Mintzberg's theories for grouping were listed. Through combining requirements for organs and organ-relations with process performance requirements a set of five sequential design rules were defined. Although Mintzberg theory for unification of control was taken as a starting point for the reasoning it is interesting to test if other theories of grouping are also covered by the five design rules. Regarding the effects of grouping all effects listed refer to the creation of a unified control system for the organ. Regarding the basis for grouping, various options are listed: knowledge or skill, work process or function, time, output, client or place. These options refer to the aspect that is considered most important as a base for unified control. In the structure design rules this is covered by the effectiveness or customer value proposition that drives the first and the second rule. In other words if either one of these aspects is considered an important ingredient of control, it is also part of the effectiveness criteria, which in the process design and structure rules primarily drives the control design.
Mintzberg's criteria for grouping also refer to mainly control aspects except maybe the social inter dependencies. These are covered by the fourth structure design rule.

Yet, looking for opportunities to unify control and applying his theories for grouping, Mintzberg seems mainly interested in organizing scale; making organs as big as they can possibly be. This is where the first as well as the fifth structure design rule can make a difference. Both these rules state explicitly the conditions under which unified control is not desirable.

Another interesting frame of reference to use for process and structure design is the so-called business balanced score card (BBSC)[Kaplan 1992]. In the BBSC, a successful vision/strategy is the result of balancing four perspectives: customers, financial, internal processes and the learning/growth perspective. As can be seen, effectiveness, efficiency, quality of labour and innovation capacity also cover all perspectives in the BBSC and productivity/performance are useful optimization drivers in balancing these perspectives. Thus Bikker's process requirements as well as the minimal set of process performance requirements is complete when compared to the perspectives of the BBSC.

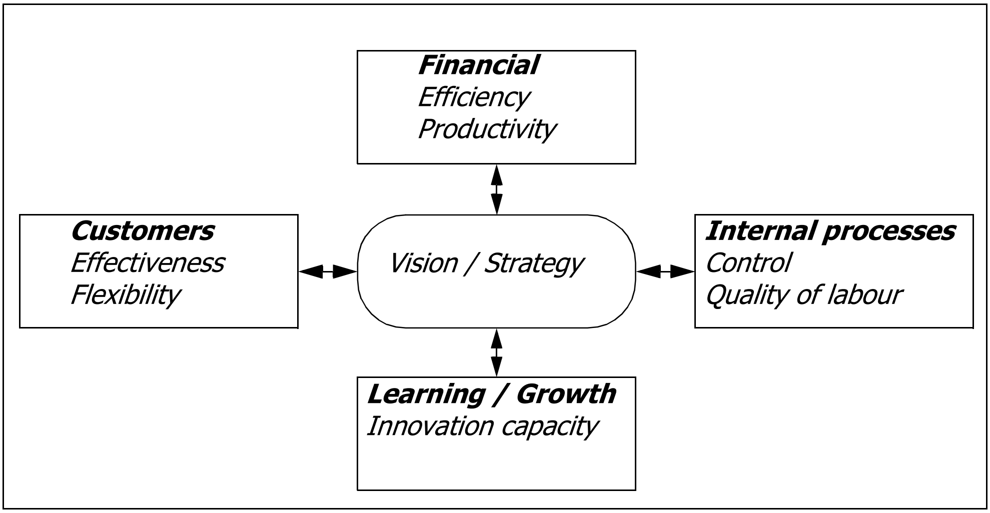


Figure 2.14: Process Performance Variables in BBSC perspective

2.7 Organ structures in practice

To illustrate how organization structures can be used to solve effectiveness problems, while maintaining efficiency, three cases are being presented in this paragraph. These cases are also considered typical for situations that occur in practice, regardless of the branch of industry.

2.7.1 Manufacturing engineering

A large manufacturer in the tobacco industries had the following problem. To support its high speed mass production facility, the manufacturing engineering department was responsible for: production monitoring, production planning, maintenance planning, improvement projects, information support, technical support and troubleshooting. Despite the existence of a dedicated sub department for improvement projects, manufacturing engineering failed to develop and implement these projects in time[1]. It was also felt that technical support and troubleshooting

1. In fact this observation triggered the installation of the dedicated improvement team, two years earlier; but apparently with (too) little success.

seemed to eat a growing part of the capacity. The initial question was to develop a set of key performance indicators that could be used to measure the effectiveness and efficiency of the manufacturing engineering department and that could serve as a basis for bench marking other support departments as well. A brief initial process analysis to establish how manufacturing engineering dealt with the diversity of questions of figure 2.15, revealed that there was no priority filtering. Capacity

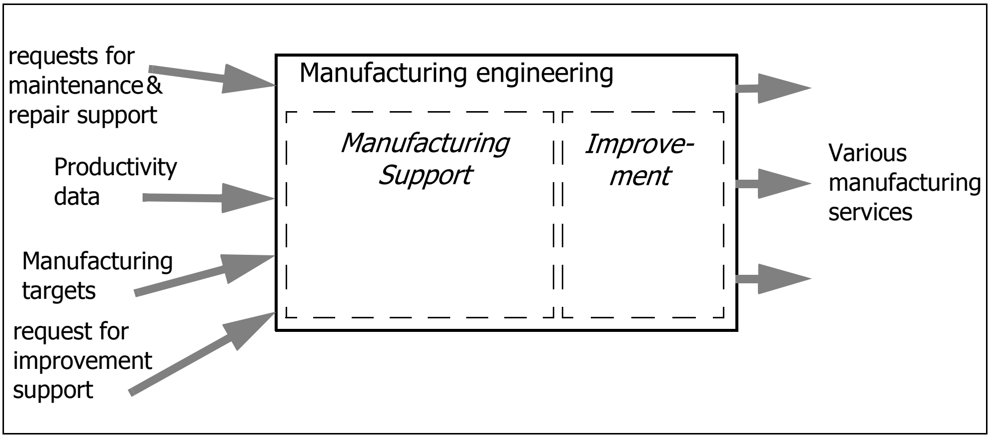


Figure 2.15: Manufacturing Engineering

allocation was merely done on the basis of technical- and historical knowledge. Members of the improvement team were also loaded with routine as well as troubleshooting tasks, in cases where this was considered more efficient then asking somebody else who lacked some past experience with the particular issue. Also there was no capacity or budget limit that would work as an incentive for manufacturing operations to rationalize their need for manufacturing support. Analysis of the input revealed that the stream of support requests could be categorized into three classes: adhoc or immediate requests, routine information requests and improvement project requests. The manufacturing engineering department was restructured to facilitate these three streams with specialized dedicated processes and a capacity overflow system that in specified exceptional situations could allow routine information requests to be postponed temporarily in a case of too many adhoc requests. Under even more strict (exceptional) conditions, the improvement team could also be bothered with adhoc requests.

In addition KPI's were developed for monitoring these separate streams. Service budgets were initiated to improve the involvement of customers in prioritizing the use of the limited manufacturing engineering capacity [Tietge&Vughs 1995].

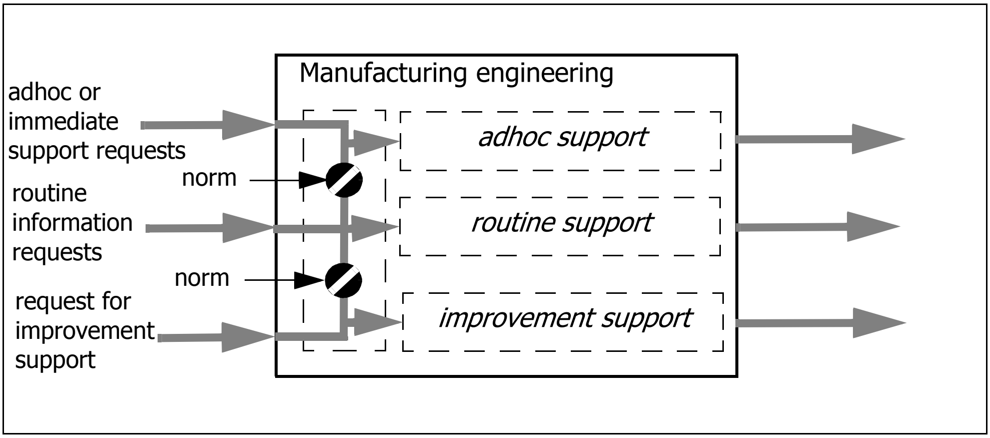


Figure 2.16: Restructured manufacturing engineering

The situation of mixing, trouble shooting and maintenance with improvement projects is often found in manufacturing engineering departments. Trouble shooting and maintenance is the prime source of technical information for improvement projects and improving machine or process availability is the prime incentive. Yet the execution of these tasks has to comply with different sets of logistic requirements that can only be met with dedicated capacities. Serving both from the same source, will always give priority to troubleshooting and maintenance. This starts a downward spiral towards autonomous growth of the capacity demand for troubleshooting. The lack of planning and implementing improvement projects means that the existing manufacturing system is mainly being kept alive through troubleshooting and repair. The solution is to separate the capacities for troubleshooting and maintenance from improvement projects and to feed the improvement team with structured information on troubleshooting, repair and maintenance. In smaller companies this may seem difficult to implement, since these manufacturing engineering departments are very small, if they exist at all. Time division may offer a solution here; e.g. work on improvement projects on mondays and tuesdays and only do troubleshooting support from wednesdays on in a fashion that makes manufacturing survive without help through mondays and tuesdays.

2.7.2 "Wide band logistics"

A printing and publishing greeting card company that was primarily fitted for delivery to (small) retail shops directly, is experiencing problems with the handling of large orders from multi-retailers[1] and foreign agents. For at least one foreign agent, the product itself also differs from the product delivered to the retailers. The packaging is done by this agent locally conform local requirements. The multi-retailers may also have special requirements to support their internal distribution and sales policies, causing small modifications to the production process.

Although the sales department has appointed dedicated account managers for foreign agents, as well as for multiple retailers, production planning and operations did not have a dedicated process for serving these customers. For some time it was thought that the main difference for these export orders is, skip the retail packaging department and deliver straight out of flat print stock. Also multiple retailers may order quantities that completely drain some products in packaged stock. As a result some products were frequently "out of print" or completely out of stock, thus starving some of the retail shops for weeks. To the retail shops, accustomed to same-day shipping from packaged stock, this meant most orders were not shipped in full. Over 35% of these shipments, had order lines, often referring to the best selling card-designs, missing. This was unacceptable.

The solution to this problem was to reconsider the print planning process in such a way that print-on-demand became possible for export accounts. The packaging department that was used to picking from print stock, became an internal customer comparable in order size and order frequency to the export accounts. Large multi-retailer orders also received their own pack-to-order process. This was possible since most customers with large orders were happy with a two week delivery

1. Retail chains with typically between 20-100 shops.

leadtime. To qualify as large customer a significant minimum amount per order line

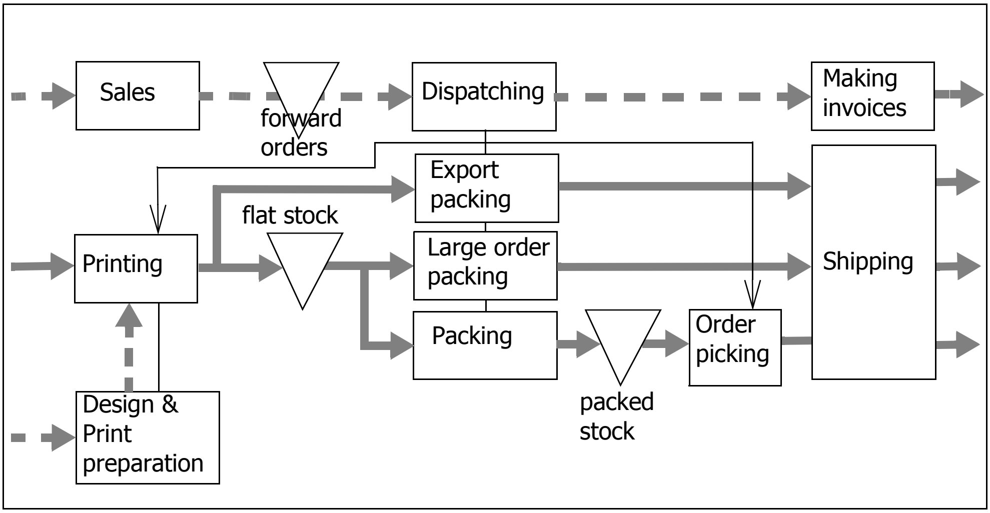


Figure 2.17: a 'single product' multiple market company

was introduced. The resulting process model is presented in figure 2.17[1]. Such solutions are capable of serving customers and markets that span a wide range of logistic requirements without losing efficiency through high work in progress (WIP) levels or excessive capacity. Note however that the notion 'single product' is not strictly true under the product notion discussed in paragraph 2.4.1. This product notion includes all logistics and packaging requirement relevant for the customer's potential to experience value. This means that the differences in packaging and logistic requirements result in three different product offers[2] for this company [Peppel 1997].

2.7.3 Smart buying

A multinational company with a very broad range of products and activities in the field of electrics, electronics and telecommunication had the need to integrate their

1. Most of the logistic filters and controls and production planning functions are not drawn in this model. The partial separation of three material streams caters for the implementation of a separate logistic control model for each stream.
2. For the sake of simplicity, the influences of different card designs and different card sizes, have been disregarded in this example.

portfolio of procurement support systems into a more consistent and coherent, company wide approach. The decision was taken to develop a framework for a supplier relationship management (SRM) system, that should be capable of company wide support of procurement processes [Mackor 2003].

The primary function of SRM is to support procurement processes.

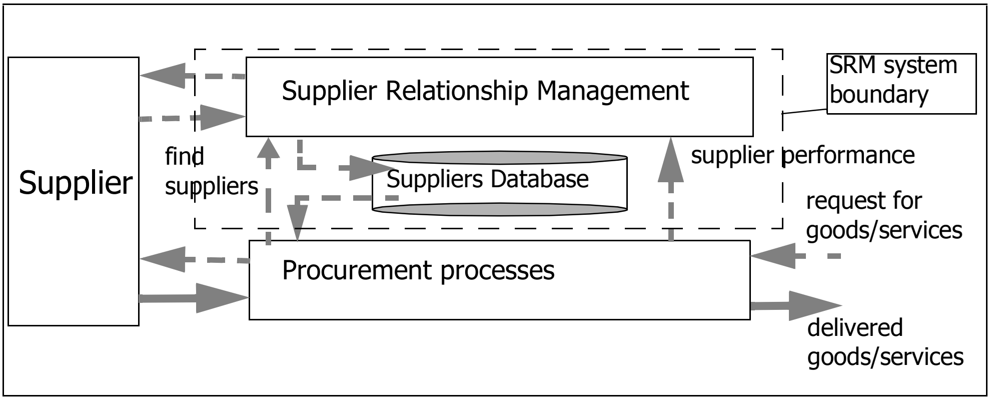


Figure 2.18: SRM and Procurement processes.

In some sense supplier relationship management (SRM), mirrors the processes and practices of customer relationship management. As a matter of fact, the supplier may have a customer relationship management (CRM) system in place to assess the importance of having this multinational as a customer. However the initial “gameplan” may be different. Where CRM seeks opportunities for cross- and upselling to strengthen the supplier’s position (make the customer more dependent), SRM may look for the same opportunities with the idea to increase bargaining power being a bigger customer. This is not what real relations are about. A good relation should support both interests of the supplier and the buyer. Together they both can create more value and achieve better efficiency in their joint business. This requires a shift from a short term procurement perspective to a long-term co-development perspective.

This change in perspective is also reflected in the definition of the value contribution of a supplier relation.

Definition 2.10 Value contribution of a supplier relation

The value contribution of a supplier relationship is the difference between the potential value contribution of a supplier and the real value contribution of a supplier.

The main goal of the SRM system is to identify a suppliers potential and to facilitate the conversion of a potential value contribution into a real value contribution. A supplier relation with a negative value contribution should be downgraded to a level where there are no investments done in maintaining the relation. This should not be interpreted as a negative supplier qualification. It simply means that there is no strategic match between future business developments from the customer perspective and the present and future capabilities of this supplier. Therefore it does not make sense to start or sustain investments in that relationship. The SRM program

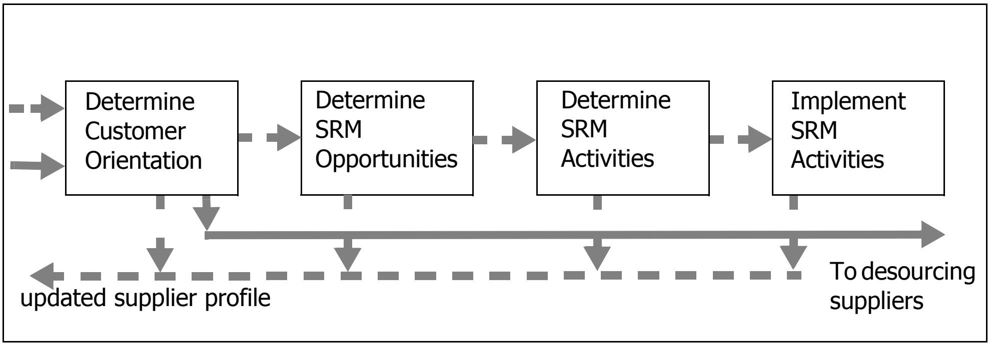


Figure 2.19: Relationship Management Program

itself is aimed at assessing and developing the potential for a value contribution from a supplier. The process that implements this program is presented in figure 2.19.

Most of these steps are information gathering steps that are needed to assess a suppliers value-contribution potential. The first step is to determine your own orientation. Without information on present and future sourcing needs, it is not possible to assess a suppliers potential. If a "quick-match" between this self-image and what is known as standard information about the potential of this supplier does not give a match, then this supplier will most likely be desourced. The next step is to identify SRM opportunities. These opportunities need not to be developed instantly. They could be related to internal future developments and sourcing needs that have not been decided yet. Determine SRM activities is aimed at defining what activities are necessary to fill information or experience gaps that prevent exploiting the suppliers potential for value contribution. Finally these activities need to be implemented. All steps taken in the SRM process and their results are recorded in the supplier database (figure 2.20).

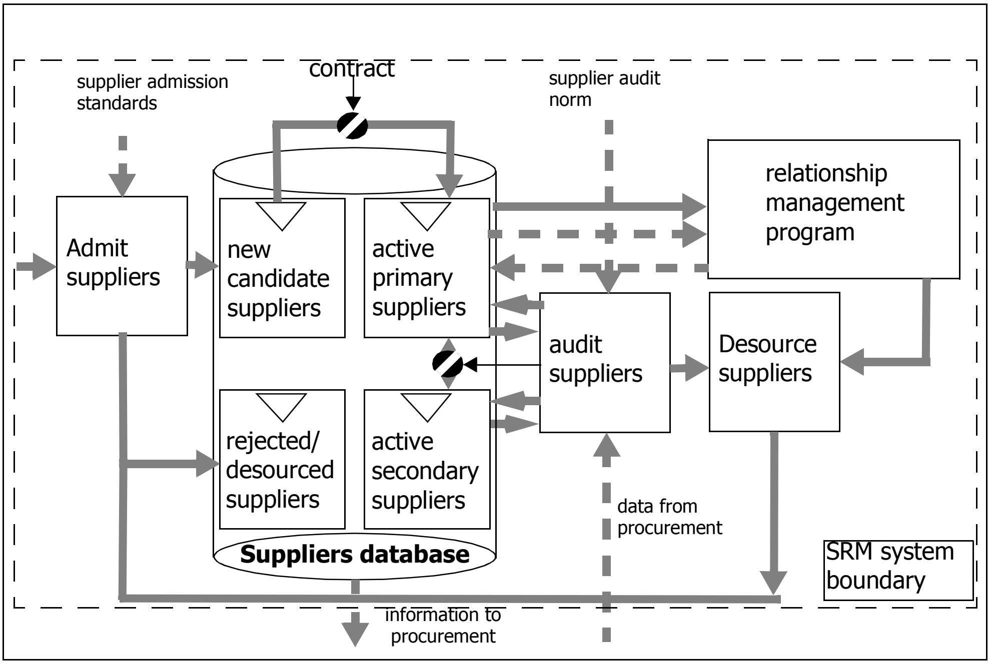


Figure 2.20: Supplier relationship database management system.

Like real relationships, the SRM program activities are intense and costly. These costs should be justified by the increase in value contribution potential. To increase the efficiency of this process in principle only active primary suppliers will be involved in the SRM process. In addition regular supplier audits are used to categorize all active suppliers. Only in case of a real shortage of active premier suppliers for a strategic need, secondary suppliers or non-active suppliers may become engaged in this process. Suppliers that become desourced will not disappear completely out of the system. A limited amount of information is kept in a database section labelled desourced or rejected supplier. This information helps to signal other divisions not to start doing business with these suppliers in case of bad past performance or other trouble of some kind. If the desourcing was only caused by a temporary lack of strategic interest then this memory can provide the information necessary to make a quick start in requalifying this supplier.

The procurement process (figure 2.21) makes use of the information in the suppliers database and report to the supplier audit function of the SRM system.

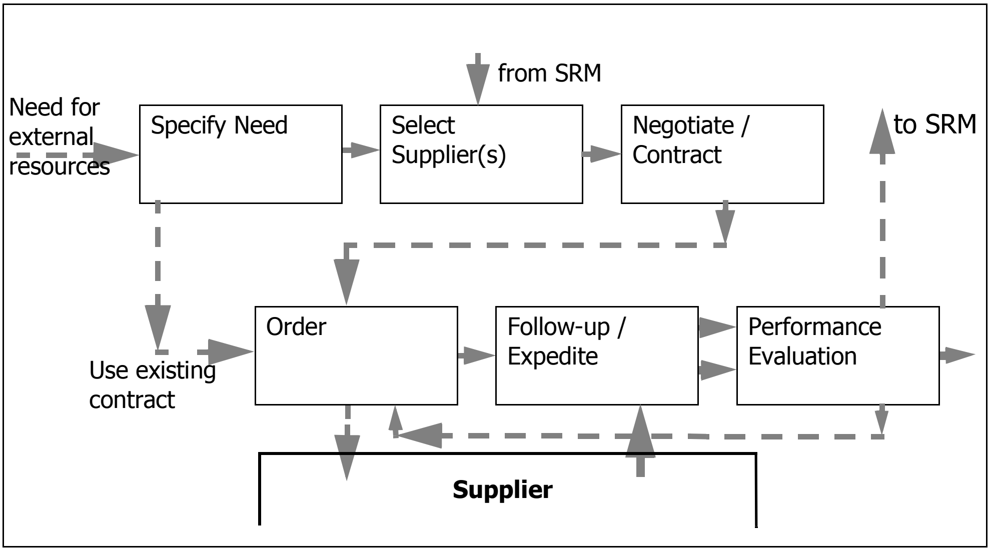


Figure 2.21: Procurement process

After having developed these processes, a typical organization structure problem surfaced. Differences in value propositions between various divisions require that each division would get its own copy of the SRM system. However, suppliers that are strategic to one division could be involved in the supply of commodities to another. In other words a single supplier can have a different supplier status across different business divisions. This also means that one division could deploy a commodity sourcing strategy that could be potentially harmful to the strategic interests of another division sourcing a different product from the same supplier. How to cope with these sometimes conflicting interests?

The following solution was proposed. Corporate level coordination and control are needed to maintain process congruence regarding the entire procurement system across all divisions. The database is shared and made transparent across all divisions to maximize the potential for cross buying and learning across divisions regarding reuse of contracts, practices etc. Corporate level coordination is also responsible for supplier acceptance and auditing standards. The SRM programs and desourcing functions are specific for each division, but they are executed in shared mode under corporate supervision, to implement and maintain a "one-face" to the supplier policy. The interests of divisions that source from secondary suppliers are included in the relationship management program if this supplier is engaged in other business as well that is categorized as a premier category supplier. The audit functions are standardized but not shared. This caters for auditing against standards specific for each division.

The resulting organ structure is presented in figure 2.22. There are multiple copies of the SRM process of figure 2.19 plus corporate level supervision. In figure 2.22, the shared and coordinated functions and database are filled with a grey pattern to indicate that they are shared and operated under corporate supervision.

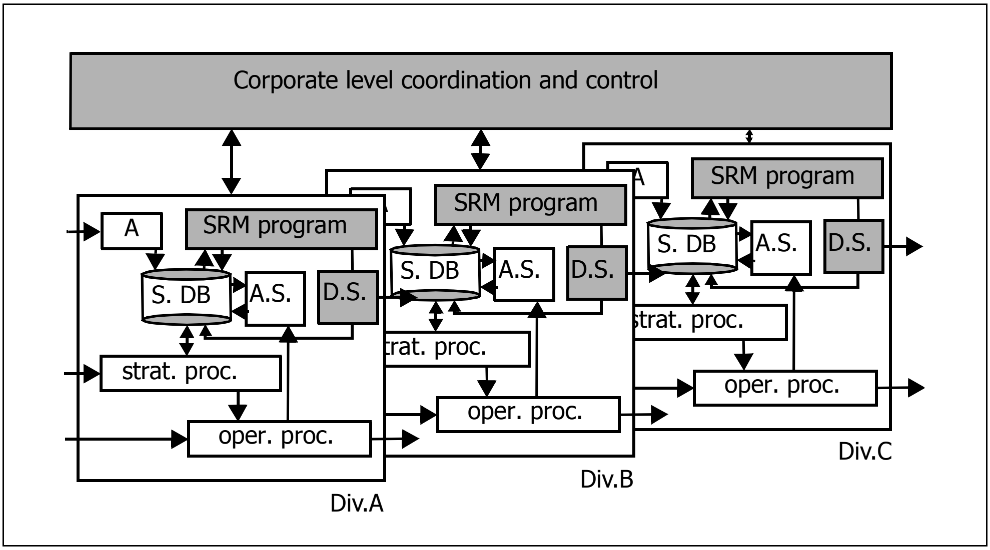


Figure 2.22: Organ structure, corporate wide procurement system.

2.7.4 Discussion

All cases discussed in paragraph 2.7 are mainly illustrations of the importance of the first structure design rule: *Do not combine value propositions that are too far apart into one process.* To cater for differentiated control strategies, separate processes were designed that resulted in additional controls to maintain tailored sets of performance indicators, one set for each stream. At the same time the smart buying case illustrates that being able to act consistently as one company also requires coordination functions and constraint management to be implemented. These functions have a profound influence on the resulting organ structure of the set of supplier relationship management systems. Similar problems can be found when implementing ERP systems, CRM systems and HRM systems for companies, customers and workers, engaged in multiple relations or should we say value-propositions.

For the sake of agility, each business division and production facility need to be in control of their own operations. This usually results in for instance multiple ERP-domains. Yet, being part of one company and to facilitate efficient aggregation of financial performance, these domains need to be coupled in some way through supervisory systems.

The same principle applies for CRM and HRM systems. Individual customers or workers may be engaged in multiple relations with a company. Agility requires that controls regarding most aspects of these relations should be implemented at operations level. However without coordination this would result in multiple policies and practices regarding the same customer or worker.

3. TOWARDS QUANTITATIVE SUPPORT FOR STRUCTURE DECISIONS

The structure design rules of chapter 2 are aimed at creating organs capable of executing tasks and offering sufficient capacity for handling the complexity of these tasks. As a general rule, the complexity passed between organs should be much less than the complexity handled within the organ. Complexity can be expressed in terms of variance as a measure of uncertainty. Reducing variance is one of the key practices in improving the performance of logistic processes in general and manufacturing processes more in particular. The first rule for structure design is aimed at reducing demand variance through splitting a multi-market job stream into separate streams that can be served by "tuned" processes, each running at their own pace. A single cell queue server simulation experiment that includes order leadtime tracking has been carried out to provide support for this rule. In this experiment, the main focus of these experiments was leadtime. Economically, a short leadtime is rewarded by the customer as well as internally through cost reductions. If multiple markets require different leadtime regimes, then the tightest regime dominates or a multiple queue-server solution must be considered.

3.1 Introduction

The main thinking behind the rules for structure design was to segment the full strategy of a company into parts that could be served with simple dedicated processes. If a company wants to operate a complicated, many faceted strategy, then the company would have to operate a (large) set of simple dedicated processes. Of course these processes would share many commonalities, such as the way they are controlled.

The goal of this chapter is to provide quantitative support for structure design decisions. Ideally one may want to define some figure of merit that could be used to optimize process structures. For this thesis dealing with complexity is the main focus. As argued in chapter 1, Shannon's complexity models the combination of uncertainty, diversity and interrelations that may cause difficulties for managers or systems to successfully accomplish their task. Therefore complexity should also qualify as base for a figure of merit.

However structure changes can have a very large and non-linear effect on control states, behaviour and performance of the system. Especially if the control states are changed one could question the comparability of both systems. Thus a sequence of structure designs is not expected to be correlated positively with a monotone sequence of complexity scores. The theoretical minimum for complexity is zero, but this corresponds to a situation where there is no uncertainty. Such a situation cannot be achieved through structure methods alone. So, in practice we do not know the complexity minimum, but as long as the complete set of control states is not changed, we can compare different designs on a relative scale.

The structure design rules are applied in the context of a set of process designs. But at the same time process design changes may be invoked as well. If an organ boundary is placed over a relation that shows a high dynamic range possibly accompanied by a high level of uncertainty, a buffer can be placed in this link to relax the complexity attached to this relation.

The structure design rules can be applied at any aggregate; at the level of a local manufacturing cell, at factory level but also at the level of a division of a multi-national company. The same applies for the complexity score, but it is clear that the numbers have a different meaning and interpretation, depending on the aggregation level. But the modelling and the decision making is the same. At each level it is possible to define a blackbox model that has a buffer and capacity to provide a specified functionality to its surroundings. The focus in this chapter is mostly on the level of a single queue server model as the key recurring model in structure decisions.

To provide quantitative support for structuring decisions means two things. First it must be demonstrated that a complexity based figure of merit can be used to reason about structure decisions. Secondly it must be demonstrated that there is a physical or maybe logistic reality behind the problems indicated by the complexity figure of merit.

The first part of this chapter is devoted to demonstrating the impact of complexity. Through modelling the behaviour of processes as a Markov process (a sequence of states) and using the probability distributions of the states as input for a complexity figure, it is possible to build a theory how to decompose a large collection of states into subsets that are less complex and that can be handled independently. However the decomposition is not for free. Decomposing problems requires coordination of the interfaces between the subsets. Looking at the symmetry[1] of these coordination problems it is possible to build a set of theories on how to deal with the coordination. In addition, a cost model is proposed that relates the coordination cost to the communication distance in organisations. This cost grows more than proportional with the length of the coordination path, which implies that subsets that require intense coordination should be located close to each other.

As an example of applying a complexity measure for the analysis of manufacturing systems, the work of Sivadasan on complexity in supplier-customer systems is described [Sivadasan 2001,2004].

In addition the optimal span theory of Glickstein offers another complexity based figure of merit for the intricacy of a hierarchical structure. The intricacy of a hierarchical structure, which could represent an organisation, is proportional to the complexity handling capacity of that structure. If dealing with the imposed complexity requires coordination and communication within an organisation, the optimal span theory provides a rule for a communication optimal organisation.

To demonstrate that there is a physical or logistic reality behind complexity, the second part of this chapter is dedicated to the performance of the single queue server model. This model is described and analysed with queuing theory. In simulation, the model is loaded with a mixture of order patterns to demonstrate how the uncertainty over the maximum leadtime performance, increases as the characteristics of the order patterns are further apart. This is a situation the first structure design rule is trying to prevent.

1. A symmetrical coordination problem means that the mutual-dependency between departments and/or subproblems are equal. An asymmetrical coordination problem means that one subproblem depends on the other, but not vice versa.

3.2 Quantitative use of complexity

Shannon's information measure of complexity is based on a state model of the system we want to consider. Each state has a probability of occurring. Depending on the present state, some future states are more likely than others. Markov processes or Markov chains are commonly used to model such systems. Markov processes support both the use of Shannon's complexity measure as well as the application of queuing theory. A Poisson process is a special case of a Markov process.

3.2.1 Markov processes and Markov chains

Markov processes are stochastic processes for which the outcome depends one or more previously occurred outcomes. Systems that can be modelled by Markov chains exhibit processes that have memory and that are subject to stochastic influences. For a Markov chain of order 1, the outcome depends on only one, immediately preceding outcome. In this case the following condition holds:

$p\langle X_t | X_{t-1}, X_{t-2}, X_{t-3}, \ldots\rangle = p\langle X_t | X_{t-1}\rangle$.

This notion can be extended to so called Markov-chains of order k. For a Markov chain of order k, the probability of X_t depends on k, preceding values of X_t:

$p\langle X_t | X_{t-1}, X_{t-2}, X_{t-3}, \ldots\rangle = p\langle X_t | X_{t-1}, X_{t-2}, X_{t-3}, \ldots, X_{t-k}\rangle$.

Because the probability ofX_t depends on the set S_t of k preceding values, this set $S_t = (X_{t-1}, X_{t-2}, X_{t-3}, \ldots, X_{t-k})$ is called the state of the Markov-chain. If we consider the state of a Markov chain as a stochastic variable, then regardless of the order of the Markov chain, the next state will only depend on the previous state. Thus a Markov chain of order k can be reduced to a Markov chain of order 1 [Lubbe 1997, p80]. If we assume that the stochastic variable X_t, can have m possible outcomes, then a Markov chain of order k, can have m^k different states as every state is determined by k symbols each chosen from m possibilities. Since after each state a new choice can be made from m possibilities, the number of state transitions is m^{k+1}, each with an associated state-transition probability. For each state, the sum over all transition probabilities should be 1:

$$\sum_{j=1}^{m} P\langle S_j|S_i\rangle = 1 \qquad \forall i \qquad i \in \{1, 2, \ldots, k\} .$$

For a Markov chain we can draw a state diagram, indicating the states and transitions with their transition probabilities. Note that $P\langle b|a\rangle$ is not necessarily equal to $P\langle a|b\rangle$.

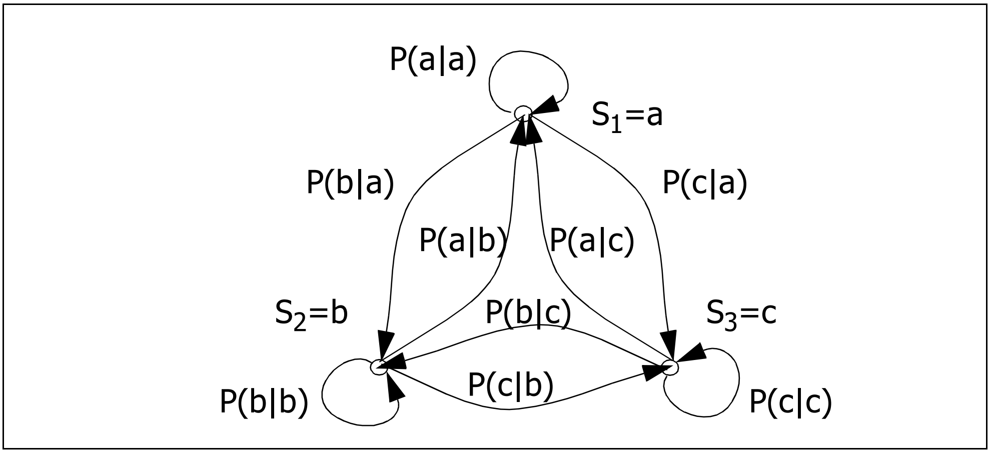


Figure 3.1: State diagram for a Markov chain of order 1 and *m*=3.

For the system of figure 3.1 we can calculate the probability of individual states as follows: $P\langle S_i\rangle = \sum_{j=1}^{m} P\langle S_j\rangle P\langle S_i|S_j\rangle \qquad i = 1, \ldots, m$ and $m = 3$.

Markov chains are widely used in information theory for modelling coding and communication systems, language and finite state machines as well as manufacturing systems. If one imagines in figure 3.1 that a, b and c are codes for, *a=idle, b=busy and c=not-functioning* and we add triggers that can cause state transitions, then this Markov state diagram represents a simple job-processing system.

Markov process models represent a reality that can be linked to complexity by considering the set of probabilities of the state-transitions.

3.2.2 Some simple structure experiments with complexity

Consider two state diagrams with probabilities assigned to the final states. Suppose all states represent independent events. Both systems have the same starting point and set of possible outcomes, but the second system also has an additional intermediate state with a probability as well as transition probabilities to and from this intermediate state.

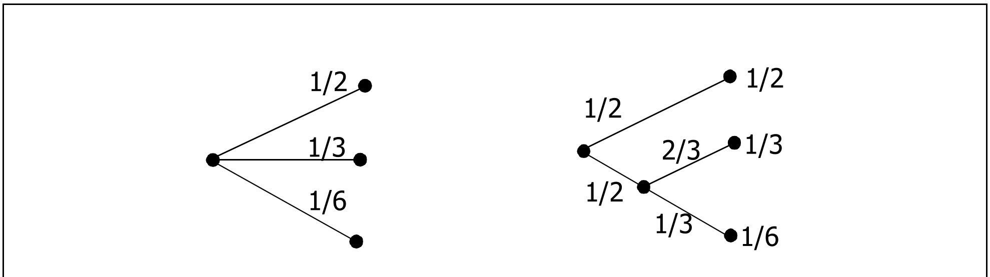


Figure 3.2: Two simple state diagrams with probabilities

The overall complexity of both systems is the same, but for the second system we can use the intermediate state to reduce the complexity of decisions at all starting and intermediate nodes. Shannon's entropy H for this system is defined as follows[1]:

$$H = -\sum_{i=1}^{n} p_i \log p_i \tag{3.1}$$

1. Note the base of log is 2 and this will be so throughout this thesis.

For the first system this means $H=H(1/3,1/6,1/2)=1.46\ bit$[1]. The total entropy for the second system is the same, but it is composed differently: $H=H(1/2,1/2)+1/2*H(1/3,2/3)=1+0.5*(0.92)=1.46\ bit.$ Introducing the intermediate state converted one problem of $1.46\ bit$ into two aggregated problems of 1 and $0.92\ bits$. This is exactly what segmenting or structuring can do. A three-state situation can be decomposed into two two-state situations that are less complex to handle. Regardless of the values of the state probabilities, the maximum

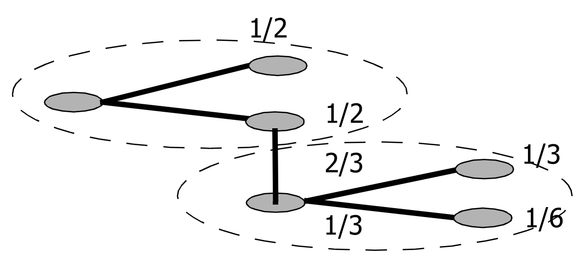


Figure 3.3: Aggregated problem structure

complexity[2] is only a function of the number of states N. For the above situation this

$$H_{max} = \log N \qquad \textbf{(3.2)}$$

means that the decomposed problems cannot have complexities of more than 1 bit.

The same structuring principle also applies for situations with higher number of states. Consider a set S containing N_s states. We can split S into k subsets such that: $S = S_1 \cup S_2 \cup S_3 \cup \ldots \cup S_k$, $S_k \neq \varnothing, \forall k$ and $N_S = \sum_{i=1}^{k} N_{S_i}$

(figure 3.4). Now if all states belonging to S are independent then all subsets S_k are independent and we can define the a-priori probability of a state x belonging to S_i as

1. Bit stands for binary digit and represents the quantity in which information is expressed.
2. Maximum complexity refers to a situation where all states have equal probability.

follows: $p_i = p(x \in S_i) = N_{S_i}/N_S$. The complexity of this system can be calculated (3.1) as follows: $H = -\sum_i p_i \log p_i = -\sum_i (N_{S_i}/N_S)\log(N_{S_i}/N_S)$, which can be rewritten as:

$$H = \log N_S - \sum_i (N_{S_i}/N_S)\log(N_{S_i}) \qquad \textbf{\textit{(3.3)}}$$

Now this is an interesting result. By changing our perspective from working with

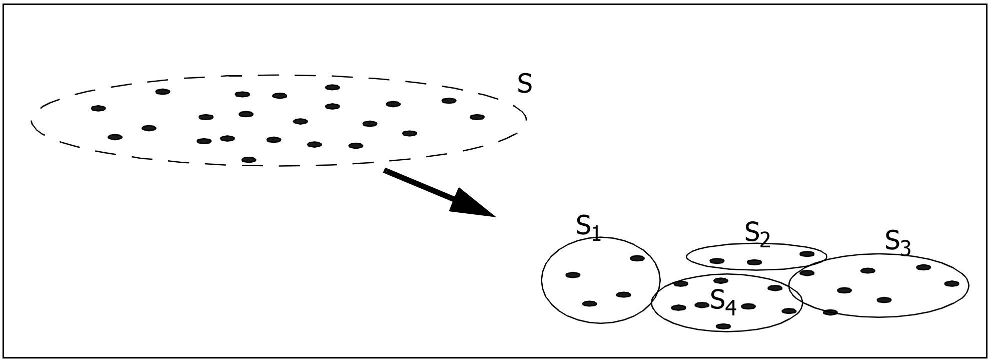


Figure 3.4: Decomposing a set of states

large set of N_S individual states (x) to a collection of subsets containing a smaller number (N_{S_i}) of states ($x \in S_i$), we have effectively reduced our complexity with the probability weighted sum of the maximum complexity we may find within each subset. In design this is a very powerful principle. If we can decompose a complex problem into a set of smaller problems by defining constraints[1] (set membership functions) then we get problems with smaller complexity in return: the problem at the highest aggregate has a complexity given by (3.3), the other problems of choosing the right state within a subset have a maximum complexity given by (3.2). Another interesting property of (3.3) is that it relates the reduction in complexity to

1. The constraints usually have the form of set-membership functions. So we need to define conditions for S_i such the $p\langle x \in S_i \rangle = p\langle x | S_i \rangle = 1$

the maximum complexity in each set. If we combine this result with the aggregation principle explained in the example of figure 3.3, then it follows that we did not loose complexity. The overall complexity remains the same. By structuring we define ways of dealing with problems of less complexity, sequentially or concurrently.

3.2.3 Structuring and decomposing complexity

The principle of decomposition works best if subsets are independent. Regardless of the internal dependencies within the subsets, it follows from (A.10) in appendix A that: $H(S) = H(S_1) + H(S_2) + \ldots + H(S_k)$. However if the subsets are not independent, with (A.9) in appendix A, we get the following:

$$H(S) = \sum_{i=1}^{k} H(S_i) + \sum_{i,j=1}^{k} H\langle S_i | S_j \rangle \qquad \textbf{(3.4)}$$

The last term in (3.4) is the information that has to be communicated between the sets to solve the dependencies. Choosing organ boundaries or in other words defining sets that will minimize this term, will also minimize the coordination needs between the organs. In fact we have two options to minimize the influence of the conditional information term. The first and already mentioned way is to redefine the sets, such that the coordination needs are contained within the sets. The second way is to define constraints such that the coordination problem is shifted from coordinating direct interaction to monitoring the space for compliance with the constraints defined.

In other words, if we can define constraints for all elements that belong to S_i such that $P\langle S_i | S_j \rangle \cong 1$, the coordination need with respect to S_j will be approximately zero as long as the constraint on S_i is not violated. In design processes this may be interpreted as defining a range of possible values for design variables belonging to this set. For processes in manufacturing logistics, this often means placing an inventory point or buffer in a material flow, such that the processes on either side of the buffer do not experience mutual influences any more. The constraint imposed on the processes on either side is that they operate within conditions that cannot cause overflow or stockout at this inventory point.

The choice between confining the coordination need or to apply constraints for limiting the coordination need can be made by studying the values and the symmetry present in the conditional information. Suppose we have decomposed the set of states S into a collection of sub-sets S_i, $i=1..k$, then for each pair i,j, depending on the values of $H\langle S_i|S_j\rangle$, we can define a matrix of theories how to deal with the coordination and resolve the uncertainty still present in the individual sets.

		$H\langle S_j\|S_i\rangle$	
		BIG	SMALL
$H\langle S_i\|S_j\rangle$	BIG	Strong mutual dependence. Consider joining sets or keep sets together by redefining the coordination mechanism.	S_j is strongly determined by S_i, but not vice versa. Keep sets together and solve S_i first or set constraint for S_i to decouple S_j after which both sets maybe treated independently under this constraint.
	SMALL	S_i is strongly determined by S_j, but not vice versa. Keep sets together and solve S_j first or set constraint for S_j to decouple S_i after which both sets maybe treated independently under this constraint.	S_i and S_j are (almost) mutually independent. Sets may be handled concurrently, even in different organs.

Table 3.1: Theories for handling mutual dependencies

Note however that imposing constraints is also identical to the principle of abstraction or defining aggregates. At the aggregated level a new state has been introduced to signal whether or not all sub-sets can comply with the constraints imposed on these sets. Suppose the probability that these sets can comply with the constraints is P_S, then the overall complexity of the system is as follows:

$$H(S) = -P_S \log P_S - (1-P_S)\log(1-P_S) + P_S\left[\sum_{i=1}^{k} H(S_i) + \sum_{i,j=1}^{k} H\langle S_i | S_j \rangle\right] \quad \textbf{(3.5)}$$

The upper term represents the aggregated "compliance" problem, the lower term represents the complexity of the (sub)sets that function within the constraints. Since P_S is usually close to 1, $H(S)$ is largely determined by the lower term. The advantage is that both aggregates may be dealt with independently and the lower term may be smaller than the unconstrained version, since some combinations may contain zero complexity under the influence of the aggregate constraint.

3.2.4 Application: complexity in supplier-customer systems

Suja Sivadasan applied an information theoretical complexity measure to measure the *operational complexity* of supplier customer systems [Sivadasan 2000, 2004]. In this work a distinction is made between *structural complexity* and *operational complexity. Structural complexity* of manufacturing systems is defined by Frizelle and Woodstock as the complexity associated with the static variety characteristics of the system; that is, the complexity without taking into account, the influences of control, disturbances and wait-queue states [Frizelle 1995]. Roughly speaking the static complexity comes down to the number of organisational units times the base 2 logarithm of the number of product-specifications per manufacturing unit. Thus structural or static complexity is merely a type of coding complexity for the product-process combination. In the face of the complexity definition adopted for this thesis (definition 1.1) and in the context of Shannon's complexity measure, the definition of static- or structural complexity may not be correct. The source of uncertainty is not clear or uncertainty may not even be present. Yet, structural complexity is considered to be a part of the *dynamic complexity* which does include sources of uncertainty and which is closely linked to the concept of *operational complexity*.

For modelling the operational complexity of supplier-customer systems, Sivadasan considers the following "state" variables:

IN CONTROL	*Probability of being within manageable tolerances*
FLOW VARIATIONS (*I*)	*(F) flows: forecast-order and delivery-order*
PRODUCT CATEGORIES (*J*)	*(U) product categories.*
OUT OF CONTROL STATES (*L*)	*(NS) situations of being out of control related to volume (short or excess), timing (early or late) and difficulty of returning to the in-control state.*
REASONS (*K*)	*(R) groups of reasons (here 3): due to supplier (S), due to customer (C) or other (O)*

Table 3.2: State variables of operational complexity [Sivadasan 2004].

The amount of information necessary to describe this supplier-customer system is then given by [Sivadasan 2004].

$$H = -\sum_{i,j=1}^{F,U} [P_{ij}\log P_{ij} + (1-P_{ij})\log(1-P_{ij})]$$
$$-\sum_{i,j=1}^{F,U} (1-P_{ij}) \sum_{k,l=1}^{R,NS} Poc_{ijkl}\log Poc_{ijkl}$$

The structure of this system is similar to that of formula 3.5. The upper term describes the complexity of the aggregated in-control/out-of-control problem (P_{ij} is the in-control probability, Poc is the out of control probability). The lower term contains the additional information necessary to determine why and how the system is out of control and what measures are required to regain control[1]. The lower term is defined as the *operational complexity* of a supplier-customer system.

Since this measure of *operational complexity* is mainly determined by the out of control states, it also reflects what managers experience. The exceptions and problems cause the worries and stress. With the analysis of the drivers of the operational complexity, one can pinpoint to a certain degree, where the problems originate.

However, what measures to take, still depends on the experience of the operator. The out-of-state levels, reasons and flow variables are only observable at the interface between customer and supplier. These observables are not directly linked to internal[2] control-states that may have caused the problem. Experienced operators have a short-list of probable causes to be checked.

If the system is subject to long leadtime then the operators task is even more difficult. A long leadtime obscures the relation between the order-forecast flow and the order-delivery flow at a later moment. If forecast quality is poor, then delivery performance will degrade as well if production planning is based on unreliable forecasts. But this degraded performance only surfaces long after the forecast failures occurred.

3.3 Structure design rules and complexity

From paragraph 3.2.2 it is clear that complexity is mainly driven by the number of state-transitions and their probability. Less state-transitions means less complexity. However, uncertainty is captured as probability and uncertainty also drives how many state-transitions we need to model the uncertainty correctly. Preferably, each cause for uncertainty is linked to its private state and state-transition. If this can be

1. Note however the choice of perspective, relative to formula 3.5. Here the focus is on out-of-control or non-compliance, rather than compliance and operations within compliance.
2. Control states that are local to either the customer or the supplier.

accomplished, then making that state active, also activates the transition to a state where the uncertainty is cured. If more causes of uncertainty become mapped onto one single state with multiple transitions, we may not have a clue how to recover from the uncertainty signalled by that state.

If the structure design rules help to reduce complexity, then the rules must support the reduction of the total number of states and state-transitions to be dealt with simultaneously while preventing conditional information or uncertainty to be introduced as a result of a structure decision. For convenience, the structure design rules are repeated in table 3.3.

1	*Do not combine value propositions that are too far apart into one process.*
2	*Identify organs that can be responsible for a distinct contribution, but do not cut important control cycles.*
3	*While maintaining 1 and 2, try to create economies of scale (efficiency).*
4	*While maintaining 1 and 2, try to achieve some flexibility and reduce the vulnerability associated with small departments and product oriented structures by merging these (support) functions into larger multi-service units.*
5	*If after 1-4 value propositions are left that do not add value to the company, outsource them or give up these lines of business.*

Table 3.3: Structure design rules

The *first rule* aims at avoiding uncertainty about leadtime and margins when customer orders are executed. Suppose we have complete information on sales price, delivery date, product- and packaging specifications, then we can make a planning for that product, minimizing the use of company resources, while reliably delivering that product. Such a plan would also maximize the margin on that product. The problem is how to make such a plan if there are other orders like this one, but also ones that are very much different competing for the same limited company resources. If we focus on margin, then orders with a lower margin may be

postponed, thus generating overdue states and possibly penalty fees. If we focus on delivery reliability but forget about sales value, then a lot of extra capacity is needed to guarantee a maximum order leadtime.
Mixed priorities lead to scheduling problems that may cause the Markov-chain of sequenced operations to develop into higher orders that cannot be decomposed. As a result, the coordination terms in formula 3.3 will start to dominate the complexity of managing operations.

The *second rule* aims primarily at minimizing the coordination needs between organs as indicated by the conditional information in formula 3.4. But the argument goes even further. The organ structure is the basis for building the personnel structure. Human based, inter-department coordination and communication suffers from a number of defects that can easily induce additional uncertainty[1]. Different departments have different value systems and priorities, which could lead to uncertainty in coordination. Their language and understanding may be different. Human based communication often suffers from variable leadtime. The discipline to answer to asynchronous "calls" (email, letters, fax, voice mail) may vary from person to person. Variable leadtime in control cycles could easily lead to instability of the system controlled by that cycle. When modelled, variable response times and uncertain interpretation of coordination messages leads to additional states and thus raises complexity.

The *third rule* may actually cause interdependencies between orders from different markets. Sharing the same resource, orders may have to wait for one-another. However, these interdependency effects on leadtime performance may be dealt with through scheduling or sequencing techniques, but at the expense of extra states and additional complexity. Economies of scale are essential in minimizing the use of resources per order. Economies of scale can reduce dynamics and uncertainty in the order flow if the relative dynamics and value propositions of these orders do not differ too much. The advantage of looking for economies of scale after applying the

1. See also paragraph 5.6

rules for splitting processes and keeping control together is that now we can compare the investment in scheduling techniques to compensate for leadtime dependencies with the cost of having separate flows with no interdependencies.

The *fourth rule* is aimed at reducing the uncertainty that may be caused by individuals in a small department. If there are only few of them, the department may become vulnerable to absence of one of them. As a result variance in output and leadtime is induced. Merging small departments into larger multi-service units may solve this problem as long as scheduling, prioritizing between different services is not too difficult and as long as personnel can be trained to execute multiple services. If this kind of training is not possible, then it does not make sense to combine such units, except maybe for social and safety reasons.

Finally the fifth rule supports the first rule by suggesting that there may be other ways outside the company to implement customer wishes or to reduce the variability of customer requirements by not acknowledging every customer's request.

3.4 The cost of coordination

Coordination requires communication. Communication is a process that requires coding C, transmission making use of a medium and decoding DC of plain messages. The meaning of these messages is driven by the intent or interest of the sender and filtered by the value system V of the receiving party (figure 3.5). A more detailed explanation of this model is presented in paragraph 5.6.

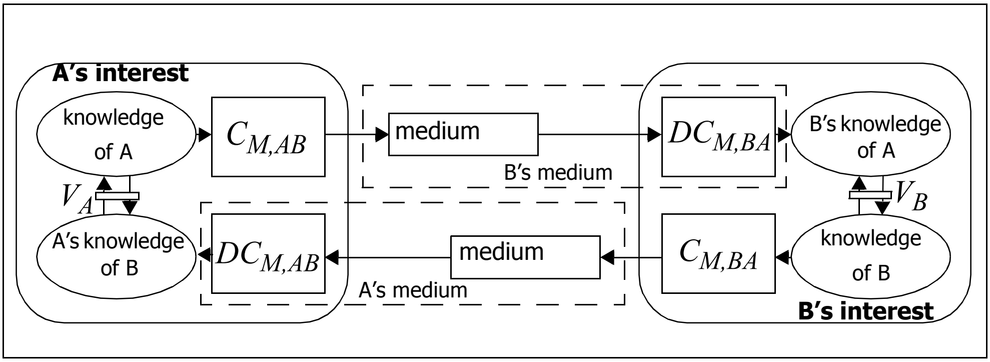


Figure 3.5: Communication Model

Referring to this communication model we may argue that the coordination cost between departments or organs are at least three times as high as the coordination cost within departments. The coordination requires, one full cycle. This means the coordination is initiated by one party and is confirmed by the receiving party. Assuming that the influence of coding, the media and decoding can be neglected and assuming that parties A and B have a shared interest in communicating, then succesfull coordination requires in addition at least two full cycles to adjust both value systems.

Within organs, the investment in adjusting value systems, coding/decoding conventions are being done once, at the initiation of the organ. Thus the recurrent part of internal coordination requires only one full communication cycle. Between organs or departments these initial adjustments are not being done as part of their initiation. Even if these value system adjustments have been done, we cannot expect them to last. Especially if both departments are driven by financial incentives that are not constrained to support the overall profitability of the company, it is to be expected that values as well as interpretations will differ. Thus if coordination is necessary to achieve a common goal, we can expect the two additional communication cycles to realign the value systems.

If the communication path includes more departments, then all of these need to have mutually aligned value systems. Thus the coordination cost becomes proportional to the number of different pairs of value systems that need to be aligned. Each pair requires at least two cycles to accomplish the adjustments necessary.

In a hierarchical organisation we may assume a linear path between two units that may need to coordinate. We can model a transition between two departments as adding one unit length to the length of the coordination path l. The number of value systems involved in a coordination path of length l, equals $l+1$. It is assumed that the cost of completing one cycle does not depend on the purpose of executing that cycle. This cost is denoted as C_u, the unit cost of one full communication cycle. Thus we can model the cost of coordination (C_{coord})[1] as follows:

$$C_{coord} = C_u\left(l + 2\binom{l+1}{2}\right) \qquad \textbf{(3.6)}$$

As the length of the coordination path increases, the cost of coordination expressed in multiples of C_u grows more than linearly with the coordination distance. This may not even be a worse case, since aligning value systems may take more then two cycles. However under the condition that company goals and incentives are clear and transparent and that incentives at department level are constrained to remain aligned with the overall goals, serial communication and coordination may be sufficient. In such a case, the coordination cost model reduces to a linear model (l*3) or even better a linear model for value system coordination and only 1 cycle for the detailed coordination directly from department to department (1+2*l) (figure 3.6).

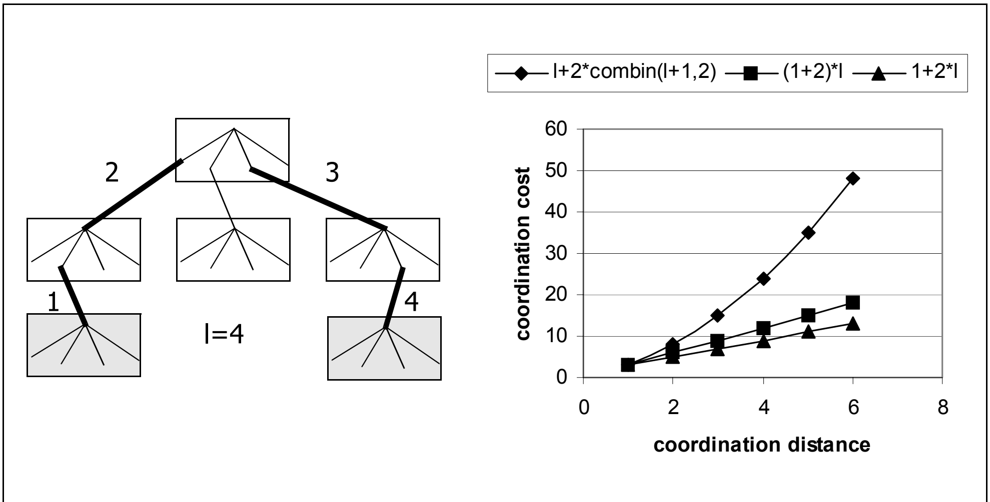


Figure 3.6: The length of a coordination path and cost of coordination

1. This is the minimum cost. Negotiation tactics that require more cycles would cause a higher number in front of the binomial coefficient. Furthermore it is assumed that we can reach an overall adjustment by mutual adjustment. Additional concurrent fine tuning in triple, quadruple or sets with higher numbers is not taken into account either.

In addition there is the problem of the time it takes to do these adjustments. If we assume that the coordination cost and the time needed for the coordination are proportional[1], then even if the coordination cost are acceptable, the coordination time may well be beyond limits dictated by the dynamics of external disturbances and internal control needs.

We may interpret adjusting value systems in various ways ranging from mutual adjustment of mental processes between two individuals to the involvement of two departmental managers agreeing on a framework for the actual coordination that is to be implemented by the workers[2]. In line with the findings of paragraph 3.2.3, this means that between departments it is more efficient to work with a framework that is built on constraints that reduce the immediate coordination needs to zero and keep all issues where this is not possible within department boundaries.

For adjusting value systems between individuals as well as between departments, one has to realise that these adjustments should be mutually beneficiary. Thus, if the rules of financial budgeting and control do not support such adjustments, then these adjustments are not likely to be accomplished.

3.5 Glickstein's hierarchical span theory

In his hierarchical span theory, Glickstein uses information theory as a figure of merit for the information holding capacity of an organisation. This theory states that for any degree of connectivity within a group, there is an optimal group size, specifying the number of members needed to maximize the amount of information per node. The amount of information that can be held by that group is named the intricacy of that group, I_g.

---

1. The cost-model is based on unit-coordination-cost and the number of different value system pairs involved in the coordination. Thus we can exchange cost and time in this model.
2. The unit cost of coordination C_u depends on the type of coordination. C_u between two individuals will be lower than C_u between departments.

Intuitively, if all nodes of a group are fully interconnected, there is no routing specifically coding information to be held by that structure. All routings are equally likely and the corresponding intricacy is low. In fact this type of structure cannot hold pieces of information bigger than the individual node can hold. The same situation occurs when there are no connections at all. Thus the organisation of the structure determines how much information can be held by that structure. The optimum holding capacity occurs when there are not too many and not too few connections between the nodes. This argument will be explained later on in this paragraph. The number of connections in relation to the maximum number of connections for a given node is defined as the connectivity ratio.

Note that this situation is very similar to the use of probabilities in the definition of Shannon's measure of information. In fact a connectivity ratio of 1 is complementary to a ratio of 0, like the probabilities 1 and 0. This means that the connectivity ratio is treated as a kind of probability of being connected. If we then apply Shannon's entropic measure of complexity on these probabilities, we get a measure for the coding or information holding capacity or *intricacy* for this structure. This is how Glickstein, started his theory of optimal group sizes and levels of interconnectivity. In the next section the optimal span theory is explained. The notations come from Glickstein's 2003 paper [Glickstein 1996, 2003].

Consider a group of S connected nodes. Assuming unique, equally weighted, bi-directional links, the maximum number of unique edges M, to interconnect S nodes is: $M = S(S-1)/2$. If we know the average node degree[1], $\overline{D}$, we can compute the average number of unique edges (A) per group connecting the nodes: $A = S \cdot \overline{D}/2$. A/M is the connectivity ratio. Now suppose we have a graph with S nodes. This graph is intended to hold or codify information. If we consider all nodes equivalent, then the information to be stored in this graph is proportional to the

1. The node degree is the number of connections per node.

number of nodes and the marginal information per node. By definition, *the intricacy of a group,* I_g of S_g nodes with on average $\overline{D_g}$ connections per node, is:

$$\begin{aligned} I_g &= -S_g \cdot (A/M) \cdot \log(A/M) = \\ &-S_g \cdot (\overline{D_g}/(S_g-1)) \cdot \log(\overline{D_g}/(S_g-1)) \end{aligned} \quad \textbf{(3.7)}$$

The total intricacy for a hierarchy of G, groups can be calculated as well:

$$I_{total} = \sum_{g=1}^{G} S_g \frac{\overline{D}_g}{S_g - 1} \log \frac{\overline{D}_g}{S_g - 1} \quad \textbf{(3.8)}$$

Now for very low connectivity ratio's I_g approaches 0. For a very high connectivity ratio ($A/M=1$), I_g also equals 0 like already explained.

Would there be an optimal value for the connectivity ratio that maximizes the amount of information per node for a group of size S_g? For this we divide I_g by the number of nodes S_g and try to maximize the remaining part. The maximum for I_g/S_g occurs when $A/M = 1/e$, where e is the natural number. From this it follows that the optimal group-size or span S_{go}, for a given average number of connections ($\overline{D_g}$) can be determined with the following relation:

$$S_{go} = 1 + \overline{D_g} \cdot e \quad \textbf{(3.9)}$$

The next question is, is there also an optimal value for the average number of connections, $\overline{D_g}$, such that I_g reaches a maximum value regardless of the size of the group, S_{go}. To prove that this is possible, Glickstein introduces a recurrent cost criterion to define some kind of operational productivity per node. Through simulation it is demonstrated in his thesis that regardless of the cost factor and the group size, the group intricacy I_g maximizes for $\overline{D_g} = 2$.

With this result, the optimal span theory is taken one step further:

$$S_o = 1 + 2e \approx 6.4 \qquad \textbf{\textit{(3.10)}}$$

In less formal language, Glickstein states the optimal span hypothesis as follows: *"All else being equal, hierarchical structures with spans in the range from five to nine yield the maximum possible intricacy with the minimum possible set of resources (i.e. components and connections), and therefore tend to be competitively selected"* [Glickstein 1996, p.9].

There is an alternative way to show that the optimal average group connectivity $\overline{D_g} = 2$. The group intricacy I_g reaches its maximum value if the connectivity ratio: $A/M = 1/e$. A graph has an information holding capacity that depends on the connectivity between the nodes. Thus we can argue that this connectivity ratio needs to be minimum for I_g to be maximum. The minimum value for the connectivity ratio for a group of size S_g is as follows. To be connected and capable of joining in a collective information holding process, the minimum number of connections per node equals 1. So the minimum number of connections per group equals the number of nodes in this group S_g [1]. The maximum number of unique bi-directional connections is: $M_g = S_g(S_g - 1)/2$. The minimum ratio is now: $\frac{S_g}{S_g(S_g - 1)/2}$.

1. There seems to be 1 open (extra) connection. This connection is needed to connect to other groups. In case managers were identified, these were not counted as member of the group, but as member of the super group they are part of as well. Each member needs at least one connection to be connected to the manager.

If we make this result equal to the "real" connectivity ratio $\frac{\overline{D_g}}{S_g - 1}$, we can conclude that $\overline{D_g} = 2$ yields the minimum ratio that gives a maximum value for I_g. Thus the optimal value for $\overline{D_g}$ was present all along in the definition of intricacy and the use of Shannon's information measure as a basis for *intricacy*.

This result has many implications. Glickstein's optimal span hypothesis offers an additional insight in the organisation and evolution of many structures, both from nature as well as man-made. This result can also be seen as an additional explanation for the result presented by Miller in his famous paper "The magical number 7 plus or minus 2" [Miller 1956]. But this is not a real surprise; The result of Miller's work "plagued" Glickstein long before he even started this research.

This is also where the theory stops. In Glickstein's theory we could link the connectivity between nodes to the need for coordination, but this is not the same as the principle for unification of control that Mintzberg uses and that is at the heart of defining organ structures. Glickstein's theory explains what would be an optimal hierarchical organisation of almost identical resources for a complex task, such as software development [Glickstein 2003].

3.6 Resume, organisation of complexity theories

Through modelling management and control problems of business processes as Markov-chains and states in a probability space, we are able to use Shannon's information entropy as a reference for complexity. Taking structure design decisions to ease the control of business processes is like segmenting sets of states into smaller but connected sets. With Shannon's information entropy, smaller sets also means less complexity.

Through analysing the conditional information that models the coordination between sets, we can decide for specific pairs of states if these pairs can be treated by different organs or else we better keep this pair together in one set. In addition, the cost of coordination between organs may bias this type of analysis towards a limited number of sets and thus large organs.

There are two limiting theories to prevent this favour for large organs. Glickstein's theory of optimal spans sets a limit to how much complexity may be accumulated into one organ. If internal coordination is required for performing the task or function of the organ, the span or group-size should be between 5 and 9. In combination with the coordination cost model, the optimal span theory also implies that we should apply methods such as defining constraint frameworks to:

1. *limit the degree of inter connectivity.*
2. *allow the application of optimal group-sizes.*
3. *minimize the number of value system adjustments required.*

Optimizing the group intricacy also favours linear communication and coordination paths. For a circular connected group of 7, this means that the maximum coordination distance is 4. With the coordination cost model in mind, this means that untill this distance individuals may experience a coordination effort that grows proportional with distance. Beyond 4, however the experienced effort grows more than proportional with distance; thus reaching the point where the investment in communication and coordination is considered not worthwhile any more.

To sustain their internal productivity, larger groups are likely to break up into multiple, possibly connected circles and adopt multiple value propositions. This development may not be undesirable at company level as long as the first structure design rule that states not to combine possibly conflicting value propositions into one organ is honoured. This means, these losely connected (sub)organs should be considered as separate organs. Only when there is a necessity for resource sharing, keeping these organs together as a result of the 3rd. structuring rule may be considered.

In the next paragraph, the reality of conflicting value propositions is demonstrated through looking at logistic interference of order streams that are clearly different in terms of order frequencies and job-sizes[1].

3.7 Queue server experiments

In manufacturing systems as well as business operations processes, the reality of complexity is often driven by uncertainty over leadtime, delivery performance and availability of supplies or other resources. Zooming in on these processes, one may see problems with technical product quality and manufacturing process stability. Being one of the biggest causes of waste and profit-degradation, it goes without saying that these technical and in many cases avoidable problems need to be solved. In fact most causes of waste mentioned in the lean manufacturing philosophy must be addressed and reduced to a minimum to improve delivery performance and reduce uncertainty over process leadtime [Womack 1995].

But even if these factors are addressed and causes of waste are reduced to a minimum, the structure of operations may still cause uncertainty over leadtime and delivery performance. As long as there is only one server process, different order streams will be competing for the same resource. If a server process has not much slack then this type of competition will induce uncertainty over process leadtime and worse, no control over the leadtime of the individual order streams. Some streams will loose customers and margin as a result of poor delivery performance or excessive capacity will be needed to implement a leadtime policy that was based on small but very frequent orders. Especially when processing times per job are very much apart, then it makes sense to consider separate queue server systems. Similar arguments have been put forward by van Dijk when proposing different queuing strategies for the Dutch post offices [Dijk 1996]

To study this effect, a simulation of a queue-server model is being built and analysed to provide further support for especially the first structure rule.

1. The job-size is the number of products per order.

This limited study is considered sufficient for the following reasons:

1. *For the design of business processes, aggregations and black box models are often used. For the planning and design of logistic systems, the single cell queue server is a realisation of a black box model that can be (re)used at any aggregation.*
2. *Design decisions on how to deal with complexity are implemented at the level of organs for which the single cell queue server is a sufficient representation.*

3.7.1 Queuing theory and (maximum) leadtime prediction

Queuing theory

In appendix C, a more extensive introduction to queuing theory is given. The most interesting result from queuing theory that is relevant for the simulation experiments is the comparison between the performance of the M/M/2,μ servers and the M/M/1,2μ server[1]. Theory predicts that a single server solution with a high capacity is to be preferred over multiple servers with smaller capacities (figure 3.7). However this performance comparison was based on the performance of the mean values of queue occupancy or leadtime.

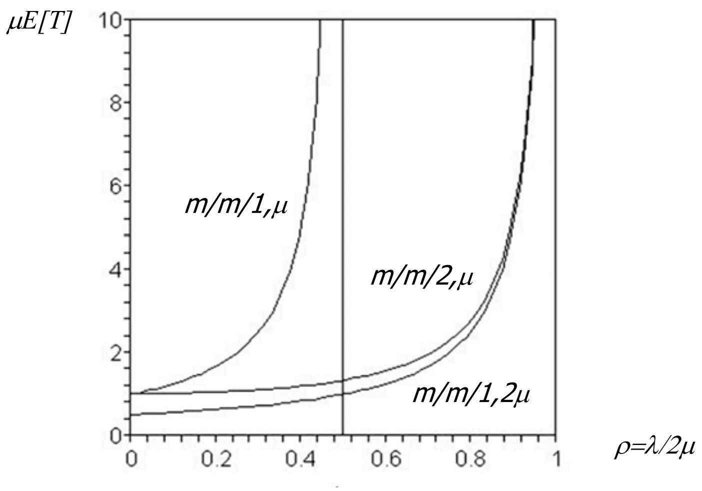


Figure 3.7: Normalize time delays compared M/M/1 and M/M/2

1. M/M/2/μ represents a queue-server system with a Markov-arrival process, a Markov-server process, 2 servers with capacity μ.

The theory also predicts that regardless of the distribution, at high load-ratios ($\rho \to 1$), the behaviour of leadtime as well as work in progress is largely determined by a factor $1/(1-\rho)$. In practice this means that beyond a load-factor of 80-85%, control over the average leadtime or average work in progress is almost impossible.

Little's law

At macro level queue server performance can be described with Little's law that relates *throughput* (average output rate of a process r), *leadtime* (T) and work in process (WIP) (3.11)[1]. Note that leadtime is the sum of waiting and processing times. The *capacity* of a process or machine is the maximum value that the throughput may take if the system is fully loaded. Thus *throughput* is the expectation of the *capacity* under stochastic influences such as Poisson arrival rates, blocking and queue empty probabilities [Hopp 2001]. By definition Little's law is based on expectations of the mean values for these performance parameters.

$$\begin{aligned} WIP &= r \times T \\ E[WIP] &= E[capacity] \times E[leadtime] \end{aligned} \qquad \textbf{(3.11)}$$

For manufacturing lines, the achievable throughput is determined by the bottleneck rate r_b. The minimum leadtime is the raw process time T_0, which for a line is the sum of the long-term average process times. Note that in practice, taking an average means careful consideration about the period of observations. If infrequent variations are not included, T_0 will be too optimistic. For any manufacturing system a critical WIP level can be defined: $WIP_0 = r_b T_0$. At this level, a line with no variability in process times, achieves maximum throughput (r_b) with a minimum

1. For this explanation notations of Hopp are used. In the remainder of the thesis the following more common formulation of Little's law is used: $E[n] = \lambda E[T]$ where λ is the rate or average capacity.

leadtime T_0. Beyond this level, the bottleneck rate will determine throughput and waiting times will be added to T_0 and will increase T.

Variability laws

If we consider maximum bottleneck capacity utilization and maximum inventory efficiency as the optimum, then it is clear that variability will always degrade the performance of these systems. Variability in either r_b or T_0 causes WIP_0 to vary, which will lead to starvation of the bottleneck resource or to degrading the inventory efficiency. In practice this leads to the following law of variability buffering [Hopp 2001, p.295]: *Variability in a production system will be buffered by some combination of: inventory, capacity and time.*

The use of the term variability here has the same interpretation as the stochastic term variance. Variance is linked to probability and uncertainty in the sense that a large variance can be linked to a large spread in the probability distribution of the underlying variable. As a consequence, the uncertainty of a particular value of that variable is also high. Thus variability is related to complexity and the variability law also applies to complexity: inventory can hide complexity, capacity and time can resolve complexity. However, the exact mixture of capacity, inventory and leadtime to cope with input variance is not always predictable.

The variance of probability distributions can be directly related to complexity through the following formula's [Lubbe 1997, Blommestijn 2002].[1]

$$\begin{aligned}
H(X_{\text{normal}}) &= \log\sqrt{2\pi e\, var(X)} = \frac{1}{2}\log(var(x)) + \frac{1}{2}\log(2\pi e) \\
H(X_{\text{exponential}}) &= \log e\sqrt{var(X)} = \frac{1}{2}\log(var(x)) + \log(e) \qquad \textbf{(3.12)} \\
H(X_{\text{uniform}}) &= \log\sqrt{12 var(X)} = \frac{1}{2}\log(var(x)) + \frac{1}{2}\log(12)
\end{aligned}$$

1. x is the stochastic variable whose values have a normal distribution (X_{normal}).

Regardless of the type of distribution, complexity is proportional to the logarithm of the variance plus an offset that is determined by the type of distribution. If we can assume that the shape of the distribution will not be influenced by the queue server, then managing variance will be equivalent to managing complexity. In practice, with large numbers of orders, order size distributions will behave as normal distributions. Although (3.12) was developed for a continuous case, it can also be used for discrete cases with large numbers of orders.

Maximum leadtime

Optimal performance of the mean values of leadtime and work in progress corresponds to optimizing efficiency. However, the market usually differentiates between fast or reliable delivery and lowest price. In the simulation experiments the prime interest will be how much capacity is needed to guarantee a maximum leadtime with a service-level of say 99.9%. This means, the probability that all leadtime values T are smaller then a value T_{max}, is larger than 0.999.

$$Pr[T < T_{max}] > 0.999 \qquad \textbf{(3.13)}$$

Note that leadtime is the sum of waiting time and service time. To solve this problem we need to know the cumulative density function of the leadtime distribution.

If we assume the leadtime distribution is exponential, then we can solve this problem as follows. The cumulative density function for the exponential distribution with capacity or rate parameter μ is: $Pr(T \le T_{max}) = 1 - e^{-\mu T_{max}}$.

If 1-α is the service-level we want to guarantee then the capacity should support the following relation between the capacity μ and the maximum leadtime T_{max}:

$$\mu > -\frac{\ln\alpha}{T_{max}} \qquad \textbf{(3.14)}$$

So α is the imperfection in the service-level that we can allow. Now if we let α take values between 0.1% and 10%, then we see the influence of a high service-level specification on $\mu * T_{max}$. For a given service-level, lowering T_{max} may be achieved through a proportional increase in capacity μ, as long as the product of μ and T_{max} is larger than a minimum value for that service-level. Demanding, a high service-level results in a more than proportional raise of the minimum requirement for the product of T_{max} and μ. Thus we will experience a more than proportional increase of T_{max} if the capacity is not raised.

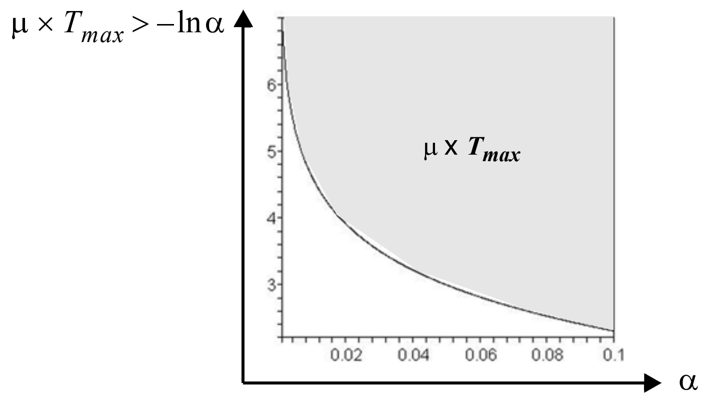


Figure 3.8: Service-level versus capacity demand.

If the leadtime T has a normal distribution with parameters $\bar{T}$ and σ_T. Then T_{max} can be found as follows:

$$T_{max} = \bar{T} + Z(1 - \alpha) \cdot \sigma_T \qquad \textbf{(3.15)}$$

where $1\text{-}\alpha$ is the required service level and $Z(1\text{-}\alpha)$ is taken from table 3.4 [Kreyszig 1970, p444].

Yet formula 3.15 does not relate capacity to T_{max}. For normal jobsize distributions and fixed capacity, the service-time will have a normal distribution as well with parameters $\bar{T}$ and σ_T. $\bar{T}$ as well as σ_T will be proportional to the jobsize and inversely proportional to the capacity. From table 3.4 it can be seen the product of T_{max} and the capacity show a similar characteristic as that of figure 3.8. A service-level beyond 97%, causes a more than proportional increase in T_{max}, which requires an equally large increase in capacity. For finding the maximum leadtime, we need to add the maximum queue waiting time to this result. Regarding the waiting time we only have expectations for the mean waiting times.

1-α	Z(1-α)
0.90	1.645
0.95	1.960
0.97	2.170
0.99	2.576
0.991	2.612
0.995	2.807
0.997	2.968
0.998	3.090
0.999	3.291

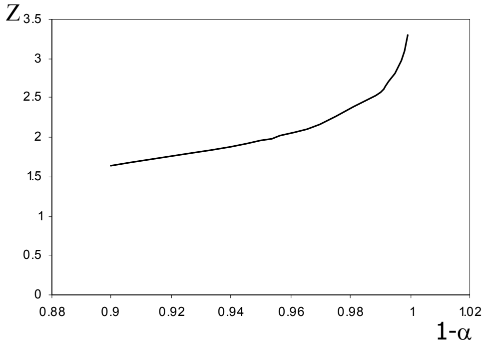


Table 3.4: Z-values for cumulative normal distribution [-z,z].

Prediction of maximum leadtime

Formulas (3.13), (3.14) and (3.15) indicate that we need to know the cumulative density function of the leadtime distribution to predict the maximum leadtime for a given server-capacity with a given reliability (service- level). For composite load-patterns a parametric model of the probability density function may not exist. A simulation study producing a leadtime histogram is needed to find a relation between the maximum leadtime, service-level and capacity.

3.7.2 The single server queue model

The smallest unit under the influence of structure design decisions is a (manufacturing) cell. This cell provides the capabilities, capacity and controls to execute a single task. The cell is a simplified version of the steady state model of In't Veld. The cell has an input buffer to receive jobs and a fixed, deterministic capacity C to process orders from this input buffer. When the orders are finished, immediate acceptance by the next cell or buffer is assumed (figure 3.9). Thus the functionality of the input zone is limited to the buffer, there is no output zone or intervention zone (see also figure 2.5).

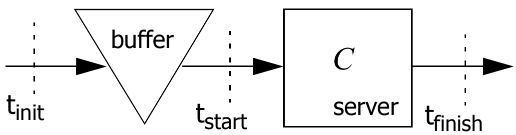


Figure 3.9: A single cell queue server model

To study the effects of concurrency of orders and jobs in a single simulation process, a discrete time simulation was built. The advantage is that in discrete time systems tracing detailed behaviour is easy and problems caused by concurrency surface sooner because there is no debate about the precision of time and the definition of being concurrent. The precision of time and the boundaries between time slots are fixed for all processes and events. However there are disadvantages as well. Continuous time statistical properties may be distorted when mapped onto a discrete time scale. Discrete time or moment based models require more code to describe, than continuous time, period based processes [Veeke 2003].

Scheduling

In the simulation experiments we want to study the performance of a cell under single and multiple order stream loads. Multiple order streams, means the cell is serving jobs from different markets. Because our main interest is to study the effect of structure decisions (e.g. use one or more cells), it is assumed that the cell does (almost) no scheduling but takes the orders from the buffer on a first in first out basis. Only where jobs have arrived simultaneously in the buffer, the largest jobs are taken first to achieve the best efficiency for a given set of leadtime requirements[1]. Jobs that are postponed[2] receive priority in the next time slot. Regarding the performance of the cell, the main interest is to find how capacity is related to leadtime and work in progress under various load conditions.

Jobs

The cell is supposed to process jobs that consist of a number of products (job-size). The proportionality of the job-size to the processing capacity determines how much time is needed to process this job. Thus for the interpretation of the experiments there is no difference between a job containing multiple products or a large job taking several units of time to complete. Jobs are considered inseparable[3] and uninterruptible. Since all products are equal, the influence of set-up or change-over times is not included in these experiments[4].

---

1. Starting with larger orders will reduce their effective leadtime, which reduces their capacity demand for finishing all of these orders within a specific maximum leadtime. Since the larger orders dominate the need for capacity, this rule can be expected to improve the efficiency.
2. Their start is delayed until the next time-slot.
3. Jobs that are not finished are continued in the next time-slot; no other job will be started before the ongoing job is finished.
4. The job administration record, caters for different products. Thus the job-server can be modified to deal with setup- and changeover-times as well as product-related processing times; however as this would complicate the interpretation of leadtime comparisons related to different markets, this effect was kept out of these experiments.

The jobs are requested by customers that belong to a particular market[1] (market number). When an order is created, a job administration record is created with 9 fields (table 3.5). With this job administration record, it is possible to track individual orders and record waiting time before the order is processed (t_{start}-t_{init}), the total leadtime (t_{finish}-t_{init}) and the process leadtime (t_{finish}-t_{start}). The job administration record also has a provision for different products (product code), but for the moment it is assumed that all products are identical.

1	order number
2	market number
3	product code
4	job-size
5	back-order size
6	job t_{init}
7	job t_{due}
8	job t_{start}
9	job t_{finish}

Table 3.5: job administration record

Capacity

The cell takes jobs. Jobs consists of a number of products also named the jobsize. The cell processes these jobs with a capacity of C products per unit of time. If the job queue is empty and if all jobs in the server process are finished, then the

1. Markets are sets of customers and competitors that engage in transactions under homogeneous logistic (and value) requirements (paragraph 2.5.1).

remaining capacity is lost. Jobs cannot be pulled forward in time ($t_{start} \geq t_{init}$). The capacity is fixed for one simulation run and contains no stochastic influences. It is assumed that all stochastic influences are accounted for in the job arrival process and in the generation of a job-size distribution. The reason for this limitation is that the main interest at this moment is the influence of interfering job streams on order leadtime. At the level of *leadtime*, the influence of capacity fluctuations would be indistinguishable from the influence of order size variance or the influence of job stream "interference".

The service-time distribution for a queue with a fixed and deterministic capacity is now determined by a linear transformation of the stochastic job-size variable. The following relations for the service time distribution apply[1]:

$$\mu = C/E[Jobsize]$$
$$\sigma^2 = Var(Jobsize)/C^2 \qquad \textbf{(3.16)}$$

These relations may be used as parameters in the Pollaczek-Khinchine formula's for calculating mean leadtime and mean WIP for M/G/1 queue-server systems (formula C.11 and formula C.12).

3.7.3 General experiment set up

The main goal for the simulation experiments is to compare the performance of a single queue-server system with a multiple queue-server system under the load of orders that originated from markets with different characteristics. A multiple queue-server system is not the same as a single queue/multiple server system as in figure 3.6. But for demonstrating the potential of diversified organs, multiple queue-server systems are more desirable, because preselection or sorting of the queue has

1. To compare with the exponential service time distribution, μ is a service rate or capacity parameter, not a time value; σ^2 is the variance of the service time.

taken place which is difficult to implement or model in the single queue/multiple server model.

In the description of the experiments and their results, care has been taken to be as explicit as possible about the correct interpretation of mean and variance values, because there are two interpretations possible. There is a load interpretation which refers to the time average scores of mean and variance. There is also the series interpretation, which refers the mean value and variance of the job-series without giving notice to the job-initialization times of each order.

If under 'equal' performance measures Mo, the required total capacity $\sum C_i$ for the multiple queue-server system is less than the total capacity C of the single queue-server system then the multiple queue-server system is preferable over the single queue-server system. Experiments with that result provide support for the first structuring rule.

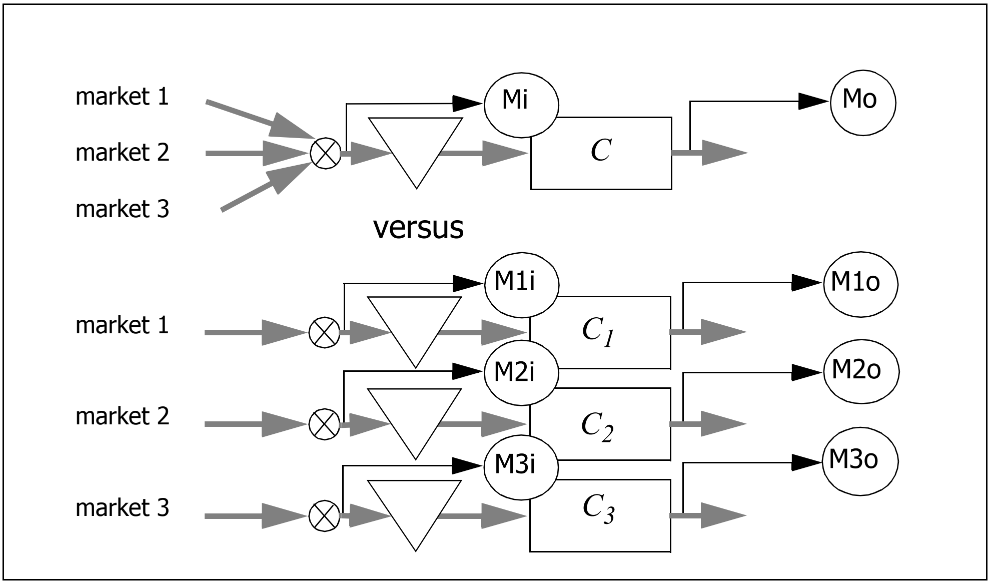


Figure 3.10: Queue-server simulation experiments

3.7.4 Order stream generation

To study the influence of order stream "interference" we need to generate series of customer orders that are different in terms of order size and order frequency. However the average demand is kept the same for all series of customer orders. A total of 6 order series were generated. For each series the following relation (3.17) between the job-size (*Jsize*) and the time between jobs $1/\lambda$ was maintained. For

$$(Jsize_i) \times \lambda_i \approx 10 \quad \forall i \in \{1, 2, \ldots, 6\} \tag{3.17}$$

the timing of the orders a Poisson process was chosen. The distribution of the delay time between orders is exponential. With (3.17) the following targets for λ_i (mtbj[1]) were chosen:

$$\lambda_i^{-1} \in \{1, 2, 4, 8, 16, 32\} \tag{3.18}$$

This leads to mean *Jsize* targets of:

$$\overline{Jsize_i} \in \{10, 20, 40, 80, 160, 320\} \tag{3.19}$$

The simulation period was set at 3200. A simulation period of 3200 time steps was chosen to have a significant amount of orders (100) in the stream with the largest job-sizes, while keeping the calculation times on a laptop-computer sufficiently low to allow some experimentation with the software[2].

1. Mean time between jobs.
2. DELL D800 laptop at 1.6Ghz Pentium-M, 1.25Gbyte DRAM. Maple 9.5 on Windows-XP-Prof-SP2. A typical single capacity run over 6 streams, 6300 jobs, 3200 time-steps, including analysis and display-graphics takes approximately 5 seconds.

Integer time series

To achieve integer values for the time series, the following procedure was implemented:

1	generate a series of exponentially distributed time-between-events values	tlist1[n], $n=1...n_{max}$
2	convert tlist1 into a series of timed events: tlist2[1]:=tlist1[1]; tlist2[m]:=tlist2[m-1]+tlist1[m]	tlist2[m] $m=2...n_{max}$
3	round all event times to integer time values: tlist3[n]:=round(tlist2[n]).	tlist3[n], $n=1...n_{max}$

Table 3.6: Three step approach to generate integer time series

Implementing the procedure of table 3.6 requires special attention at the beginning and end of the simulation period. All time-events that were rounded to zero, were stacked at the starting time of the series: 1. When a job-series has been processed in simulation by the queue-server, there may be jobs that have a finishing time beyond the maximum time initially considered for the simulation run. For correct estimation of the mean output and output variance values, the time scales are extended after simulation to the finishing time of the last job.

In addition to tlist3[n] that contains the time-value for each job-index, a list named TLIST[t] was generated that contains the number of jobs for each time-value t. Thus TLIST[t]=0, means there are no jobs for this time t; a value TLIST[t]=x means there are x jobs that were initiated at time t. Multiple jobs for one value of t, occur in single streams because of rounding time values to the nearest integer or as a result of merging multiple job-streams.

The characteristics of the generated time series are listed in table 3.7.

STREAM-ID	1	2	3	4	5	6
mean time between events	1.014	2.017	4.128	7.697	15.45	32.96
variance of time between events	1.148	4.150	16.75	61.66	252.1	855.8

Table 3.7: Time series properties

Considering the target values for mean time between jobs and jobsize (3.17), the mean time between event values for stream 4 and 5 are under target, which corresponds to a load above the target of 10. Except for stream 1 and 6, the variance values are within 6% of their theoretical expectation.

Job-size distribution

Job-sizes were generated with a normal distribution. The mean job-size is given by (3.19) and a standard deviation σ given by $\sigma_i = \overline{Jsize_i}/6$. With these targets it is "guaranteed" (99.7% confidence level) that all jobs $Jsize_{ij}$ satisfy the following condition $(\overline{Jsize_i}/2) \le Jsize_{ij} \le (3\overline{Jsize_i}/2)$.

If we merge[1] multiple job-series into a single stream then we get the properties of the streams with ID's 12, 123, 1234, 12345 and 123456 respectively[2]. In table 3.8, a distinction is made between job-data and load-data. The job-data refer to the job-size series as they are generated without being mapped onto a time line. The load-data refer to the size-time-series after the jobs have been mapped onto a time-line. For the composite series, only load-data-analysis is available, because the jobs, each individually marked with a order-number and a market-number, have not been

1. The merging is done through putting all jobs on a single timeline. The jobs maintain their individual identity. There may be more jobs per time-step.
2. The full identification of a job-Series with a Normal jobsize distribution is: SNi*, where i represents the series id and * indicates that in case of multiple digits, this is a merged series; e.g. SN1 is a single stream, SN123 is a merged stream SN1+SN2+SN3.

changed and can be sorted after processing to identify interference influences. In table 3.8 the most important properties of these job series are listed.

	NORMAL						NORMAL				
stream-id	1	2	3	4	5	6	12	123	1234	12345	123456
min-job	5	10	18	40	79	206	n.a.	n.a.	n.a.	n.a.	n.a.
min-load	5	11	18	40	79	206	5	5	5	5	5
max-job	17	32	59	119	222	446	n.a.	n.a.	n.a.	n.a.	n.a.
max-load	65	91	148	211	379	506	111	150	272	416	526
mean-job	10.02	19.91	39.62	79.80	159.9	311.2	n.a	n.a	n.a	n.a.	n.a.
mean-load	9.892	9.861	9.605	10.37	10.36	9.515	19.71	28.98	38.65	48.34	57.78
job-size-variance	2.721	11.63	44.03	197.1	663.9	2608.	n.a.	n.a.	n.a.	n.a.	n.a.
load-size-variance	100.6	198.8	392.3	868.1	1650	2989	298.5	682.3	1476	3072	5850

Table 3.8: Job series properties

The mean time between order values for series-id's 4 and 5 (table 3.7) are a little under their targeted values, 8 respectively 16. As a result, these time lines are shorter than the targeted period and the average load figures that result from that are a little higher than 10. This is also demonstrated in the graphics of the single order streams (figure 3.11). The time lines of job-series 4 and 5 do not reach a maximum value of 3200, whereas the time lines of job-series 3 and 6 go well beyond 3200.

The interference between different order-series is visible from the maximum load figures of composite streams in table 3.8. The maximum values for the composite streams are consistently higher than the maximum load for the single streams contributing to the composite stream. This means there are moments in time when multiple jobs from both streams coincide and cause bigger composite loads.

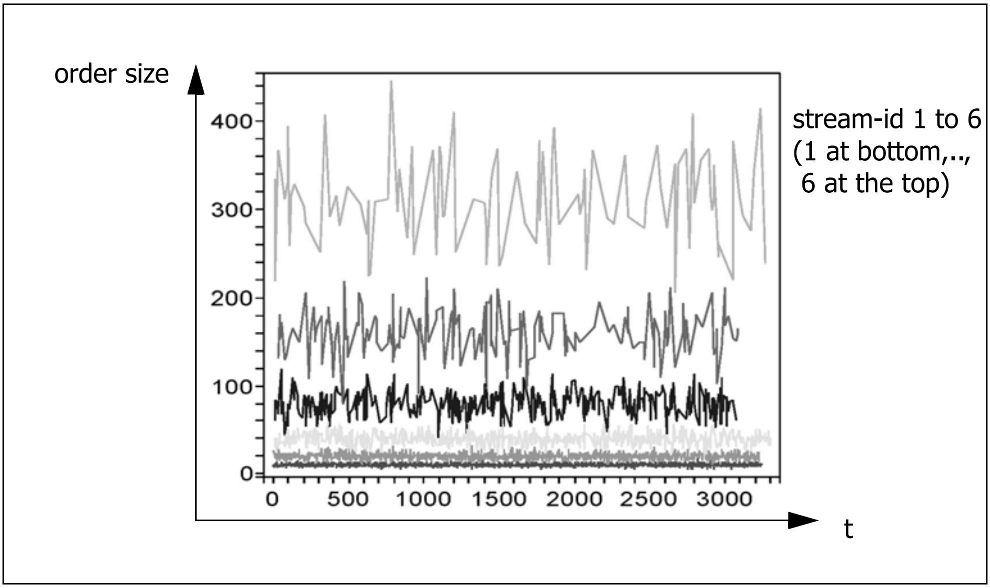


Figure 3.11: Order series, non-composite streams

3.7.5 Job-size variance and the variability laws

To demonstrate that capacity alone is not sufficient to absorb the job-size variance present in the input stream, the following experiment was done. The composite order series SN123456 was loaded onto a single server. The average capacity demand for this series is 59.63. The server was tested with a series of different capacity values. For each run, the sample variance of the job-size time-series was calculated for the input as well as for the output of the server. The difference (output variance minus input variance) was recorded. Basically, this difference describes the development of the output variance offset by the constant input variance.

The idea to study variance was inspired by formula 3.12. Variance drives complexity and in this experiment, job-size variance of composite streams, is expected to cause additional uncertainty over job completion times. The minimum capacity necessary yield equal input and output job-size variance, is also a kind of stability condition for the queue-server system. Another reason for doing this experiment was to study the variability laws (paragraph 3.7.1) to see if it is at all possible to find a condition for capacity as the main tool to absorb variance.

Results

In many experiments, the minimum capacity required to achieve almost equal variance values, was the first integer above the average load-size. But this result was not consistent. Some combinations of streams required a much larger capacity than expected to have equal input- and output variance scores. To have a better understanding of how the variance difference develops, a capacity series from 1 to 180 was tried. The result of the experiment can be seen in figure 3.12.

At a very low capacity, the variance difference is completely determined by the input variance. The output variance is close to zero. Most of the time no orders are completed and the ones that are completed are spread over a very large period of time. The buffer before processing and leadtime variance are absorbing the input variance.

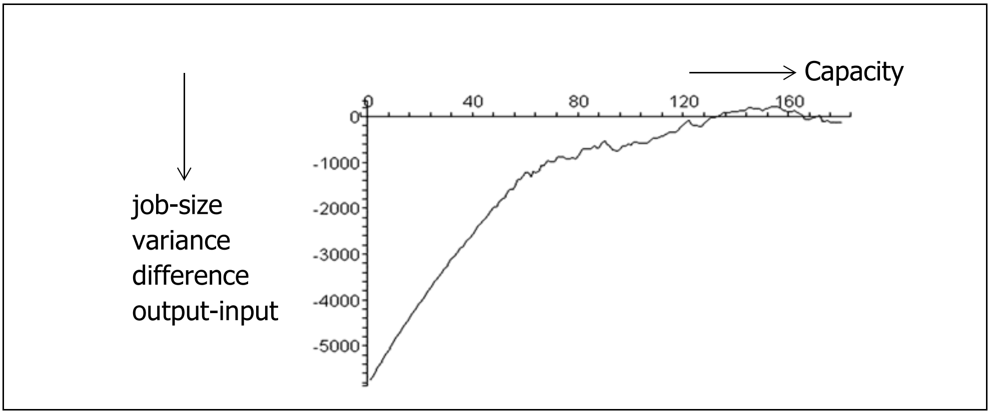


Figure 3.12: Order series 123456 through a capacity series 1-180.

As capacity increases up to 60, this situation gradually changes. For the smaller orders, an increasing proportion is processed without delay, only the larger orders still take long leadtime due to a lack of capacity. The output variance grows proportional to capacity. Still, the input buffer and leadtime variance are the main factors in absorbing the input variance.

Beyond 60, capacity becomes the main factor in taking care of the input variance. Compared to the mean load size, there is extra capacity available, which can be used to recover from job-size peaks that cause delays. Delays no longer accumulate into ever growing leadtime. The number of jobs in the buffer as well as the leadtime variance will become small.
At the same time figure 3.12 shows that the variance difference is no longer a positive definite function of the capacity. The behaviour of the queue-server system differs from the system with a capacity under 60 in the sense that the effective capacity becomes state dependent. The extra capacity not only helps to recover from peak loads, it also causes the queue to be empty some times, in which case the remaining capacity in that time slot is lost.
At a very high capacity, the waiting buffer is always empty. All incoming orders are processed and completed the same time slot. Output variance is expected to be equal to the input variance.

Although the experiment results demonstrate that at very high capacity, output- and input-variance will converge to one another, the state dependent queue-server behaviour above C=60 also demonstrates that the output variance exhibits a behaviour which makes it less useful in parameter estimations over a single run.

3.7.6 Maximum leadtime experiments

From a value creation perspective managing the leadtime performance is a sensible thing to do. Reliability as well as a short leadtime are often rewarded by customers. A predictable leadtime is essential for customers to plan their inventory, given that leadtime. A short leadtime allows customers to minimize their inventory, while maintaining agility. A short leadtime also reduces the manufacturers internal need for inventory. However capacity utilization may be low. In manufacturing settings where material-,inventory- and labour-cost are dominant over the capital cost of machines, low capacity utilization is not a problem.[1]

1. Most theories promoting high machine utilization as the main profit driver for the manufacturing industries, relate to cheap labour and low material cost.

As argued in paragraph 3.7.1, the relation between capacity, maximum leadtime and service-level is a highly sensitive one and for a general case simulation is required. For SN1 as well as for SN123456, the maximum leadtime was recorded over a range of capacity values. Results are presented in figure 3.13. In these results, because of limitations[1] in the simulation software the service-level is implicit and thus very high. All jobs of the series are completed within a leadtime indicated by T_{max}. Both graphs resemble the normalized behaviour of M/M/1 queue-server systems (proportional to $\rho/(1-\rho)$) for the expectation of the leadtime. But the T_{max} values do not match sufficiently, to claim that these graphs are just scaled versions of the expectations for the mean. The Pollackzek-Khinchine predictions for the mean lead-time, seem

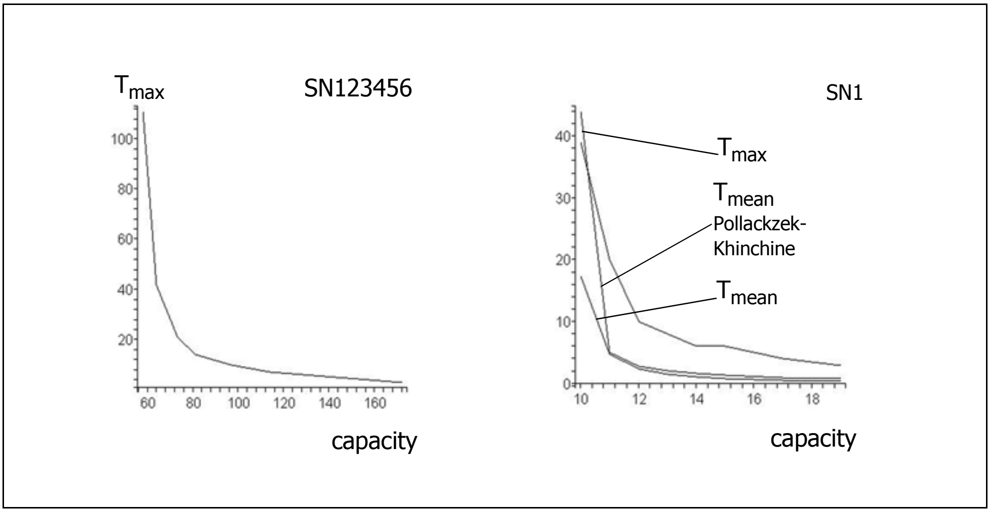


Figure 3.13: Capacity, maximum- and mean leadtime

biased towards slightly higher leadtime scores for loads of 90% or lower. Above 90%, the prediction is very sensitive to the load factor. In fact a capacity increase of 2% at $C=10$, will result in an exact match between the simulation results (T_{mean}) and the Pollackzek-Khinchine predictions. The slightly longer leadtime predictions may be due to imperfections of the process of mapping jobs on a series of timed events and the imperfections of the generated size-distributions.

1. There is no automatic histogram integration to determine T_{max} for a given service-level.

Stability

The minimum requirement for a queue-server system to be asymptotically stable is that the server capacity exceeds the average load. For a stable queue-server system we may expect that leadtime does not accumulate into growing delays. The results presented in figure 3.14 demonstrate this. In figure 3.14 the horizontal axis of all

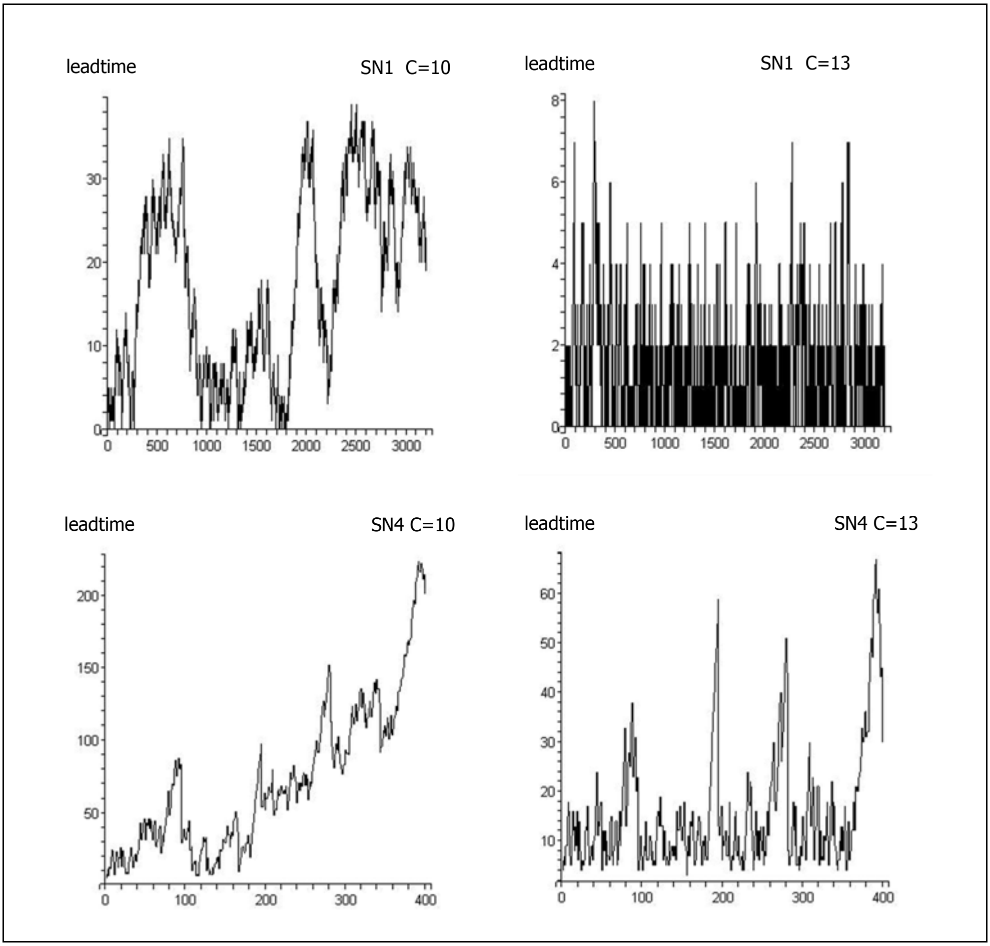


Figure 3.14: Stability and accumulating leadtime

figures refers to the order-index. Because of the relationship between the number of jobs and the average time between orders for each stream (3.17) and (3.18), the job-index location on the scale, refers approximately to the same moments in time across the streams. Thus, the orders with index 3200 from stream 1 and index 400 from stream 4 have been initiated around the same point in time: 3200. As we can conclude from the load-data from table 3.8, a capacity of 10 may be only just sufficient for SN1, but it is not for SN4. For SN4, leadtime does accumulate into growing delays, whereas for SN1 until t~2200, there is enough capacity to catch up in periods where the load is less than the capacity[1]. For comparison, the same run was repeated at a load of ~80% (capacity at ~1.25 times the average job-size). Results show that the peaks in the leadtime graphs are much less and sustained accumulation of delays does no longer occur.

Composite order streams

If a composite stream is loaded onto a single queue/single server system, then there is only one capacity control-variable available to manage leadtime. As a consequence we'll have to choose between either installing a very high capacity to complete all orders regardless of their size in a short time or we are confronted with a situation where only the stream with the largest job-sizes may perform with an acceptable leadtime whereas all others are late as soon as a large order is in process. As to be expected, at the output all streams will be synchronized to the 'slowest' stream. The results of this experiment are shown in figure 3.15. In this experiment composite stream 123 was loaded and processed with a capacity of 29, which corresponds to a load of 100%. After processing, the market-id was used to sort the composite output

1. For SN1 C=10 at t>2200, leadtime also seems to accumulate beyond recovery.

stream and sequence the sorted orders with their order number.

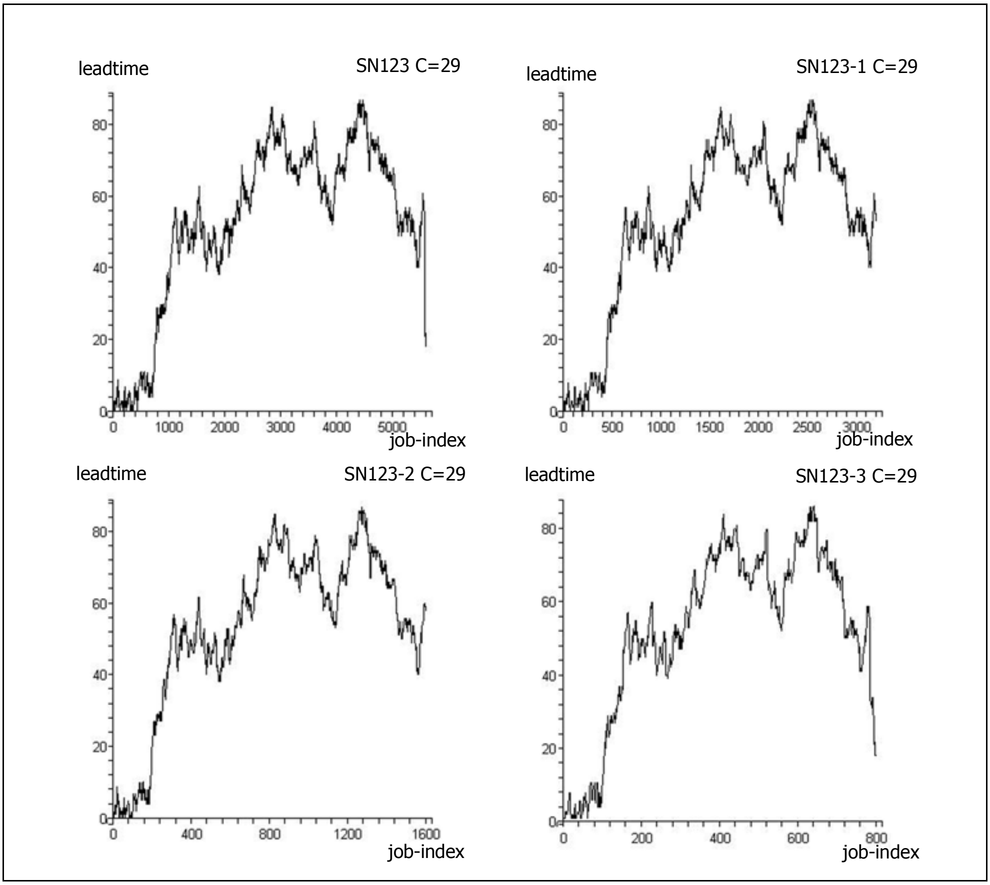


Figure 3.15: Stream 123, composite and sorted after processing.

The average leadtime in all cases was around 53 and the maximum leadtime topped at 87. Processing SN123 with a capacity of 36 (~80% load), makes the maximum leadtime drop to 13 and the average leadtime is now less than 2. This is a significant improvement. Yet a maximum leadtime of 13 will undoubtedly cause some customers to leave. The corresponding histogram demonstrates that over 80% of the orders are processed with a leadtime up to 4. For sub-streams 123-2 and 123-3 a maximum leadtime of 4 may be acceptable (leadtime < 2 * mtbj). For sub-stream 123-1, consisting of 50% of the orders, the histogram shows a leadtime performance that is slightly better than the leadtime performance of the composite case. But compared to a maximum value of 2*mtbj, the histogram shows that 40% of these orders are late. Processing all orders with a maximum leadtime of 2, would require a capacity of 61 and the resulting load is only 47.5%.

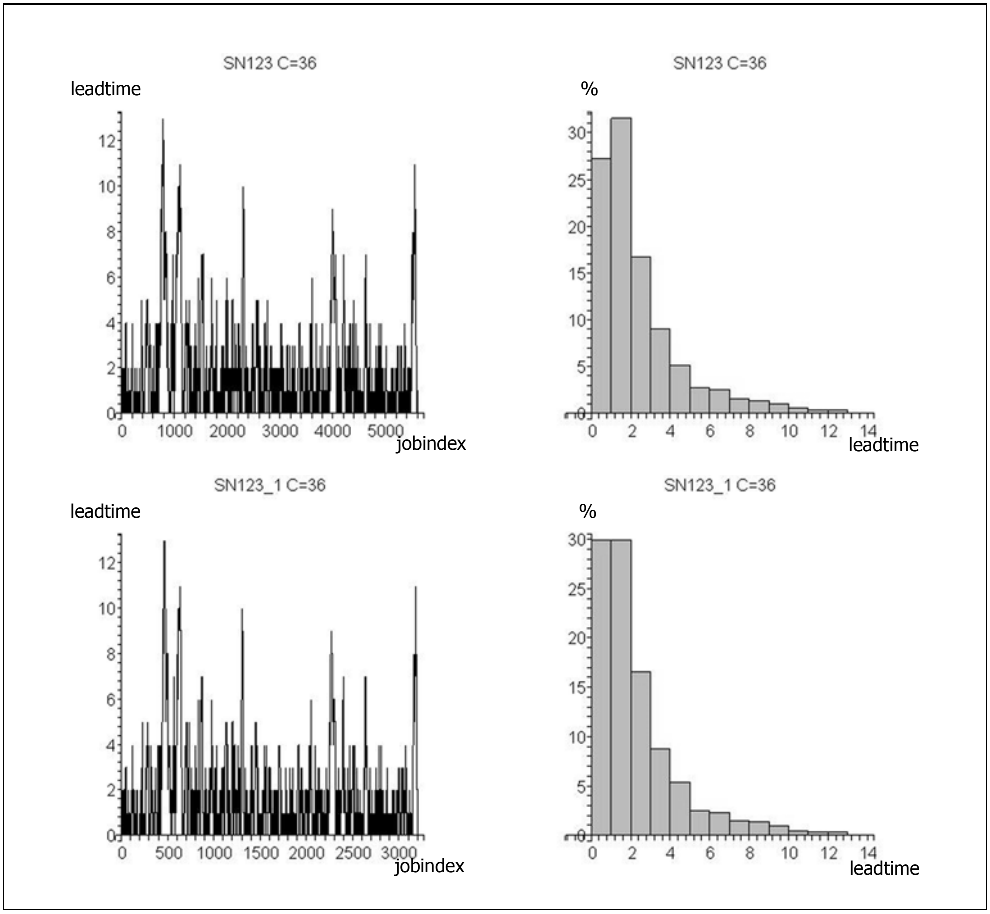


Figure 3.16: Leadtime SN123 at 80% load (C=36).

This example demonstrates that it is not possible to satisfy two different leadtime conditions with only one capacity variable, unless we can reduce the two conditions to one condition for the composite stream. Taking the most severe leadtime constraint as the target value, may accomplish the leadtime requirements but, always results in a low load-factor (low efficiency).

3.7.7 A case for splitting

The argument for splitting presented in the previous section, is a relatively simple one. If the set of requirements is larger than the set of design variables then the system needs to be redesigned with more design variables; in this case separate queue-server systems.

In this section the influence of interference in composite streams handled by a single queue-server is studied and compared with a double or twin queue-server system handling non-composite streams. A special interest is the influence of dynamics. How big is the span of logistic requirements that can be handled by a single server?

To make this comparison, first the characteristics of composite and non-composite streams handled by a single queue-server were recorded for two capacity levels:

- *~80% load*
- *the capacity needed for a maximum leadtime of: ceil[1](2*weighted mean time between orders).*

The last criterion leads to the following target values for the maximum leadtime. The

STREAM-ID	1	2	3	4	5	6	12	123	1234	12345	123456
max lead	3	5	9	16	31	66	3	4	5	6	7

Table 3.9: maximum leadtime targets for single and composite streams.

weighting is chosen to do justice to the dominant stream, such that mean leadtime and the service-level for the composite stream are not too much influenced by a single large job in the middle of many small ones. It is believed that these values are realistic for a market situation. But even if they are not, these target values are useful for getting an idea of what is possible with maximum leadtime control. In case of a maximum leadtime target, the server experiment is repeated until all order-leadtime scores are at most equal to the max lead target. The capacity value to achieve this, is recorded as well as other performance characteristics such as

1. Ceil is a Maple-procedure for rounding upwards to the next integer.

minimum- and mean leadtime scores, load, mean WIP[1] and mean jobsize. In table 3.10 the results for non-composite streams are recorded. Table 3.11 contains the results for composite streams. To test how far we can stretch the span of logistic

STREAM-ID	ρ	Cap	max-LT	mean-LT	mean WIP	mean-input	mean-output
1-80%	0.76	13	8	1.47	14.35	9.892	9.892
1-maxlead	0.52	19	3	0.42	4.09		9.892
2-80%	0.76	13	23	3.72	18.21	9.861	9.840
2-maxlead	0.41	24	5	0.52	2.56		9.861
3-80%	0.74	13	28	6.65	15.46	9.605	9.573
3-maxlead	0.44	22	9	2.09	4.85		9.602
4-80%	0.80	13	67	15.66	21.10	10.37	10.27
4-maxlead	0.41	25	16	3.92	5.28		10.37
5-80%	0.80	13	103	41.19	27.62	10.36	10.28
5-maxlead	0.43	24	31	9.23	6.19		10.32
6-80%	0.73	13	113	42.43	12.25	9.515	9.460
6-maxlead	0.56	17	65	25.87	7.47		9.474

Table 3.10: Results single-server, non-composite stream.

requirements, non-adjacent composite streams (13,14,15 and 16) have been tested as well as adjacent composites of more than two streams (123, 1234, 12345 and 123456).

1. WIP, input and output and job-size are expressed in number of products.

STREAM-ID	ρ	Cap	max-LT	mean-LT	mean WIP	mean-input	mean-output
12-80%	0.76	26	10	1.21	17.76	19.71	19.71
12-maxlead	0.55	36	3	0.31	4.57		19.71
13-80%	0.74	26	12	1.70	20.1	19.33	19.32
13-maxlead	0.51	38	4	0.50	5.96		19.32
14-80%	0.76	26	27	3.42	38.37	19.73	19.73
14-maxlead	0.41	48	4	0.50	5.63		19.73
15-80%	0.76	26	34	6.33	67.12	19.76	19.76
15-maxlead	0.19	102	4	0.16	1.73		19.76
16-80%	0.74	26	41	7.54	73.45	19.32	19.27
16-maxlead	0.17	111	4	0.21	2.07		19.32
123-80%	0.81	36	13	1.85	30.73	28.98	28.97
123-maxlead	0.58	50	4	0.43	7.07		28.98
1234-80%	0.79	49	16	1.95	35.09	38.65	38.65
1234-maxlead	0.64	60	5	0.74	13.33		38.65
12345-80%	0.79	61	17	2.64	49.63	48.34	48.34
12345-maxlead	0.50	96	6	0.43	8.04		48.34
123456-80%	0.79	73	21	3.34	63.19	57.78	57.78
123456-maxlead	0.51	114	7	0.57	10.79		57.78

Table 3.11: Results single-server, composite streams.

Discussion

The 80% load figures, indicate that making composite streams of 1 and 2 and feeding them to a bigger single-server is does not harm too much. SN1 is delayed a little more in SN12, but the leadtime of SN2 benefits considerably from the mix. The mean leadtime of the combination is even better than the mean leadtime of SN1. Repeating the same experiment for SN1 and SN3, makes the leadtime performance much worse for SN1, which now shows an increase of 50% in the maximum leadtime. Mean leadtime performance deteriorates as well for SN1. But suppose this is still acceptable, then it turns out that filling the slack between SN1 and SN3 with the jobs of SN2 can be done without a problem and needs less capacity than the individual streams.

Experiments SN14, SN15 and SN16 demonstrate that at 80% load, the leadtime performance for SN1 has become completely unacceptable. The less frequent, but larger orders clearly act as a disturbance.

Thus we can conclude that at approximately 75% load, streams SN1 and SN3 should already be considered for a multiple-queue/multiple server solution: *At equal loads, a factor 4 in job-sizes and job-frequencies cannot be accommodated if a "tight" leadtime control regime for the more frequent and smaller jobs is required. However if this leadtime performance is still acceptable, then economies of scale are possible. Filling the slack with other jobs, smaller and more frequent then the larger ones is permitted as long as the total load does not exceed 80%.*

Looking at the max leadtime experiments, one can draw similar conclusions, *however the factor between the job-sizes and job-frequencies where leadtime control for the more frequent jobs becomes unacceptable is now raised to 8.* The reason is that the max lead experiments mostly result in load-factors of 50% and less. This gives the slack needed to quickly recover after a large job. Also the maximum leadtime criterion forced to install enough capacity to serve even the largest jobs in small time. Again, if the leadtime constraint on the smaller orders can be relaxed, then filling and deploying a large single-queue server is a good solution as the SN123456-maxlead experiment demonstrates.

Without the filling, under a weighted max leadtime regime, composite streams that differ more than a factor 8 in size and frequency, should always be handled by separate servers.

All experiments demonstrate that for these streams, one cannot achieve a load-factor or efficiency of 75% or more and have a tight maximum leadtime regime at the same time. Thus if high machine efficiencies are needed, buffers and make to stock strategies are the only way to guarantee short delivery times. This means that inventory efficiency will be low, which implies that such a strategy is only economically feasible in case of low material cost and a low added cost per product[1].

3.8 Structure decisions; the first rule in practice

Primafoon shop vs. Belcompany.

In the Netherlands one can still experience how serious the problem of job-stream "interference" can be with the Primafoon shops. Primafoon is the equipment outlet as well as customer front-office for all products and services KPN-telecom has to offer. At the highest level, Primafoon has two kinds of shops: Primafoon business centre for professional clients and the Primafoon shops for all other customers. This means Primafoon recognizes that business clients have different leadtime requirements for their services than consumers. But this is also where their acknowledgement of the first structure rule stops. In shops selling telephone peripherals as well as being front-office for subscription services the problem is the service time requirements for different types of customer needs. Buying small materials such as pre-paid telephone cards, battery packs, wiring do-it-yourself kits, should be no more than 30 seconds to 1 minute transactions. Subscription services may be in the order of 5-10 minutes. The same service time applies for buying fixed line peripherals. However, buying a cell-phone including making a choice out of 4 different types of subscriptions may require 30 minutes or more. The same services times apply for PC's and PC-network equipment. At Primafoon everybody lines up in the same queue. Thus despite multiple servers (shop-assistants), the waiting and service time for every customer is about 30-45 minutes, regardless of their need. Fortunately there is competition. Belcompany demonstrates that a different organisation or task-allocation for the shop-attendants produces better maximum waiting and service time results. First of all their cash-register is manned all the time by 1 shop-attendant who is also responsible for selling small goods in the 30 seconds

1. Added cost rather than added value; bearing these cost requires financing and cash decisions. added value is a 'virtual' appreciation that only becomes cash after sales.

to 1 minute category. For all other needs there are 2-3 sales-advisors that take care of their customers until the moment they have made their choice. After the customer's choice, goods and invoices are handed over to the staff at the cash-register to finish the transaction. Front-office services regarding subscriptions are either dealt with by specialized staff or by the product-advisors if the shop is not so busy and specialized staff is not available. At Belcompany the shop attendants are aware that customers need time to think about their options, time to feel the product and get a first impression about its use. Staying with these customers while they think and feel is not only annoying, it is a waste of service time. To give customers their "private" time with the information and the products, it is acceptable to share the service-time and service up to 4 customer's concurrently. The length of the queue is kept small by offering only a limited amount of shop-space to the customers. If the shop is full with people, most potential customers pass and consider coming back another time. At Belcompany most customer's experience a waiting time of less than 5 minutes and even shorter for buying of the shelf commodities.

Aircraft maintenance.

Most airline companies operate a fleet of aircraft in a so-called hub and spoke service network. This means that smaller aircrafts are flying along the spokes to smaller airfields to collect and bring travellers to and from the hub where larger aircrafts are available for hub-to-hub long distance transport. The hub-feeding schedule usually means that most of the smaller aircraft have a daily start and end of their schedule at the hub. This is perfect for in line maintenance (so called weekly inspections, A- and B-checks) that can be scheduled and executed overnight and in the weekends without disturbance to the operations schedule. Line maintenance operations and weekly inspections are planned and executed in close coordination with the fleet planning department of airline operations. Due to unexpected problems, flexibility in postponing and forwarding certain inspections is also needed and airline companies do not want to compete for capacity with other airliners that may result in their aircraft not being available for operations. This is exactly why Easyjet together with

FLS Aerospace has set up Easytech.
However, not all airline companies reason the same way. Despite their responsibilities according to EASA or FAA ruling[1], some are reconsidering their core-competencies and aircraft maintenance is often considered as an activity to be outsourced to an external specialized party. In terms of organisation design two problems may surface as soon as maintenance is outsourced. Maintenance facilities are expensive. This means we cannot expect a lot of slack[2], unless the maintenance company is well rewarded for maintaining slack. The airline company no longer has a private facility. Competition with other airline companies also outsourcing their maintenance will occur and the largest or most profitable customer will most likely dominate waiting- and service-times. Moreover the tight coordination with fleet planning, which requires mutual access to both planning systems, often ceases to exists. A second problem may be that the service provider can not offer facilities or capacity at the hub or base. This means aircrafts may have to perform extra flights to their maintenance facility. In both cases airline companies will need a larger fleet of aircraft to operate the same schedule. As a general rule, only large maintenance operations that are not scheduled within daily or weekly operations such as C- and D-checks may be outsourced to specialized companies. Line maintenance activities planning are the responsibility of fleet planning. This planning includes prioritizing the available maintenance capacity and facilities.

Universities

Universities are typically organized in faculties, departments and laboratories. A laboratory is the smallest working unit, specialized around one discipline. But also at department level and faculty level, identities based on traditional disciplines, are the main driving forces behind the choice of organization principles. This is no surprise since the quality assurance and accreditation mechanisms for faculties as well as for

1. EASA (European Aviation Safety Agency) and FAA (Federal aviation authorities) both rule that the aircraft operator is the prime responsible party for doing all that is necessary to maintain airworthiness of the aircraft.
2. In this case keeping your own maintenance facility is a more profitable option.

the journals that play an important role in research assessments, are based on peer review of traditional mono disciplines.
At the same time, society and industry express their need for multi-disciplinary academic education. Economy of scale is believed to bring more efficiency and scale or critical mass is believed necessary to perform large scale research aimed at innovation and breakthrough. Public funding is reduced to force universities to reorganise themselves into bigger units and to force them to offer more opportunities for private parties to influence research agenda's through participating in private funding.
The conflict between traditional accreditation mechanisms and the demand for multi disciplinary education and research is not helping universities to change their organisation. Universities have established institutes and centres to offer multi-disciplinary research capacity. But these institutes are manned with groups that primarily depend on mono-disciplinary activities for their funding and accreditation. Without a strong financial incentive to participate in multi disciplinary projects and without mechanisms for solving resource conflicts between work for the centres and work for the department, these centres are not likely to be successful. As long as the success of participating laboratories does not depend on the success of the centre, their commitment to invest in the centre is limited unless the centre provides good opportunities for creating mono-disciplinary value and preferably even generates free cash.

Big may not be beautiful. From the viewpoint of the coordination cost model and from Glicksteins hierarchical span theory, one may seriously question the efficiency of large faculties and departments. As long as student-, PhD- and staff-colloquia are the only working methods to create a scientific community that shares a vision, tools and skills and that works effectively on scientific problems as well as education (training) of new generations of scientists, then the total number of persons involved should not exceed 50 persons[1].

1. 7 times 7, 2 layer organisation.

Clustering laboratory groups into departments offers only added value if departments can be formed as faculty front-ends for multi-disciplinary activities targeted at certain industry branches or applications. For some faculties, departments could take the role of centres or institutes. For a mechanical engineering faculty, a centre for the automotive industries or for the process industries, would be a well recognized market oriented institutes.
A faculty's prime task is to offer education programs. The research facilities are an essential part of finalizing both undergraduate and graduate academic education programs. At the same time the curriculum should prepare for participating in research programs running at these facilities. However, this does not mean that the research is leading the education program. If research ambitions are leading over undergraduate and graduate course programs then faculties may end up with running as many programs and degrees as there are laboratories and they will loose their identity as faculty. A faculty should be responsible for organizing as few programs as possible. Preferably a single undergraduate curriculum should contain the commonalities to support all laboratories in their ambition to push the frontiers of science as well as in their need to participate in multi-disciplinary centres. If such a common program is not possible, then it is better to have more faculties. If a faculty does not attract enough students to operate the common program efficiently then it should consider a position as centre or department maintaining the research training facility with students from various groups or even faculties.

Privatisation of public infrastructure

The examples above represent only a small subset of many strategic and organisation structure problems encountered in practice. Today many public service organisations are being sold to private investors based on the believe that the market can do a better job in running these companies more efficiently. Examples are: mail, telecom companies, electricity and railway companies. The privatisation of mail and telecom companies can be called successful because when these processes were started competing technologies (cell-phone and internet) and competing infrastructures (optical fibre networks, world wide logistic service companies as UPS, DHL) already existed. For railway companies and electricity companies this is not the case. Alternative infrastructure (power lines or railway lines) is not available and is too expensive to install. The same applies for drinking water and natural gas distribution. Thus competition on the level of the main capital intensive part of the service is unlikely. Splitting these companies into a network part and exploitation companies to allow competition at the level of exploitation will cause problems to guarantee the long term service level and stability of these services. The problem is investments. The network companies loose control over cashflow from the end-user. The cashflow needed for maintenance and investments in the network infrastructure may go to non-value adding activities or to the owners of the exploitation companies. For electricity companies there is an additional problem. The options for storage of generated electricity are very much limited. This means technical infrastructure needs to balance supply and demand over the entire network at all times. If this coordination becomes the responsibility of privatised companies with different strategies and value propositions, then it is likely that service at peak load levels cannot be guaranteed any more. For railway line companies, the problem is that their primary customer value driver is the reliable operation of a network of connections with sufficient capacity and service frequency. The reliability and safety of these operations is a combined responsibility of the network company and the company that owns and operates the train. A split between the network and exploitation for railway lines corresponds to splitting an important control cycle. Such a decision requires new control strategies and technology that allows economic decoupling. To my knowledge and experience such solutions have not been developed or installed yet.

3.9 Conclusions and Discussion

Complexity as argued in chapter 1 is driven by variability and uncertainty. Both of these factors are closely related to the troubles of managers, also often coined by the term complexity.

Quantitatively, variability and uncertainty can be represented by Shannon's information measure or sometimes short: Shannon's complexity. Two arguments support this choice:

1. *Shannon's complexity measure is a relative notion of complexity. It's value depends on the level of variability and uncertainty, which is influenced by the required detail of modelling and the choice of the level of aggregation for which this level of detail is relevant.*
2. *Shannon's complexity is a quantitative figure that represents the amount of information needed to resolve the problems of variability and uncertainty for the chosen level of aggregation.*

The main question for this chapter was, can we use Shannon's complexity as a figure of merit or as a reference for supporting structure design decisions? To answer this question, it was demonstrated that the theoretical framework of Shannon's complexity can be used to built theories for structure design decisions that influence the relative complexity in the desired direction. Especially by looking at the value and symmetry of the conditional information that represents the coordination efforts between sets of issues, possibly handled by different departments, additional support for splitting or merging these sets or for defining constraint relations between these sets could be derived (table 3.1).

If inter department coordination is necessary, then these departments should not be too far apart in terms of the length of the coordination path. A coordination cost model shows that these cost or time grows more than proportional to the length of the coordination path. Assuming that coordination within a department does not need the exchange of criteria and value systems, coordination over two management layers takes over 20 times the unit cost of coordination within a department. Thus this model gives additional support for keeping related controls together (2nd structure design rule).

In addition, Glickstein's optimal span theory was discussed to indicate that there is an optimal group-size for optimal working group efficiency if coordination between members of the working group is required. Glickstein's theory is also based on Shannon's complexity. The optimal span theory shows that the optimum number of coordinated workers is 6.4. This number is also considered as additional support for Miller's well known "magical number 7 plus or minus 2". On a smaller scale, the coordination cost model could also be used for inter person coordination. The unit communication cost factor will be lower than for inter department communication. If we suppose there is an upper limit to what is acceptable for inter person communication cost, then as a consequence working groups or departments will brake down into smaller sub-groups as soon as there is no topology possible to connect all people within the communication distance that corresponds to this upper limit.

To demonstrate that there is a logistic reality behind the organs and complexity thinking, the problem of "interference" between job-streams was studied. The main hypothesis for this study is that markets, characterised by their order frequency, jobsize and a certain level of variability may have different leadtime requirements. If job-streams from different markets are mixed into a single stream handled by a single server, then it becomes impossible to serve these different leadtime regimes while maintaining efficiency. For this experiment a total of 6 job-series were generated each representing a different market, with job-sizes and job frequencies that were approximately a factor 2 apart. Yet all streams were generated under the condition that the average demand for capacity is the same for all streams.
From these experiments it became clear that multiple job-series handled by only one server results in the same leadtime performance for all streams. Whether this is acceptable depends on the adopted capacity regime. If efficiency is the main target and a load of about 80% is the prime goal, then the leadtime performance is dictated by the stream with the largest jobs. All job-streams become synchronized. Depending on the requirements for the most frequent and smallest jobs, up to a factor 4 in frequencies and job-sizes may be acceptable. Beyond 4, more servers are strongly recommended to implement differentiated leadtime regimes.
If a maximum leadtime performance for all jobs is the main target, then the efficiency is usually low. In the experiments, a job frequency weighted maximum leadtime regime was used for target values. Thus these targets are dominated by

the more frequent and smaller orders. Yet the capacity needed to achieve this leadtime regime for all orders is determined by the largest orders in the stream. As a result, the load factor and thus the efficiency is low. But it also means that filling the slack with smaller orders up to approximately 80% load, does not harm leadtime performance very much. Regarding the range of jobsizes and job frequencies that can be accommodated by one server under the max leadtime regime, up to a factor 8 can still lead to acceptable results for the job-series with the highest order frequency.

4. ORGANIZATION IN DESIGN AND DEVELOPMENT

In the preceding chapters, structuring processes as a means to reduce management or control complexity has been discussed in a quasi-static context. For most operational processes as can be found in sales, manufacturing, distribution and service this is not a problem since the context for learning, reuse of knowledge and exploitation of experience is quasi-static. However for breakthrough innovations and high complexity design and engineering projects the structure needs to be more dynamic. Learning and experience are at the heart of such projects, but the amount of reusable knowledge is limited and it is often not clear from the start what knowledge is needed, where and when.
In this chapter a concept named knowledge logistics is introduced and principles of self organization are being discussed as a frame of reference and a source of ideas for new design processes that can deal with more complexity and achieve results in shorter time. Set-based concurrent engineering is discussed as an example of the use of principles of self organization in design. Taking the idea of set-based concurrent engineering one step further, an evolutionary organization for design processes is being proposed.

4.1 Introduction

Most academic institutions still teach structured design to their students [Pahl 1996, Hubka 1996]. Not because it is the best method, guaranteed to lead to good designs, but merely because it addresses all the relevant areas of design processes in a comprehensible way to students unaware of their own early design experiences.

Although Suh may claim so, axiomatic design [Suh 1990, 2001] is not fundamentally different from structured design. The design phases are roughly identical. Axiomatic design is characterized by the design matrices that represent an efficient data-representation to show where design decisions are complicated (coupled) and where they are not. A decoupled (triangular) design matrix provides information on the sequence that needs to be followed to solve the design problem. However, the problem of developing a set of uncoupled or decoupled design matrices for a specific design is as complex as solving the design problem using other methods such as structured design.

Following the principles of structured- or axiomatic design, one can easily see a phased plan, perfectly fit for a work breakdown structure and presumably fit for effective and efficient development processes. Industrial practice shows that relying on a-priori developed work breakdown structures often results into either risk-averse incremental development of a known concept, or into cyclic, hard to finalize development processes in case a new concept was pursued. It just is very hard to predict up-front what relations and which concepts prove to be crucial for a successful design.

As a consequence, the product architectures of cars and aircrafts have not changed significantly over decades of their existence. Despite claims that technology developments are speeding up, the impact of new technology or new materials is often limited to redesign of sub-systems. The problems of introducing new and unknown relations are avoided as much as possible. Also, the opportunities of new business models with new technologies are often kept outside the scope by first implementing new technologies into existing product platforms.

Where new concepts do appear, prototyping and testing often take a more prominent place in the development process. In other words, critical relations between subsystems are being discovered through trial and error. This rather laborious process continues until a working systems-architecture is discovered. Although this approach is not really a self organized process, principles of self organization can serve as a reference for understanding how (sub)system boundaries are settled such that the interaction needs between (sub)systems is being relaxed to a level where the interaction becomes manageable under all operating conditions. But trial and error is not a rapid development process if the trials are done sequentially in a slow pace.

Global competition has increased the required pace for product development and innovation. As an answer to the need for an increased pace for product development and innovation, Dill and Pearson recognize the need for a focus on structure and communication mechanisms that enhance cross-functional and cross-disciplinary knowledge exchange as well as learning [Dill 1991].

Enhanced cross-functional and cross-disciplinary knowledge exchange will improve development processes directly since most development problems involve multiple disciplines. In addition, the opportunities for learning and reuse of knowledge can be enhanced as well. Reuse of mature concepts can save a lot of development time. Being able to migrate concepts back and forth between disciplines makes the source to pull concepts from, much richer. Being able to learn from other disciplines helps to identify problems and critical relations earlier in the development process.

Another element in speeding up development processes is to further increase the level of concurrency. However increasing the concurrency of the development phases as in concurrent engineering is only increasing the need for coordination. This problem has much in common with the problems of structuring and decomposing complexity as described in paragraph 3.2.3. The solution to this problem is the same: define constraints to decouple the phases. Yet the gain that can be expected is limited unless the phases themselves can be executed much faster.

Speeding up the phases requires an increased level of concurrency within the scope set by the development phase definition. In practice this means: work on multiple solutions and concepts concurrently to establish a level of functional redundancy. Two beneficial effects can be expected from this approach. First, it allows to put more capacity on the job as long as the job is not dominated by coordination and communication issues. Second, the diversity and redundancy created through concurrently working on alternative solutions reduces the chance that developed concepts will fail in the coordination between different phases and different parts of the design, even when new unforeseen requirements are added later on to the program.

In this chapter two concepts are presented that help to facilitate reuse and learning and that support the increase of internal concurrency.
The first concept is called knowledge logistics. This concept was inspired by process-structures commonly used in material logistics. Through segmenting the development process into phases, each with a different scope and timing, a

"buffered pipeline" of development processes is created with capacities dedicated to each phase. All of these processes can be executed concurrently creating an inventory of concepts that can be assembled to order to create a complete design. The second concept is called evolutionary organization of design processes. Evolutionary processes effectively implement trial and error on a massive scale. Rather than sequentially testing individual concepts, many concepts are tested and developed concurrently. From studying principles of self organization and evolutionary problem solving processes as genetic algorithms, it can be demonstrated that relatively novel design processes such as "set-based concurrent engineering" have the potential to come up with new concepts and exploit redundancy to reduce the risk of getting stuck in cyclic, non-convergent reasoning processes. Taking this idea one step further, an evolutionary organization for design processes was conceived.

4.2 Knowledge logistics

Dealing with the time aspect seems crucial in bringing engineers and business managers together. For one thing the business processes at product level operate on a different time scale than the technology development processes. Yet both processes are connected. The technology should provide for product ideas that must enter the market within three months and that may be on its decline only nine months from now. However the same technology may have taken three or more years to develop and must support future product ideas for at least another two.

The link between product and technology development runs through scenario's and estimates. The larger the gap between technology development leadtime, the product lifecycle and the product development leadtime, the bigger the risk of developing technology that is obsolete before it can be applied commercially. This situation is not much different from manufacturing processes that because of their physical properties take more time than the customers are prepared waiting for. In such cases, manufacturing to stock with planning based on demand prognosis is the preferred strategy.

From a complexity perspective it is important to realize that within research and development departments, the dynamic range they can cope with is also limited. Knowledge intensive development with hard targets and deadlines every quarter requires different processes than very advanced long term visionary developments with a leadtime of 5-10 years or more. The visionary projects are usually managed through yearly budget cycles.

4.2.1 Production logistics

In production logistics it is common to introduce decoupling points if the manufacturing leadtime is (much) longer than the acceptable order leadtime in the market. To complete the manufacturing process in time, these processes need to start early on customer order prognosis. The manufacturing-order is usually completed before the customer-order arrives. Thus manufacturing planning and control becomes decoupled from sales planning and control. Another reason for introducing a decoupling point is a significant change in market risk. The customer value attributes may be to unpredictable to keep stock of customer ready products. The decoupling point is introduced at the point where the commonalities between products reduce to the level where manufacturing on order becomes the only commercially viable option.

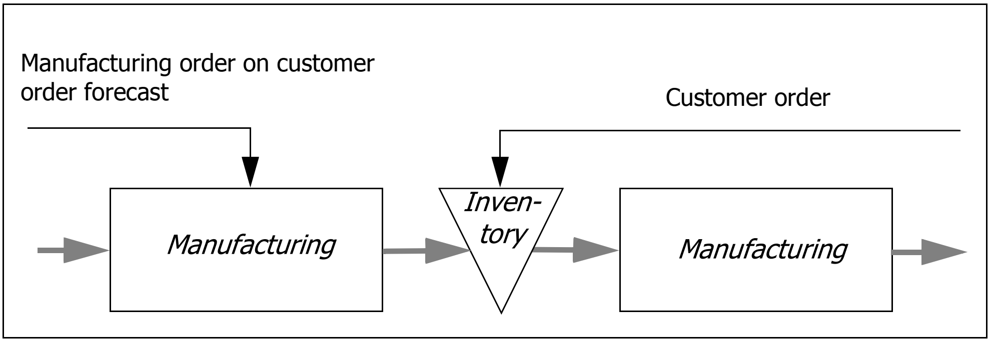


Figure 4.1: Manufacturing decoupling point

In general, a decoupling point allows the use of separate material processing control strategies on either side of the decoupling point. At the decoupling point one usually finds an inventory point that contains the material that is manufactured on prognosis. In practice five generic decoupling points or order-entry-points (OEP), referring to different manufacturing strategies are distinguished: sales from local stock (1), sales from central stock also known as make-to-stock(2), assembly-to-order (3), manufacture-to-order (4) and purchase-to-order (5). The latter can be found with project industries that make one-offs. Note that in a real company one seldom finds a pure OEP 2, 3, 4 or 5. In practice combinations of OEPs may be chosen depending on risk and leadtime values associated with particular parts and the demands of customer orders. Especially when companies serve more markets concurrently, more OEPs will be found in the manufacturing processes.

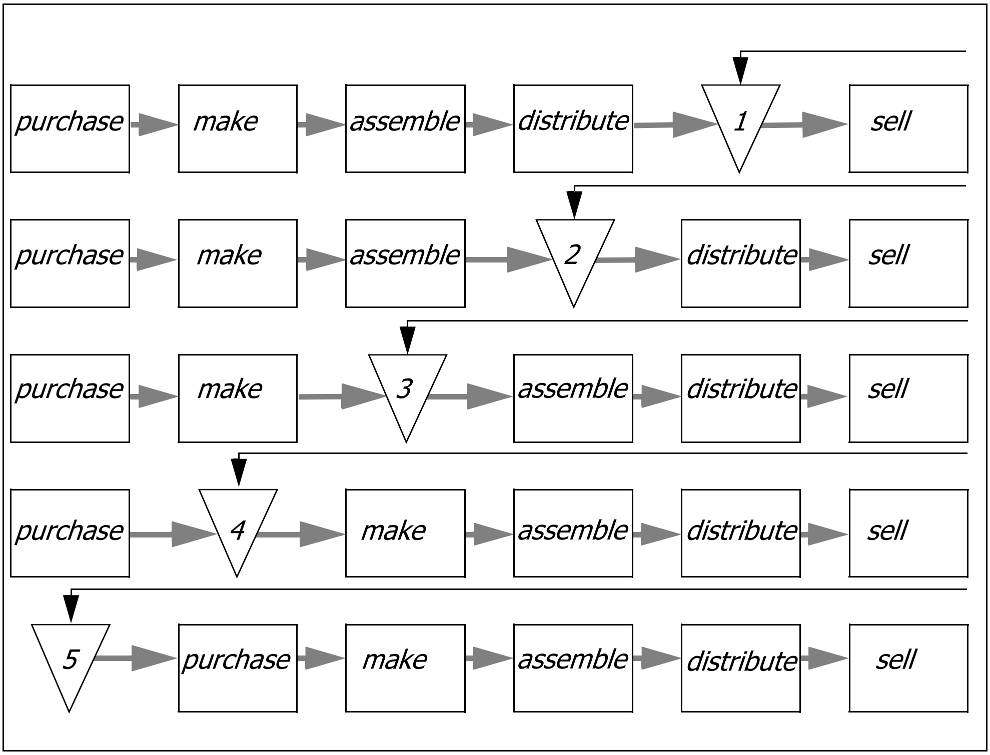


Figure 4.2: Generic decoupling- or customer order entry points

In manufacturing processes stocks can be found wherever steps in dynamics and risk have to be levelled. These stocks serve as the point of reference for both upstream and downstream manufacturing planning processes.

Bikker later extended this model with the so-called order-specification decoupling point (OSDP), where he considered 5 different stages of preparation of information related to the material flow and the material transformations. The order entry point and OSDP were also combined into a matrix, the so-called order-entry matrix [Bikker 1994b, 1998]. Yet the OSDP did not consider strategic planning and innovation as influences to order execution and at the level of operations it considered preparation of material flow control information as one of the OSDP's. The latter was already accounted for in the material flow decoupling points. Later the lack of attention for strategy and innovation was addressed through defining innovation impact points in relation to the layers of the innovation model of In't Veld [Bikker 2000].

4.2.2 Knowledge logistics

To be able to apply the production logistics metaphor to knowledge development processes one has to realize that in production logistics the generic processes purchasing, manufacturing, assembly, distribution and sales are all by nature very different in their dynamics and in the way material is treated. Thus buffers or stocks have their natural place between these processes.

Slicing up a knowledge development process into sub-processes with the intention to create a production logistics metaphor requires that generic knowledge processing steps are identified that possess similar differences in dynamics. In addition, the information items produced in each phase need to contain clear, distinguishable results that can be stored and used later as needed in later phases.

The starting point of a goal oriented product development process is a set of (future) user-need-scenarios. The user need scenario specifies the development problem to be solved and the context in which the product has to function.

These scenarios can be expressed in terms of functions that support those needs. Often there is a range of technologies that in principle can support the realization of these functions (road mapping) and that can meet the context demands. The selection of a particular technology platform usually leads to modules and systems that can be configured to built a complete product. Finally the design is implemented (distributed) through scheduling and dispatching routines that trigger production.

Thus the following generic knowledge processes have been identified:

- *Road mapping is the process that transforms a functional specification of user need scenarios into sets of technologies that could support these functions.*
- *Technology development selects the appropriate technology and develops the technology to a level where both function and cost effective solutions for the customer need are feasible.*
- *Modules & Systems design transforms the technology into a modular product family platform. Modules and (sub)systems have standard interfaces and they may be functionally complementary.*
- *Configuring / knowledge assembly, assembles a complete solution design from modular subsystem designs.*
- *Manufacturing scheduling and dispatching distributes the design to production.*

Each transition between the knowledge processes can be interpreted as a decoupling point where risks and dynamics may change. The customer order role is played by the vision or idea that triggered the development process and the availability of building blocks and the market leadtime determines the entry point where development should start.

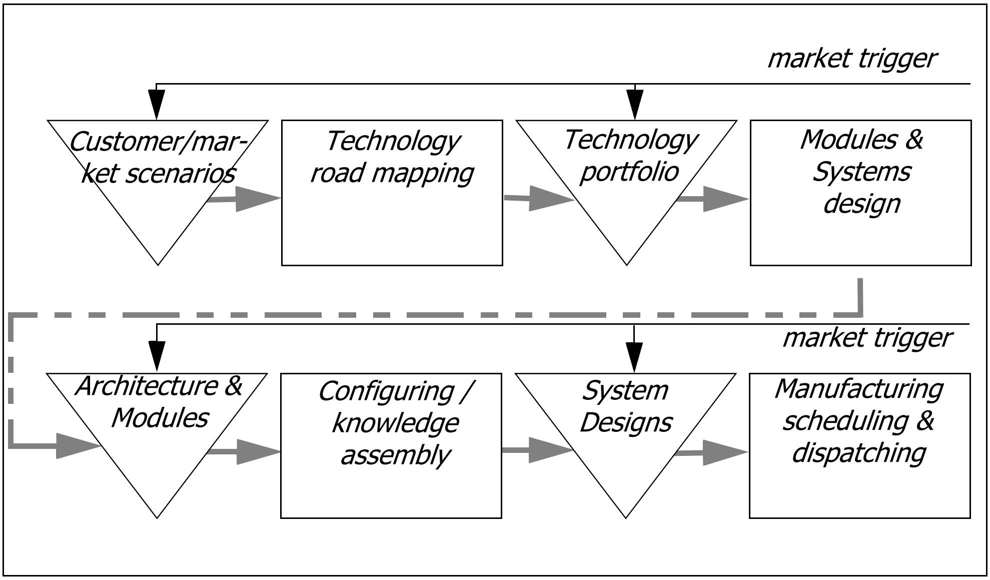


Figure 4.3: Knowledge logistics process

In manufacturing processes, the number of inventory points is usually kept as low as possible because inventory takes space and ties up working capital. In knowledge logistics processes inventories don't take a lot of space and there is no need to keep multiple copies of the same item. There is no capital tied up in the same sense as capital tied up in materials. The value of a knowledge inventory is like the value of option stock. If the options are called, the multiplier in manufacturing will generate the profits. If the option is never called, only its development cost are lost. So in practice a knowledge logistics process will show all inventory points and processes. The market opportunities trigger, where to pick-up knowledge developed in the past and start a particular development for the opportunity. All processes work concurrently on keeping the knowledge inventories full and up-to-date.

4.3 Concepts related to knowledge logistics

TAO©

Based on three domains (technology, organization and application) Voûte has developed the TAO©-model as a road-map for all the translation problems that may occur in the process of transforming a technological invention into a successful innovation. The TAO©-model presents an integral picture that even includes the customer (figure 4.4) [Voûte 2000].

As a road-map, TAO is a visual communication aid that facilitates discussions between engineers, managers and others involved in innovation or development processes. The design of TAO is a kind of marketplace. At each level or row an aspect of an organization matches Technology Supply with Application Demand. The rows are ranked in order of increased dynamics. At the top level the dynamics is low in the sense that there always will be science, there will be users and there is business development that tries to match the opportunities offered by science with the needs of users. At the bottom level the dynamics is that of the real market place. This is the level that normally receives most management attention because this is the level of the material logistics and the real cash flows that are needed to feed the technology and organization columns. The rows of TAO show similarity with the generic steps in the knowledge logistics process.

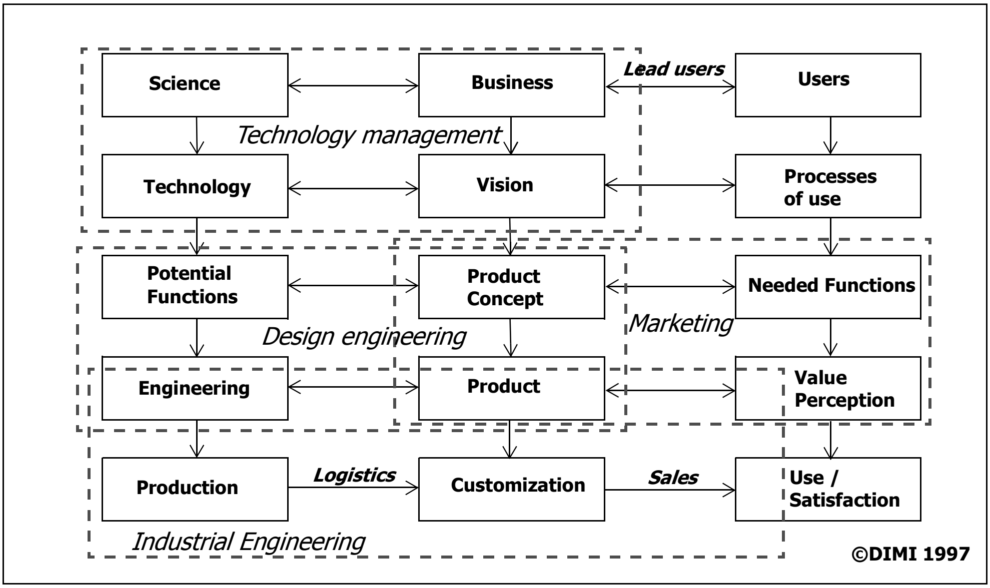


Figure 4.4: The TAO©-model [Voûte 2000]

Project Maps

Wheelwright and Clark have introduced project maps as a means to position development projects in terms of change needed in either the product or the manufacturing processes. They argue that the greater the change in either dimension, the more resources are needed [Wheelwright 1992].

Wheelwright and Clark have used this tool to visualize the strategic positioning of projects and the amount of resources allocated and they use this perspective for a more balanced reallocation of the resources. With this map it was also possible to visualize the relation between projects in different stages. This too is an application of knowledge logistics. In fact the project categories Breakthrough, Platform and

Derivative correspond to Technology Development, Modules and System design and Configuring.

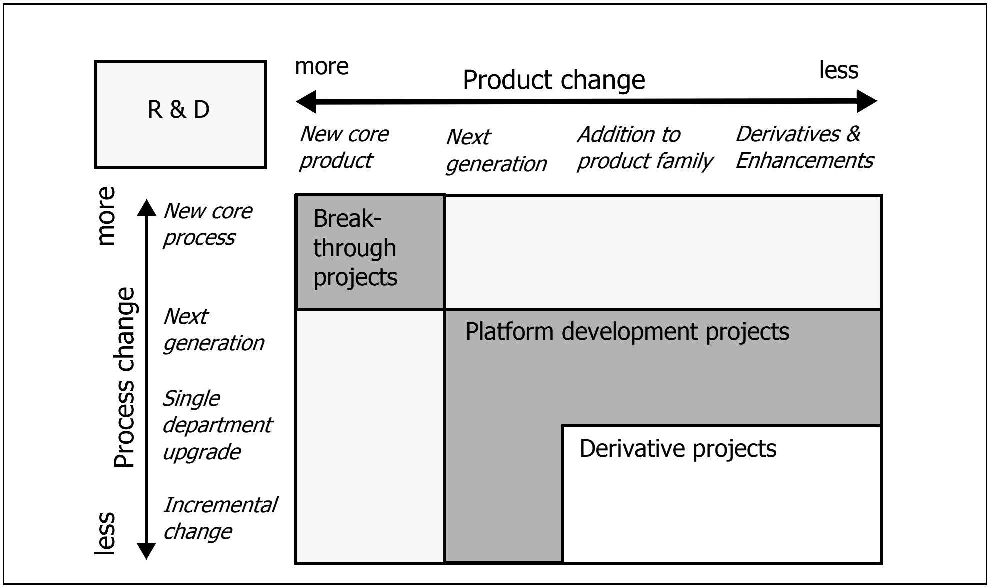


Figure 4.5: Project maps [Wheelwright 1992]

4.3.1 Knowledge logistics in practice

The concept of knowledge logistics can be very powerful in a market where technology innovation is very rapid and where the actual marketplace is showing hyper growth. Because of the growth, capital is not a problem. Investors are very eager to participate in future fortunes, so almost anyone with a good idea gets funded. Entry costs are relatively low. As a result there is no foreseeable dominant design in any of the knowledge development stages. In this situation the only certainty is uncertainty; a one horse betting strategy can prove suicidal.

Decoupling all the stages and applying a portfolio strategy for all stages is the only way to make sure that wherever the market goes, there will be a process- and product-track available to follow fast enough to stay in business. Configuring system designs from modular sub-system sets is a responsive solution as long as the designed systems are also competitive in price and performance. Therefore alternative technologies need to be in development as well to have a new technology platform available as soon as the competitiveness of the platform currently exploited starts to degrade. However the resources (capital, skills and time) to do this are never unlimited for a single company.

The value creation model at Lucent Technologies Inc. proposes a methodology to match timing, probabilities, risks and resource allocation. (investments) with projected cash flows The basic principle is simple and effective; a big investment with a high risk, should always be matched with potentially high revenues. On top of this, a portfolio strategy balances the project mix both in terms of timing and risk. Having only high-risk projects really is a one-horse bet [Walsh 1998]. The combination of timing, risk and available skills may also indicate that the acquisition of technology through alliances or takeovers may be more attractive than internal development.

4.4 Concepts of self organization

Self organization is a term that has at least two interpretations. In the area of systems control and cybernetics self organization refers to systems that are capable of changing their structure and their functionality in order to adapt to new environments and exhibit new interactions. Another perspective on self-organization originates from a systems perspective on understanding nature, life and organizations. This perspective named autopoiesis, does not take adaptation as a response to changes in the environment as axiom. Instead it claims that living structures influence or adapt their environment as a means to self-maintain and improve their chance of reproduction.

These perspectives are complementary and in some sense it is a matter of choice of systems boundaries: what elements or sub-systems are considered fixed and unchangeable and what are the mechanisms for re-configuration to create new responses for the "complete" system.

Both perspectives on self organization are relevant for understanding and improving design processes. The systems and cybernetics approach may be useful for developing adaptive design support systems. The biology driven theories may be useful in developing a better understanding of agents that can act in a design context. Agents can be designers as well as their ideas.

4.4.1 Autopoiesis

Maturana and Varela developed a theory of self organization which they coined by the term autopoiesis [Maturana 1980]. As biologists, their motivation was a desire to grasp the identity of living systems in terms of their autonomy as a phenomenon of their operation as unitary systems. They argue that living systems are organizationally closed, autonomous systems of interaction that make reference to themselves: *An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components that produces components which: (i) through their actions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in the space in which they (the components) exist by specifying the topological domain of its realization as such a network.*

The key in this definition is the recursive formulation in "*processes of production of components that produce components that……*" The theory starts with the distinction between systems that are self-referred versus systems that are *allo-referred*; in other words systems that can be characterized with reference to themselves versus systems that can only be characterized with reference to a context.

Maturana and Varela state that autopoietic systems are purposeless in the sense that any purposeful interaction of autopoietic systems with other autopoietic or non-autopoietic systems is a construct of observations that only belong to the domain of observed actions. Since there are many structures and organizations possible and capable of generating these interactions, the interactions themselves do not reveal the organization or internal structure of autopoietic systems. Similarly, if the observed interactions only belong to the domain of the observed interactions then forcing these interactions to adapt to a changed environment, if at all possible, does not necessarily lead to changes in the elements and the internal structure.

Living autopoietic systems shape their environment through selectively applying their portfolio of interactions. For living systems this is a natural thing to do as long as the capacities for self maintenance and reproduction benefit. As a consequence, the basic skills for demonstrating a supposedly new or adapted interaction, must have been latently present all along.

Changes in the internal elements and structure of autopoietic systems in order to adapt only happen through processes of selective reproduction. The distribution of the diversity of certain interactions may change through selection in reproduction. In addition, sequential reproduction with the possibility of random change in each reproductive step necessarily leads to evolution. Strictly speaking, autopoietic systems can only exist within a limited time frame of manageable external conditions. The length of this period is determined by the maximum number of non-manageable changes a species can absorb from one generation to the next.

With respect to cognition and knowledge the autopoiesis theory has implications for science in general as well as for design in particular. Maturana and Varela state: *It is very often believed assumed that observation and experimentation are alone sufficient to reveal the nature of living systems and no theoretical analysis is expected to be necessary and least of all sufficient for a characterization of the living organization. It would be long to state that why we depart from this radical empiricism. Let us simply say that we believe that epistemological and historical arguments may more than justify the contrary view: every experimentation or observation implies a theoretical perspective, and no experimentation or observation has significance or can be interpreted outside the theoretical framework in which it took place.*

This sounds very much like a definition of self-referred theories. In fact it claims that theories, self-referred systems alike, span a closed space of possible experiments and observations. As discussed in appendix D of this thesis abduction, induction, deduction reasoning modes, express this as the importance of axioms. The autopoietic theory on observation implies that axioms are in fact responsible for spanning the space of existence for both theories and facts.

Autopoiesis theory and design

Although the autopoiesis theory is an accepted system theory for living systems and although we can not deny that living systems have a capacity for autonomy and self-organization, the theory is not directly useful for self organization in design. Autopoietic systems have self organization capacities in their interactions, but not in their internal structure or elements. To understand design in terms of an autopoietic system we need to identify design agents as autopoietic systems and understand their drive for self-maintenance and reproduction, in the context of an environment that rewards specific designs. Thus the most interesting part of autopoiesis theory for design is the reproduction and evolution part.

4.4.2 Self organization and complexity, a cybernetics view

In an attempt to define criteria for self-organization, Gershenson and Heylighen conclude that self-organization is not a property of a system as such; it is the result of a system-perspective imposed by an observer [Gershenson 2003, Heylighen 2003]. This viewpoint is consistent with similar statements made by Maturana and Varela. In supporting this observation Gerhenson and Heylighen make use of the general concept of statistical entropy or Shannon's information entropy. Increased order, corresponds to a decrease in entropy. Increased disorder corresponds to an increase in entropy. However, in line with the theory presented in paragraph 3.2.2 and paragraph 3.2.3, entropy values are relative to the definitions of state variables, aggregated sets of state variables and their distributions as defined by the researcher. If the number of aggregated sets is the same, two distinct aggregations based on the same collection of states will demonstrate different entropy values. But unless a mechanism for aggregation is postulated, it is impossible to tell which one is preferred. Nor can we tell if these systems are self-organizing of self-disorganizing.

Gershenson and Heylighen define *organization* as *structure with function*. Self organization then means that a functional structure appears and maintains itself spontaneously. They mention two important properties of self-organized systems:

1. *Distributed control of mechanisms that establish relations.*
2. *Self-reinforcement of mechanisms that establish relations.*

The control needed for self-organization needs to be distributed over all participating components. The drive and capability to establish relations needs to be present in all individual elements rather than centrally coordinated. These mechanisms need to be self-reinforcing. If connecting to peers is rewarding and can be signalled to others that have not yet joined, then this is a self-reinforcing mechanism for growth. Larger clusters can attract more individuals to join because they can provide more and better opportunities to join. Growth only stops when there are no individuals left with a capability to join.

The emergent structures can change under the influence of external changes or as a response to random fluctuations in the rewarded interaction patterns. Changes occur locally and may require new relations to be established. Successful (rewarded under the changed regime) relations are signalled and copied by others. For bigger clusters, changes start outside and propagate inwards.

4.4.3 Self organization and design processes

For developing an understanding how self organization in design may be accomplished, we can look at how the information entropy of a design problem may evolve. Suppose, designers work on sets of concepts and ideas. Each concept or idea can be represented by a state variable and an associated probability of fit to the final system design. The total entropy of this system is represented by formula 3.4.

$$H(S) = \sum_{i=1}^{k} H(S_i) + \sum_{i,j=1}^{k} H\langle S_i | S_j \rangle$$

If no aggregations are considered then each set contains only 1 state and the entropy is equal to the entropy of the non-aggregated system. The natural aggregations we try to find in design are functionally related sets of problems that when solved, represent a functional component of the design. Each set S_i may relate to one component. The relation to other functional components of the design is represented by the conditional information term $H\langle S_i | S_j \rangle$. Minimizing the total entropy of this system means lowering k, the number of functional components needed for the complete design while minimizing the complexity of the interfaces represented by the conditional information terms.

In practice this means defining constraints or interface requirements that guarantee a fit. Within the sets of possible solutions, this can be achieved through promoting functional redundancy in combination with diversity with respect to interface requirements. Designers are challenged to create as many options as possible, thus maximizing the internal entropy of each set. To make this process self-organizing means:

1. *Distributed control over establishing relations. Concepts or ideas do not function in isolation. Only in combination with other ideas, they can contribute to the realization of customer value attributes*[1]*. Yet these connections need to be negotiable. Designers and their ideas recognise a good fit as well as opportunities to influence others to create a good fit for a strong concept.*
2. *Reinforcement of the reward mechanisms. The reward mechanisms need to reward support of functionality as specified by customer value attributes as well as keep the conditional information*[2] *as low as possible. Clusters grow until the support for customer value attributes is strong and the complexity of connections with the outside world is low.*

---

1. See paragraph 2.4.1 and [Meijer 2004, Kemperman 1999]
2. The coordination between functional sets.

For this process not to get stuck in local, coincidental clusters, there need to be competing functional clusters as long as the overall architecture is not fixed. Working with redundant sets of ideas and combining these ideas in a self organized manner has much in common with evolutionary problem solving. In the next paragraph evolutionary problem solving is described in more detail to develop these ideas pm self organization in design one step further.

4.5 Evolutionary problem solving

Evolutionary problem solving is based on the structure of a genetic algorithm. The basic structure of the genetic algorithm, originally developed by John Holland is as follows [Holland 1975, Banzhaf 1998]:

1. *Initialize a starting population of physically feasible solutions*
2. *Create a new generation through genetic operations such as mutation and crossover and reproduction.*
3. *Rank this population using the fitness function.*
4. *Select the top of this population and randomly select a couple of others to create a new starting population.*
5. *Repeat steps 2-4 until the top-member of a generation has a sufficient fitness score to be acceptable as a solution.*

This scheme is sufficient for understanding how the algorithm works. An optimization problem is represented as a vector or a string of variables for which good or preferably optimal values have to be determined. In genetic terminology one could call this a chromosome. We can create a set of physically feasible but not yet optimal vectors. This is the initialization of the population.

Next we create a new generation through genetic operations. Mutations can randomly change the value of variables or even replace some of the variables with others. Crossovers cut a part from the chromosome and try to replace this part with a similar part from another chromosome. Reproduction is simply creating a copy of the chromosome. Through these operations a new generation is created.

The ranking and selection is a highly non-linear and irreversible step in the process. It is a decision over life and death. The chromosomes selected survive and will participate in the reproduction at least one time again. The non-selected chromosomes are in fact dead and removed from the process. The fitness function controls the chance of survival of the chromosomes with the best fit, so far. The random selection of some others is important for the process to create diversity and not to get stuck in a local minimum of the fitness landscape, not yet good enough to be accepted as a solution. The processes of creating a new generation, ranking and selection are repeated until the top member shows a sufficient fit to the ranking function to be accepted as solution.

The success of this algorithm can be attributed to two properties that make it distinct from linear optimization techniques. The first property is redundancy and diversity. Rather than developing one solution, genetic algorithms develop and maintain multiple solutions concurrently and it does not easily get stuck in developments that lead nowhere. The second property is the non-linearity of the selection process. With linear optimization the fitness landscape is set from the start by the starting solution and the fitness function. Finding the optimum in this landscape could mean exhaustive search through the entire landscape. Although the fitness landscape is set from the start, a genetic algorithm employs multiple starting points for the search and the generation and selection steps cause the effective fitness landscape to be reshaped at the start of every generation.

For an implementation in software there are a number of additional control parameters and heuristics that are used to guide the creation of a new population. E.g. one may use the ranking or fitness function, to steer how the genetic operations are used. If the selection of the operation is a random process then control of the probabilities for each operation is used. The size of a generation and the size of the starting population are important variables that drive the diversity and the evolution speed of the algorithm. Also there are different strategies possible for selecting the next generation. All of these can be used to influence the efficiency of the algorithm if one can attach a meaningful interpretation to these controls in the context of the problem one is trying to solve [Banzhaf 1998].

4.5.1 Evolutionary problem solving and self organization

The genetic algorithm is a system model for the self reproduction principles of the autopoiesis theorem. The solution patterns genetic algorithms may generate are predominantly the result of the initial set of solutions that were present at the start. The fitness function is the context within which structure changes may occur as long as survival as a unity or species is not at stake. Changing the fitness function will cause serious changes and may also cause death in case the present elements cannot generate a sufficient fit (survival) to the new fitness function. In case of survival, biologists may recognize evolution, but they may also claim that the new organism is a different unity or species that is capable of a different set of interactions, fit for the new context.

In a design context it does not matter whether the design has a different structure and should be identified as a new "species", as long as the design supports the customer value attributes it was designed for. Evolutionary problem solving is a good model for putting self-organization principles to practice. In addition to the requirements for self organization (paragraph 4.4.3) the selection process after each generation, serves as an effective means to reduce entropy and to select more promising interfaces and boundary conditions as a point of reference for the next generation of ideas.

In the next paragraph Set-based concurrent engineering is described, since this relatively novel approach to concurrent engineering exhibits many characteristics of self-organization and evolutionary problem solving as developed in the previous sections.

4.6 Set-based concurrent engineering

Ward and his co-authors argue that in concurrent engineering there are two fundamentally different approaches to be recognized: *point-based and set-based*. In case of point-based design, a single solution is synthesized first, then analysed and changed accordingly. Even though the phases of the design process may be executed concurrently, all designers and specialists invest their efforts in the pursuit of only one concept that is to be developed into a solution.

In set-based concurrent engineering, designers explicitly communicate and think about sets of design alternatives at both conceptual and parametric levels. The efficiency of set-based versus point-based design is that in communicating sets, implicitly or explicitly all designers become more focused on relations and constraints between different aspects of the design than they would be when focusing at a point solution. All designers communicate their range of options rather than one preferred option. Sometimes to maintain focus, constraints for these sets can be tighter than they would be in case of a point based design [Ward 1995].

Ward and his co-authors found evidence that Japanese companies and more in particular Toyota and Nippon Denso deploy concurrent engineering practices that have much in common with the set-based concurrent engineering philosophy. In the next section the Toyota and Nippon Denso concurrent engineering processes are exposed in more detail [Sobek II 1996].

4.6.1 The Toyota and Nippon Denso case

The set-based engineering processes of Toyota have the following characteristics [Ward 1995]:

1. *The team defines a set of solutions at the system level rather than a single solution.*
2. *It defines sets of possible solutions for various subsystems.*
3. *It explores these possible solutions in parallel, using analysis, design rules and experiments to characterize a set of possible solutions.*
4. *It uses analysis to gradually narrow the sets of solutions. In particular the team uses analysis of the set of possibilities for subsystems to determine appropriate specifications to impose on those subsystems.*
5. *Once the team establishes a single solution for any part of the design, it does not change it unless absolutely necessary.*

Sets can be created by using design ranges which can be narrowed rationally once these areas have been explored. Toyota does this, but is not limited to design ranges for defining sets. Toyota makes extensive use of so-called “lessons learned” books, not simply to record experience but also to define the space of manufacturable designs. In the process effort is made to avoid changes that expand the space of possible designs. In this way decisions remain valid through the project's life. In the early stages of conceptual car-body design, experts from all functional areas review all the alternatives, adding their manufacturability ranges to these designs and specifying possible conflict areas where existing capabilities may be insufficient. Sometimes these conflict areas can be resolved through making changes to the concept design, in other cases these areas give rise to capability enhancement projects to make the new design feasible. The lessons learned books also provide an opportunity for institutional learning. Documenting all explored solution areas and the starting point for each development provides possibilities for backtracking developments to their roots and maintains sight at built in limitations that may not be so obvious any more, once the concept has been reused and changed 4 times over.

Nippon Denso, a partner of Toyota and a major automotive supplier of components as well as systems, also applies a process that has characteristics of set-based concurrent engineering and extends this even to pre-design R&D. In this process the degree of parallelism and redundancy is much higher than typically with Toyota. As an automotive supplier, the demand for diversity is higher and their competitiveness is much affected by new technologies and new materials. In order to push the limits and to stay ahead of competition Nippon Denso tests as many ideas as they can in order to create a platform (set) of solutions that is competitive and can be easily adapted to the specific interfacing requirements of different car makes. What may be a surprise is that Nippon Denso's development processes have a start that may be 3-5 years ahead of the start of the car development processes that adopt the new designs. Rather than pursuing rapid development once the outline of the specification from their customers is clear, Nippon Denso pursues radical breakthrough designs that are ready before their customers ask for it. When they start working with their customers, the focus is on interfacing and not on the core technology, thus avoiding the major part of development risk.

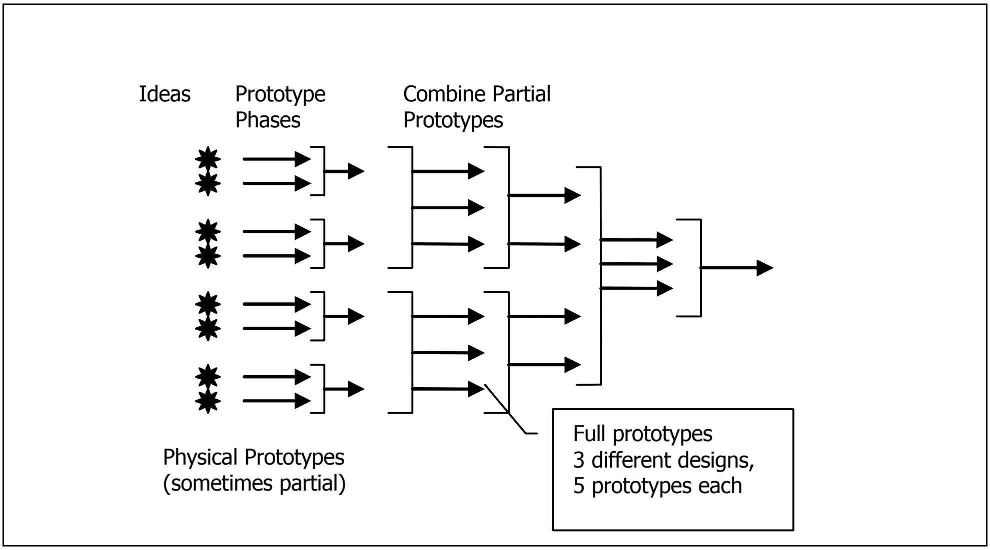


Figure 4.6: Nippon Denso's R&D Process [Ward 1995]

4.7 Evolutionary organization of design processes

It is possible to take the idea of set-based concurrent engineering one step further if we study the similarities between evolutionary problem solving and the practices of set-based concurrent engineering as implemented by Toyota and Nippon Denso. The key to this development, is understanding the importance of the lessons learned book and the idea of combining partial prototypes as implemented by Nippon Denso. Combining partial prototypes is like implementing crossings. The lessons learned book contains information on present and past fitness functions. This information is also important for assessing the fitness to the new design. The Nippon Denso process starts with many concepts in parallel; another key element of evolutionary design processes: parallelism and redundancy. Together these concepts contain the kernel of the evolution process; that is how to start and create a population of promising designs.

We can make the process organizational by adopting the structure of a genetic algorithm and by taking measures to introduce and maintain the redundancy and parallelism in the process as well as in the organization. To do this, the following assumptions are needed. Most of these assumptions are general in nature; they apply to any project that is targeted and confined with respect to time and resources:

- *A target exists and the target requirements can be transformed into a fitness function or into a set of fitness functions for the evaluation of results.*
- *There is a deadline. This means the process has to finalize before a certain date in order to meet the market window for introduction.*
- *All resources to do the project are available.*

Now the evolutionary organisation of a development process can be implemented as follows:

1. *Divide the staff into n independent teams that are all capable of executing the entire project and give all these teams an identical assignment and dead-line.*
2. *The teams will develop their concepts and solutions following set-based concurrent engineering practices and they record their achievements in lesson's learned books.*
3. *At regular intervals, a fair is organized where all teams present their progress and give insight in their lessons learned books.*
4. *At these fairs team members look around for promising partial solutions with their competitors.*
5. *After the fair, teams continue their own development, including some ideas inspired by the last fair.*
6. *If a (required number of) design(s) with sufficient fit has been achieved, stop else repeat steps 2-5.*

The processes within the teams could also have characteristics of evolutionary problem solving, if they apply brainstorming techniques for finding and selecting ideas. However the fair is really the place where crossovers and mutations occur. At the fair everyone is looking for clever ideas that could fit to their own concept (crossings) and ideas of other groups may also trigger new thoughts (mutations).

Although a fitness function exists and is used for objectively ranking ideas, the organizational form of a genetic algorithm has the advantage that besides the objective ranking, the fair causes (partial) ideas also to be ranked subjectively and implicit. Ideas that may not be very successful in one context, may receive a much higher ranking in the context of a different concept. Partial ideas can also inspire individuals to new ideas. Thus all ideas receive multiple rankings, improving their chance of survival and supporting migration of ideas between concepts.

In case of only one explicit ranking per concept these ideas could have been lost. The implicit and subjective ranking also solves a social problem of working with a large engineering force. In large engineering forces, a dozen socially dominant engineers will monopolize the decision making at centralized meetings to a degree where a significant portion of the engineering staff effectively has no influence. Because the central meeting is now replaced by a fair where implicit recognition is the mechanism for survival of ideas and not the socio-political skills of some engineers, good ideas regardless of their source stand a good chance of being inherited into the final concept.

This process can be made more efficient if overlaid with a structured design process where the fairs will be synchronized with stage gates and deadlines in the overall development process. Such a measure will synchronize the efforts of all the teams. The knowledge logistics concept could provide for such an overlay.

4.8 Discussion and Conclusion

With the evolutionary problem solving model and the evolutionary organization proposed in the previous section a recipe for implementing self organization in design is offered. The process is self organizing to the degree that the structure of the relations and the parts that provide the best fit for the target is not the result of careful causal sequential reasoning. It is the result of a process that provided focus through its fitness function and that at the same time allowed maximum freedom to copy, mutate and combine ideas from resources that at the starting point may not even have been part of the design assets.

The autopoiesis theory and the complexity/cybernetic perspective on self organization share the concept of choosing an observer or perspective. The Autopoiesis theory makes explicit that we can not expect behaviours from our agents other than the interaction patterns that were implemented through the internal elements and structure. Autopoietic systems can only evolve through evolution. The biggest challenge for design is to understand better how to deploy these principles of evolution. More specifically; how to shape fitness functions that create focus as well as self-reinforcement towards the desired properties of the design.

From theory of evolutionary problem solving it is learned that redundancy and concurrency are the most important ingredients for an efficient and effective optimization process. Developments as set-based concurrent engineering already make effective use of these qualities.

Redundancy is also one of the key properties of the knowledge logistics concept introduced in this chapter. Through creating a "buffered pipeline" of ideas that covers all aspects of research and development, including business development and product-planning, it is possible to shorten the market leadtime between recognizing the business opportunity and being on the market. In addition, systematic storage and retrieval of research results, no matter if related to market intelligence, technology or product platforms, provides a memory base that is essential for organizational learning. Only if good concepts can be memorized and reused, design organizations can become really efficient.

At the same time design and development organizations must also have the ability to forget in order to innovate. Regularly hiring new engineers and designers may "refresh" the people based memory of a development organisation. Their ambition to implement their own ideas or one could say to leave their (design)DNA supports the innovation processes. Granting these people access to databases on designs from the past helps them to become efficient by also inheriting succesfull ideas from the past.

5. BOUNDARY CONDITIONS, COMMUNICATION AND CHANGE.

The structuring principles presented in this thesis are all based on principles from mathematics or physics that have been proven effective for technical system design. For these ideas to work for organizations that involve humans as sensors and actors, it is necessary to study if and under what conditions, humans will comply with the processes and structures that result from a technical system design approach. If such conditions exist then we can safely treat socially induced disturbances as boundary condition to be treated separately from the organization structure design. In this chapter Maslow's hierarchy of needs, Belbin's teamwork theory as well as some theories for creative problem solving are studied to identify conditions under which human's will generally comply with processes and an organizational context that is based on technical system design methods. Moreover, with Maslow in mind it is argued that a vision of processes and structures is important to communicate as part of reorganization and change process. Understanding human communication is a key tool in making humans comply. For this purpose a model of a human to human communication cycle has been developed and presented in this chapter.

5.1 Introduction

Human workers can fulfil many roles in business processes, ranging from offering their capacity for manual labour to being a knowledge worker and a manager. Real jobs consist of a mixture of all of these. Unskilled manual labour workers are not human powered automates. Like any other human, they use their senses, take decisions to perform certain actions and they communicate with their peers as well as with their boss about problems of various nature. Knowledge workers as well as managers also operate "machines" for engineering, documenting and communicating their ideas. Thus although the balance may be different, all humans are involved in learning, communication and decision making processes as well as in providing some kind of manual labour to business processes.

The biophysical properties of humans interacting with machines is studied extensively in research fields such as ergonomics [Sanders 1993], the study of Man-Machine Systems and some very interesting results that can be found in the area of rehabilitation robotics [Hogan 2000]. Yet these characteristics have little influence on the organization of business processes.

This chapter focuses on the communication and decision making processes of humans that are at the heart of development processes as well as management processes. We do this by looking at process conditions for creativity, teamwork and human communication processes. However, today it is clear that computers are also important resources in these processes. The first section of this chapter investigates computer support in the context of "creative" problem solving. The aim is to identify boundary conditions for effective and efficient computer support and to state more explicitly where the human qualities in decision processes are. The remaining part of this chapter focuses on describing conditions under which humans commit their best efforts to work processes even in a situation of organizational change.

At the heart of this chapter is human communication. Inspired by Shannon's model of a noise-free communication channel, a simple model for a human-to-human communication channel was developed. With this model, understanding and locating sources of misunderstanding and mis-communication becomes easier. Engineers and managers can use this knowledge to direct their communication to those issues that need to be solved to improve the effectiveness and efficiency of their communication.

5.2 Machine intelligence and creative problem solving machines in perspective

Generally speaking, computers have more power in algorithmic problem solving then humans. From this perspective one could ask, will someday computers or machines be capable of doing more than the support role? Machines or computers are certainly capable of synthesizing knowledge. Genetic Algorithms are often applied for automatically solving very complex optimization problems. But is computer synthesized knowledge new knowledge? Does the result of an optimization, open new options? Technically speaking, we can make machines that generate new options through reasoning. However, creative machines that reason in an open knowledge space, have a talent for making mistakes. Even if their synthesis is perfect, which requires a formal knowledge domain description to exist, the computer's world model is never complete and machines have no awareness of the risks associated with incomplete models. So there is always the risk of synthesis on a partial fact-base, which could result in new knowledge that is not useful. If such a machine would act on that knowledge, the result is often a damaged machine or worse.

Beyond damaged hardware, one could question to what extent intelligent machines are desirable. If today, engineers have to built a system that is guaranteed safe and stable in its operation, then that machine has a fixed number of states or in other words a closed and known knowledge space and their possibilities for interaction with their environment are limited. Even then problems can occur if the human operator is not aware or cannot identify the operational state of these machines correctly[1]. Despite what many people expect, intelligent machines are hard to operate[2].

Also be aware that humans have a bigger tolerance for errors or faults if they are being made by other humans. We expect our machines to be a 100% safe and reliable. If failure of machines may lead to human casualties, human supervision and control is always added as an essential part of the operations. If machines fail, investigations are always carried out, possibly followed by criminal prosecution of operators, managers as well as designers that were involved in deploying a smart machine which turned out to be not so smart.

Are machines then useless or undesirable in creative processes? Computers are very efficient data storage and retrieval systems. This property may be very useful in creative processes where unlimited access to ideas, data and images is important. Moreover computers do not need to be perfect at this. Even mishits in the retrieval

1. Through their interactions with a machine and its environment, humans can always create situations outside the scope of working conditions for these machines. At best, the machine switches off but if the machine fails to identify its situation correctly, machine intelligence can cause weird responses.
2. Sometimes this is an organ-boundary problem. The boundary between the machine's responsibility and the operartors responsibility may shift as part of the operation [Degani 1997].

system can inspire humans to creative linking of concepts. In addition a warning must be set out here. Do not underestimate the capacity of team workers to select the right information out of an abundance of data. A creative process is by no means sequential or causal, so there is no way to tell beforehand what information may be needed at what time. Any attempt to filter in advance for the sake of efficiency, has a risk that important data is thrown out too soon. The only efficiency measures that really work are user programmable filters (heuristics) and speed in search and data storage and retrieval.

Computer networks are also capable of lowering communication barriers. Physical distances are no longer a problem. Social distances tend to disappear. Websites, newsgroups and e-mail allow for a-synchronous communication, which is very helpful when synchronous communication is not possible. These are all features that can support problem solving processes and managers could deliberately decide to deploy computer tools to support their program.

5.3 Creative work is human work

Creative work is the domain of human beings. Such a statement calls for some explanation on the meaning of the word creative. Hohn has reviewed definitions of creativity, many of them also based on earlier reviews [Hohn 1999]. What becomes clear from Hohn's discussion is that creativity is a process that results in something new and successful, that did not exist before in the minds and memories of everyone and everything involved in this process. Statements and terminology often used in this field of research are "a successful step into the unknown", "creative thinking as a process of seeing and creating relationships", "thinking by analogy", "heuristic rather than algorithmic". Roger de Bruyn from the centre for the development of creative thinking (COCD) had a one-liner saying: "creativity is elegant stealing". This statement is just another way of defining creativity as thinking by analogy and it also assumes an environment to "steal" from.

Personal creativity versus societal (or group) creativity is a matter that has not been resolved yet. The problem is that even if we assume that the creative breakthrough can only take place in the mind of an individual then we still cannot isolate the influence of the environment, other people, previous experiences and mindset (culture) on this person's act of creativity.

One could ask, does this really matter? If the breakthrough could be attributed to a single mind, it would not be very successful if it was not communicated or shared in some form; success implies a degree of usefulness, valued by others.

What does matter is the environment. "Elegant stealing", "thinking by analogy" and "creating relations" all suggest that things, ideas or concepts that already existed often inspire the breakthrough. Rhodes introduced the 4 P's of creativity: Persons, Process, Press and Products [Rhodes 1961]:

- *Persons* — *refers to all aspects that can be attributed to an individual.*
- *Process* — *applies to motivation, perception, learning, thinking and communication. It refers to the tools of the creative environment.*
- *Press* — *is short for pressure and refers to constraints that come from the relation between human beings and their environment.*
- *Products* — *are the communicated results of the creative environment.*

These four P's can be considered as the dimensions or variables of a creative environment. These are the "knobs" managers can turn to influence the effectiveness and efficiency of the process.

5.3.1 Process, Persons and Press

A powerful and well-known process in creative problem solving is brainstorming. The basic principle of brainstorming is separation in time of the phase to generate of ideas (divergence) and the time to criticize them (convergence). Osborn presents a set of rules for the generation of ideas in brainstorming [Osborn 1963]:

- *Criticism is ruled out, adverse judgment of ideas must be withheld until later.*
- *Free-wheeling is welcomed, the wilder the idea the better, it is easier to tame down than to think up.*

- *Quantity is wanted. The more ideas, the bigger the likelihood of the presence of useful ones.*
- *Combination and improvement are sought. Participants are invited to expand on how ideas of others can be combined or improved into better ones.*

The process of brainstorming shows a strong analogy with genetic algorithms. The separation of generating ideas and ranking or ruling out these ideas is similar. The freewheeling can serve as mutation. The combination and improvement of ideas are either the crossings or the inversions of genetic algorithms [Meijer 1996].

The ranking and selection of ideas is done on the basis of constraints and the goal of the problem solving session. Applying different modes of communication can be vital for both the ranking of ideas as for the generation of ideas. Especially when the solution has to come from the synthesis of a number of different knowledge domains, then it is unlikely that a common spoken or written language exists. Being able to use other communication modalities, such as showing drawings, images, listening to sounds or watching movies could be a necessity to keep the generation process going, but also for finding the right criteria. Defining a ranking method itself, may require a brainstorm session.

The social context in which the generation process can flourish is one of mutual trust. All participants, whatever their personal attitude is, must feel free to put forward ideas without censorship or criticism. This condition is often violated if the normal working environment and decision making is based on hierarchy. Behavioural patterns based on past experiences may cause individuals to apply internal (personal) censorship, before coming forward with their ideas. The same thing may happen if there are schools of experts involved. If these experts know each other and their respective views then they are more likely to comply with one of the schools of thought than they are to come up with something new. A tool to overcome this may be anonimization of the responses e.g. through Delphi sessions. However this puts even more pressure on the availability of communication technology to support other modalities of communication.

Belbin has done research on team roles that need to be present in teams to be successful [Belbin, 1993].

Together the set of team roles (table 5.1) represents a team that is not only capable of creative problem solving but also to complete the assignment into daily working practices. Some of these roles can be combined into one person. Depending on the size of the team, the role distribution may be different. Not all of these roles are part of the creative process, but in Belbin's definition teams do more than creative problem solving.

TEAM ROLE	ATTITUDE
Coordinator	*Chairman, makes goals and restrictions clear, able to delegate, independent yet good*
Resource investigator	*Builds networks, link with the outside world, extravert personality*
Plant	*Intelligent, creative, unorthodox thinker, dominant*
Shaper	*Dynamic personality, gets things done, focused courageous decision maker*
Monitor / Evaluator	*Visionary, strategic, foresees options and warns for major problems, independent*
Team worker	*Cooperative, diplomatic, supportive*
Implementer	*Practical, reliable, efficient and conservative*
Completer	*industrious, cautious, eye for detail, reliable*
Specialist	*Focused, skilled however no helicopter view, passive*

Table 5.1: Team roles of Belbin

Again looking at the analogy with genetic algorithms the plant and the resource investigator can act as genetic operations that combine whatever resources occur to them into options and solutions. The monitor is in the role of setting the criteria and is very important to prevent groupthink. The shaper gets the process going and

together with the implementer and completer gets the project out. They are more important in the convergence phase, although it sometimes takes a shaper to get the plant and resource investigator going.
The shaper, the coordinator and the monitor/ evaluator can put pressure on the team. The coordinator and team worker can relieve this pressure and thus keep the team socially together. Thus these roles represent a mixture of roles and attitudes that causes the team to be cohesive without the group-think risk; creative but at the same time practical and focused.

As already mentioned, people have to feel free to come forward with their ideas and opinions, in order to contribute to problem solving. Maslow sees this as an act of self actualization, that will only come forward if the supporting levels in his hierarchy of needs are fulfilled. The hierarchy starts at the bottom with physiological and safety

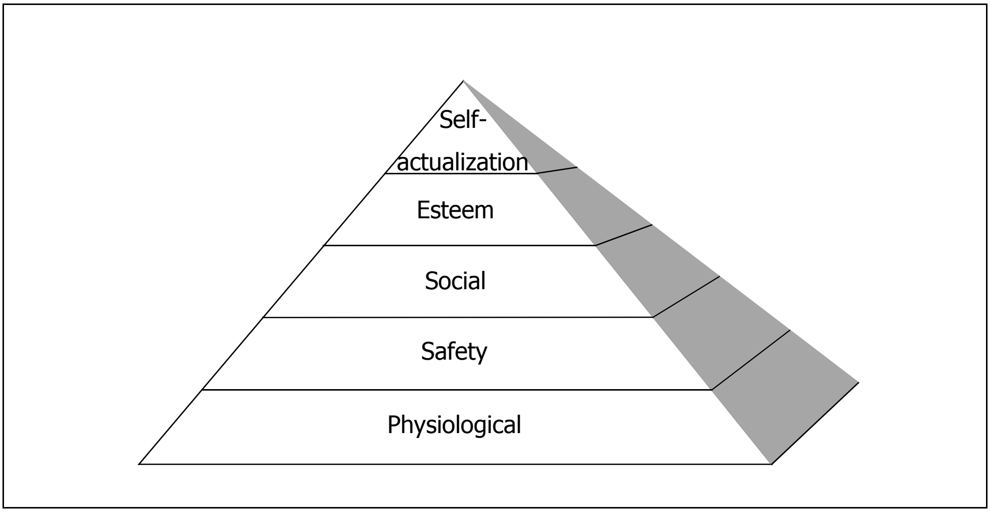


Figure 5.1: Maslow's hierarchy of needs

needs. These are people's prime interests if survival is at stake. On top of those come social needs, the need to belong to a group of human beings and the need for esteem (esteem from others and self-esteem). Esteem, self esteem and respect are factors that are not always present in professional organizations. Only when a person has no worries, does not need to keep up any guards or holds any reserves, only then his drive to come forward with his best will be out of the need to self actualize [Maslow 1998].

This may seem easy and obvious, to turn this model into a management paradigm is something else. The hierarchy of needs is designed for individuals. As a metaphor, one could apply Maslow's hierarchy of needs to organizations as well. However this is not the same as building an organization on the principles laid out by Maslow. The problem is that someone's need to self-actualize may be in violation with someone else's need for esteem or social safety. This is yet another reason for Belbin's team worker and coordinator role. They must safeguard the teamspirit so that everyone feels respected and acts out of self-actualization.

None the less building an organisation applying Maslow's principles is not impossible. A section from an interview with Anne Robinson, co-founder of Windham Hill Records, records:
"I think it was a combination of the work and my strong belief that if I worked to create and environment in where people felt empowered, they would bring to their work the very best they had to give. When a company grows large, as happened to Windham Hill, there comes a point where you have to hope that you've hired people who have a vision and integrity that complements your own.

In looking back, I also think that employees sensed that we were doing something different from the rest of the industry. They sensed that our products had meaning to people. I felt strongly about making a product that had lasting value. I think our employees took real pride in that belief. The product reflected our values. I think our employees, as well as our distributors and suppliers, bought into our business philosophy. The true struggle for me was after merging with BMG records. I had quarter-to-quarter projections to make and a bottom line to meet. Yet I realized that I needed to work to hard at keeping the same value system within the organization or the end product would suffer" [*Maslow 1998 p8*].

Maslow's puts this phenomenon in the following words:

"Learning, creativity, fairness, responsibility, and justice come naturally to people according to Maslow's theories. Why is it that we often design organizations as if people naturally shirk responsibility, do only what is required, resist learning, and can't be trusted to do the right thing?
Yet, most of us would argue that the believe in the potential of people and that people are our most important organizational assets. If that is the case, why then do we frequently design organizations to satisfy our need for control and not to maximize the contributions of people. …. For centuries human nature has been sold short". [*Maslow 1998, p.11*].

Finally without (time) pressure the work will never finish and even the quality of creative work may benefit from pressure. Deadlines are needed to plan the work towards meeting that deadline and will force the team to put in that extra effort needed to achieve that extra level required. Deadlines can also shape Belbin's team such that plants are allowed as long as concepts are needed and forces shapers and monitor/evaluators to take over to reach a decision when it is due.

5.4 Wittgenstein; from images to language

Presenting Wittgenstein's Tractatus in just one paragraph is a mission impossible. The statements below are just a very small selection of the first three "chapters" [Wittgenstein 1922]:

1	The world is a set of facts and the set of facts is the world
2	It is a fact that there exist connections and those connections represent relations between facts
2.1	We make images of facts
2.224	From the image alone we cannot tell its truth.
2.225	To establish the truth of an image, we need to compare it to the world.
3	The logical image of the facts is the thought
3.01	The collection of true thoughts is an image of the world
3.1	We can express a sensible thought in a full sentence.
3.11	We use the sensible token (audible, visible) of a sentence as a projection of the possible state. The projection method is the thought of the meaning of a sentence.

What is interesting from these propositions is that the sequence of Wittgenstein's logic suggests that individual humans first understand their world through building images of sensed connections and facts and only then start working on expressing these understandings in language for more efficient communication. The meaning of the thought expressed in proposition 3.11 is the first step in a "language" based communication process. Real communication requires at least two individuals to agree on meaning. Sentences as well as the interpretations should be exchanged both ways until there are no doubts left about each other's interpretation.

Although it is difficult to prove Wittgenstein's propositions, there are observations that support these propositions. One of them is the development process of small children. They first learn to see and understand their world by moving around, touching and by copying the observed behaviour of their parents. Any action originating from misconceptions is corrected by binary signals from their parents (approval or disapproval). Only after 18 months children start to develop language for more efficient communication and again this language is learned starting with pointing at known concepts and linking spoken words to it (shared experiences).

It is argued that this process is also valid for the (re)searching, learning, developing and communication processes of knowledge workers. The role of the parents however is now replaced by experiment and by past and personal experiences. Especially in creative processes concepts, understanding, language are often incomplete and insufficient. Yet communication is the key element in the creativity process.

This observation is very important, because it means that we must be aware of the concepts and images in people's minds if one wants to improve knowledge communication. Being able to communicate through sharing experiences, building a common understanding of a concept, may not sound very efficient, but with Wittgenstein in mind, it may well be the only way. Note that a conflict between two engineers over a solution for a particular problem is hardly ever about the solutions themselves, it is about the criteria for valuing these solutions. Differences in conceptual understanding lead to different criteria.

Designers and Mechanical Engineers only have a very limited language compared to the rich collection of their artifacts. They don't need to have a rich language because they communicate in drawings (images) and sculptures (shapes). This also implies that building a knowledge system in the domain of design engineering and mechanical engineering is a very hard task, that should start with designing a formal language that is rich enough to describe their knowledge. Conceptual design research today tends to focus on a process description language and on multi-modal communication support (rapid prototyping, virtual reality and smart annotation interfaces). This means in the absence of a formal language, technology development is mainly aimed at better and more efficient support, but not at replacing what the human mind is good at.

5.5 Numata; knowledge amplification

In their paper, Numata and Maeda analyse and describe the product development processes of Sony and they introduce a concept called knowledge amplification. In

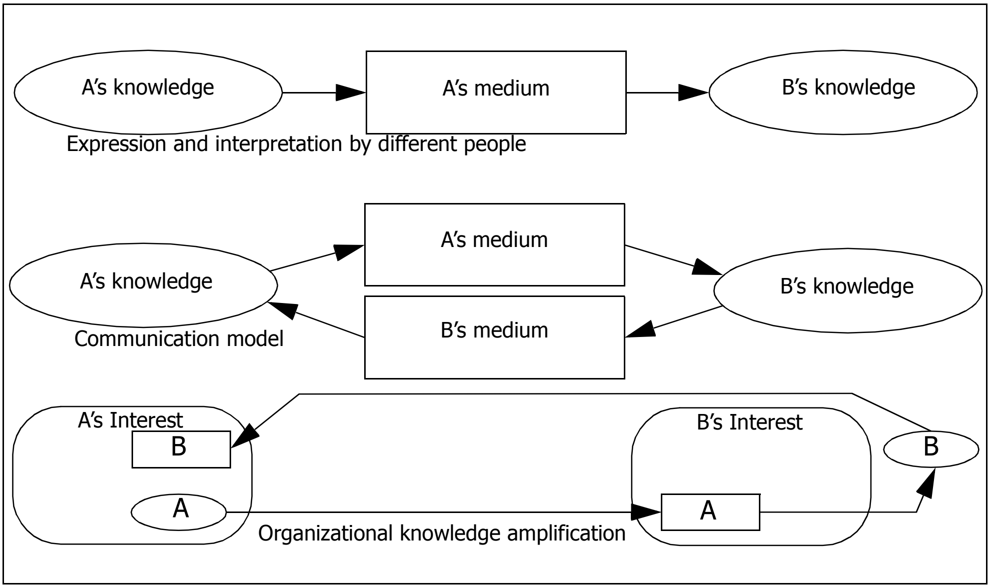


Figure 5.2: Knowledge amplification [Numata 1998]

agreement with Wittgenstein, their communication concept acknowledges the difference between the knowledge of an individual and the medium this person uses to communicate this knowledge to another person. Preferably, the medium of the receiver is used by the sender (Figure 5.2) [Numata 1998].

In addition to Wittgenstein, the concept of knowledge amplification argues that for successful communication it is also necessary to understand an individuals intent or interest behind the knowledge and communication. Intent or interest is more difficult to synchronize than conceptual understanding of the problem. In case of competition, intent can even stop communication effectively. Wittgenstein also acknowledges the importance of intent or interest in his later work, “Philosophical Investigations” published in 1953 [Wittgenstein 1953].

Intent is what manager's knowledge is about. How to find and combine the interests of individuals so that their knowledge can be combined (amplified) into new products. In fact Numata argues managers and engineers have the same 5-step development cycle but the elements they work with are different (table 5.2). Note

MANAGER	ENGINEER
Investigating information about engineers	Investigating information about technology
Select engineers	Select technologies
Coordinate communication between selected engineers	Converge selected technologies
Harmonize among team members	Adjust inter-technologies and form as one system
Observe actual operation	Verify with prototyping

Table 5.2: Engineering Management vs. Engineering [Numata 1998]

the “harmonize among team members” step makes perfect sense if Belbin's teamwork theory is applied. Web based communication technologies are proposed to relieve the barriers from social as well as physical distances, which is very much in line with the role computers can play to support creative processes. [Numata 1998].

The work of Numata and Maeda confirms the effectiveness of the combination of creativity- and teamwork-theory, psychology and philosophy, in how to organize and manage a creative problem solving organization as put forward in previous paragraphs. Yet their starting point has been rather different. Numata and Maeda have based their paper on the theory of organizational knowledge creation of Nonaka and Takeuchi, which they claim to be very much founded on Japanese culture and tradition as opposed to western traditions in science and philosophy [Nonaka 1995]. In the theory of Nonaka much attention is paid to the difference between tacit knowledge and explicit knowledge and the importance of tacit knowledge. Organizational knowledge creation is the result of any form of communication aimed at either transport or transformation of knowledge. Being able to communicate tacit knowledge is considered a key element in knowledge creation [Nonaka 1995], which is in many ways similar to the transition from true thoughts to logical images and finally to language as in Wittgenstein's Tractatus [Wittgenstein 1922].

5.6 A general communication model

Combining the ideas of Numata on communication [Numata 1998] with the communication channel concept of Shannon [Shannon 1948], we can define the following communication model, suitable for both marketing type of communication as well as directive control (Figure 5.3).

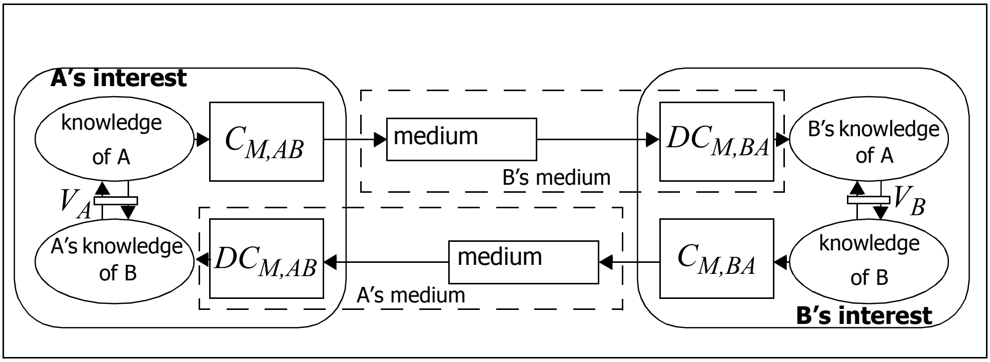


Figure 5.3: Communication Model

Suppose A has some knowledge and has an interest to communicate some information to B. To make sure that B will be able to receive this information, the information has to be coded for a medium M that B can understand and that is in line with A's interest to communicate to B. The medium also has to be suitable for the type of information communicated. This means A's perception of B's value system and interest influences both the choice of the medium and the coding $C_{M,AB}$. In this notation AB indicates the coding is done by A for B and BA vice versa. B can decode the message from A ($DC_{M,BA}$) which leads to a an interpretation of B about the information from A. Accepting this piece of information as a truth or a fact resembling A's knowledge and interest that triggered the communication process, requires another transformation, driven by the value system of B (V_B). This value system may convert B's interpretation of A's message into a piece of knowledge which can trigger a new message from B to A. This closes the communication cycle. There is no guarantee that the cycle will be closed. The value system may also consider the new piece of B's knowledge about A irrelevant and not worthwhile to react. In this case the value system of B, which also drives his interest to communicate has no incentives to communicate to A. The decoding that B does is specific for the medium and for A. Knowledge of the characteristics of the medium, as well as knowledge about A's interest is essential to do a correct decoding of the message.

In practice this means that every communicator should realize the following. As a first guess A may start with the following assumption: $C_{M,AB} \bullet DC_{M,BA} = I$, which simply means A's understanding of B, which A uses for decoding messages from B, is the inverse of the decoding that B uses for messages from A[1]. However, the medium is never loss free and coding and decoding functions in one-way communication are driven by different knowledge sources and value systems. Under these conditions, we cannot assume communication to be perfect. There is no guarantee that the knowledge communicated by A is understood 100% correctly by B: $C_{M,AB} \bullet DC_{M,BA} \neq I$.

At best A may assume: $C_{M,AB} \bullet DC_{M,AB} = I$. This means that A is at least consistent in his application of knowledge on B for coding and decoding messages in relation to B.

The same applies for the feedback. If we take the limitations of the medium for granted, the condition for a minimum loss communication channel becomes: $C_{M,AB} \bullet DC_{M,BA} \approx I$ and $C_{M,BA} \bullet DC_{M,AB} \approx I$. A’s knowledge of B, becomes part of A’s knowledge after it is filtered by A’s value system. This filtered knowledge of A about B also drives A’s ability to code messages for B. The same applies for B with respect to A. This means that for implementing minimum loss communication both ways, a common interest or common intent and adjusted or shared value systems are needed.

---

1. I is the identity transform.

Interest and value system are closely related. Interest is a momentary expression of the value system. Thus for a minimum loss communication process, shared value systems are important, since interest and value systems drive the coding and decoding steps. In practice the communication cycle may have to be executed a number of times to adjust or callibrate both coding and decoding processes and to adjust both value systems to reach a common interpretation (=shared or communicated knowledge). For innovation processes these adjustment cycles may be repeated at regular intervals since the innovation context may change as a result of progress in time and choices made in the process.

The choice of medium is also an important factor. Although losses can never be avoided, the capacity must be sufficient, the medium must be suitable for the receiver and most important the medium must match the intent and the available knowledge representation of the sender.

5.7 Implementing organisational change

The practice of implementing changes in strategy and structure today is still a complex, hardly manageable process. If one says structure follows strategy, discussions on new strategies are often frustrated with constraints from the existing structure. In other words, the uncertainties that result from changing parts of structures that are not in line with the new strategy generate forces against these strategies. On the level of the individual worker, this is no surprise. The announcement of reorganisations in terms of lay-offs is throwing them back to the social and safety level in Maslow's hierarchy of needs [Maslow 1998]. Some change managers may call this the "burning platform" that will unfreeze the existing structure of the organisation. Labour Unions often call for detailed information about the numbers of lay-offs, so that people know where they stand. The problem is that this information does not reduce complexity. For the individual, the number of lay-offs will change the restructuring process into a casino game with job security at stake. Thus the event of announcing restructuring and lay-offs, will generate as many extra states as the number of options the workers like to think of for themselves. Even without the number of lay-offs people feel insecure at their social level, because they cannot see their options and thus they cannot evaluate if their personal job situation will improve or deteriorate.

In his chapter on change management Daft describes the work of Lewin [Lewin 1951], so called force field analysis to find the benefits of change vs. resistance [Daft 1997]. The analysis of the force field gives a better understanding of where management attention is badly needed to overcome the forces of resistance. Morgan describes a management paradox, which states that any force in the direction of change will also generate an opposite force and that the real management challenge is not to choose for either direction, but to balance these forces with strategy and customer demands [Morgan 1996, p268, p291-295]. The balanced set of forces is basically an attractor. For changing structures Morgan uses the metaphor of the Lorentz attractors. An existing organisation has an attractor which locks the organisation in its existing structure, patterns of interaction, ideas about the market, its customers etc. Changing means moving the system to the edge of this attractor and providing a new attractor. This metaphor has much in common with the organisational development approach, which recognises the phases of unfreezing, changing and refreezing [Daft 1997].

The practice with most reorganisation processes is that the unfreezing and intervention is carried out by top management, with or without the help of consultants in the role of change manager. So the organisation is driven to the state where no doubt, changes will occur. But if the attractor is not communicated or simply does not exist, the organisation will most likely fall back to the old attractor but with the loss of the best people.

If the new processes and structure were available at the time of unfreezing, the organisation would have an attractor available for freezing in the new situation. The gain in reducing complexity here is to provide a future set of processes and states that is sufficiently detailed for individual workers to recognise their own future perspective. Most of these workers and especially the best people will not generate the extra states and the extra uncertainty that will cause the change process to become unmanageable. Instead they will comply and start working towards implementing the new processes and structure. The process design approach and the organ structuring method presented in chapter 2 can be used to develop such an attractor image [Meijer 2002] .

5.8 Conclusions

The work of Hohn, Belbin, Maslow and Wittgenstein has been addressed to make clear what managers need to know in order to manage creative and/or problem solving processes effectively. It is argued that the actual synthesis or knowledge fusion necessary for problem solving, only happens in the minds of humans. Unless they communicate their knowledge in some form, no one is going to benefit from their newly gained insight. It is the manager's task to provide the technical and social context in which the team members are inspired to be creative and feel free to communicate. Belbin’s teamwork theory was introduced as a means for specifying and selecting the social traits needed for the team to work effectively.

In creative processes as well as in decision making, information and communication are key aspects. With respect to information, your team cannot be overloaded, provided the team members have the knowledge and the skills to oversee what they are supposed to achieve. With respect to communication, a manager has to understand that language is a distorted projection of concepts and images that only exist in the minds of an individual. Thus, mis-communication is in the nature of language unless one recognizes the interests and past experiences of the individual that inspired these concepts, criteria and opinions. A manager can effectively facilitate communication by offering more than just verbal communication means and by transforming a discussion over solutions into a discussion over mindsets (concepts) and criteria. The works of Nonaka and Numata, show that these ideas are already applied in Japanese industries.

Combining the work of Numata with the channel communication model of Shannon has resulted in a general communication model where the coding and decoding transformation as well as value-systems that act as filters have been made explicit. This model has been used to demonstrate the importance of communicating value-systems in multi-company collaboration and innovation [Meijer 2004] as well as to give clues about the choice of communication media depending on the amount of control over partners in innovation [Voûte 2003].

In implementing organisational change Maslow's hierarchy of needs is important. Creating uncertainty through announcing layoffs, generally throws workers back to the level of social needs or even their needs for safety. A lot of resistance is the natural result. A clear vision on the future strategy, associated with new process designs and a structure supporting the strategy, can lower this resistance. The majority of workers will recognize their personal options in the new organisation and will start working towards calling these options and implementing the new strategy.

6. CONCLUSIONS AND FUTURE RESEARCH

"By starting with the design of business processes and structuring these with the intent to facilitate quality management (customer value) and to optimize productivity, a structure can be found that serves as a basis for the design of a department and management structure. Thus an implicit and natural match is accomplished between the quality control and management requirements from the market and the management and control capabilities of managers and directors." This principle still holds for the design of organization structures that don't induce unneccessary complexity. So the problem shifts to what processes to design and how to structure these processes. Being able to identify one or more flows is the basis for the processes to be designed. For productive and less complex processes, process requirements should be kept as homogeneous as possible. This is the first of five structure design rules. The second rule then states keep related control functions together. For most processes of operations, applying these rules already gives a blue-print for the organization structure. If the flow is not clear, as in research or innovation processes, then evolutionary problem solving processes are proposed. But even then, the structure design rules remain valid, to structure research organizations for better complexity handling.

6.1 Conclusions

In the world of managers the word complex is often a synonym for difficult, complicated, involving many factors and highly uncertain. A complex business decision requires careful preparation of the managers and workers involved. The preparation reduces the uncertainty and reveals the structure of the problems and processes to be dealt with. In this interpretation, complexity is a measure of the effort deemed necessary to resolve the uncertainty and solve the problem. The complexity of business problems is driven by uncertainty and diversity of- and interrelations between- aspects of these problems. Shannon's entropic measure of complexity provides a suitable model for this type of complexity.

Operating business processes requires dealing with a sustained stream of issues and problems in order to create the customer value. The organisational infrastructure can support as well as frustrate the efforts of the workers to deal with their part of the complexity. The research presented in this thesis is aimed at the development of methods for the design of organisations that can handle complexity more effectively and efficiently.

Starting point for this research was the observation that the Delft School of Organization design, founded by Prof. Jan In't Veld and Prof. Pierre Malotaux, was based on a still unique doctrine about organization design:
By starting with the design of business processes and structuring these with the intent to facilitate quality management (customer value) and to optimize productivity, a structure can be found that serves as a basis for the design of a department and management structure. Thus an implicit and natural match is accomplished between the quality control and management requirements from the market and the management and control capabilities of managers and directors.

In the context of designing organizations that can deal with complexity, this doctrine has two important consequences:

1. *The match between process management requirements and management capabilities prevents many of the induced uncertainties and unnecessary coordination leadtime that could frustrate workers and managers if this match is not accomplished.*
2. *Organization structure design starts with process design; this means organization structure designs are specific for each company and their strategies.*

The first consequence avoids complexity that is not related to the creation of (customer) value. The second consequence implies that there are no universal organization structures for dealing with complexity, but there may be common design principles for the structure of processes.
With business process design as a starting point and by qualitatively looking at the sources of uncertainty the following five structure design rules were developed to support structure design decisions:

1. *Do not combine value propositions that are too far apart into one process.*
2. *Identify organs that can be responsible for a distinct contribution, but do not cut important control cycles.*
3. *While maintaining 1 and 2, try to create economies of scale (efficiency).*
4. *While maintaining 1 and 2, try to achieve some flexibility and reduce the vulnerability associated with small departments and product oriented structures by merging these (support) functions into larger multi-service units.*
5. *If after 1-4 value propositions are left that do not add value to the company, outsource them or give up these lines of business.*

The first rule prevents uncertainties induced by orders from different markets competing for the same resources. Diversity from two or more stochastic sources mixed into one source may even become a bigger source of complexity than the sum of these sources. The second rule prevents uncertainties induced by spreading the responsibility for controls over different managers. The third rule supports efficiency but may require investments in advanced planning and control to prevent reintroducing the uncertainties avoided by rules 1 and 2. The fourth rule reduces uncertainty over the availability of (human) resources. The fifth rule promotes a strategic focus. Strategic focus or limiting yourself to doing what you are good at is in itself an effective means to reduce complexity.

The importance of these design rules is that they offer a sequence for taking structure design decisions. Especially not taking any decisions over organizing for economy of scale (rule 3 and 4) before the results from the first and the second rule are known, allows organizations to weigh the benefits of economies of scale against the additional investments in scheduling necessary to maintain customer value attributes. If the market is paying well for speed and agility, then economy of scale, which is basically an efficiency goal, is not the first priority. Organizations that start their design with the economy of scale rules often violate the principles of the first rule. For these organization typically 20% of the customers are responsible for 100% of the profit (if any). The other 80% are just eating capacity at no margin.

The structure design rules can be applied to all types of business processes; processes of operations as well as processes of development and innovation. Yet there are also important differences between these processes that influence how to apply the design rules. These differences and consequences will be discussed in more detail in the following paragraphs.

6.1.1 Structuring processes of operations

In operations, a flow of material, information or resources is easy to recognize and often this flow coincides with the primary interests of the customer. Organizational problems for this type of process are often caused by allowing customers to order the product or service with too many variants and in a wide range of quantities, serviced by a single process. Sometimes this situation is the result of organizing for economy of scale with insufficient attention for leadtime performance. The solution is to create multiple processes that can operate concurrently. The productivity for each process can be optimized for a more homogeneous set of customer requirements. Quantitative queue-server simulations have indicated that if two job streams are mixed that represent an equal average load but who's average job-sizes and average job-frequencies are a factor four or more apart, a single queue-server is not capable of achieving an acceptable leadtime performance for both streams. The leadtime performance is dominated by the stream with the largest jobs. Adding more capacity to the server certainly helps but results in an efficiency of less than 50% before the stream with the smallest and more frequent jobs is serviced acceptably.

If the processes of operations are too big to be considered as one organ, then it is important not to spread important controls over two or more organs. Usually order entry points or inventory positions are natural places to be considered as organ boundary.

6.1.2 Structuring processes of development and innovation.

Processes for redesign or variant design rely very much on reuse and learning from the past for their efficiency. The structure of the product and the information architecture are known and not subject to change. This means these structures can serve as a reference for process design as well as for the organisation structure.

However, special attention is needed for the interfaces between physical parts and systems based on different technologies. In aircraft as well as automotive design, parts of different subsystems are often operating in, and competing for the same physical space. Design decisions related to these systems become part of one control related to the allocation of physical space in that area. The second structure design rule prescribes to keep these decisions together in one organ. Some car manufacturers have a special department for the design of doors. A door is a structural part of the car body that is largely responsible for the side-impact crash resistance. At the same time it holds ventilation systems, audio systems and electrical systems for windscreen operations, locks and rear view mirrors. The interior of the door is also influenced by car interior designers that may have reduced the available space in the door for an integrated armrest and storage bins.

As a general rule, development processes should be structured around complex multiple interface decisions. By defining departments for these "design-areas", rather then project teams, learning and standardization of solutions can be enhanced, which has a positive influence on the manufacturing cost.

For development and innovation processes where these interfaces, interactions and complex controls are not clear, self organisation principles and evolutionary development processes need to be applied. Evolutionary problem solving processes also have their use within departments that deal with the complex design areas as discussed above.

The evolutionary development processes also solve a social problem that occurs in large hierarchical development organisations. In a hierarchical design organisation, social dominance of senior workers may cause new ideas or early problems spotted by young engineers without a reputation to be neglected. Even worse, a desire to maintain the balance of power among seniors, may cause group-think among them. If ideas and problems are passed on anonymously as part of team documentation intended for learning and copying by other teams, then the chance of survival of good and novel ideas is much better.

6.2 Tools for defining process requirements

Designing processes and structuring them for effective and efficient operations cannot deal with complexity if the process requirements or process constraints are still uncertain. The customer value mix of Kemperman and Free Cashflow Methods for creating business value were selected and described in this thesis to help organizations to develop better performance indicators, with less uncertainty about cashflow and profit drivers for their business. These performance indicators provide the most important ingredients for business process design and the first and second structure design rule.

6.3 Human factors in organization structure design

The structuring principles presented in this thesis are all based on principles from mathematics or physics that have been proven effective for technical system design. For these ideas to work for organizations that involve humans as sensors and actors, it is necessary to study if and under what conditions, humans will comply with the processes and structures that result from a technical system design approach. If such conditions exist then we can safely treat socially induced disturbances as a boundary condition to be treated separately from the organization structure design. In chapter 5 Maslow's hierarchy of needs, Belbin's teamwork theory as well as some theories for creative problem solving are studied to identify conditions under which human's will generally comply with processes and an organizational context that is based on technical system design methods. Moreover, with Maslow in mind it is argued that such processes and structures are essential communication means for a change or reorganization process. Understanding human communication is the key. For this purpose a model of a human to human communication cycle has been developed. The most important quality of this model is that it distinguished the functions of coding, decoding and acceptance (values) in person to person communication. Although the communication model is not a quantitative model, making these functions explicit helps managers, engineers and other workers to identify the sources of uncertainty and mis-communication. Supported by knowledge about language, culture and value-systems humans can try to focus their communication on those parts that may cause disagreement or misunderstanding.

The communication model was also used as a basis for the coordination cost model of chapter 3. For hierarchical organizations with only formal communications between departments, coordination between departments 4 steps apart and communications between them involving three layers of hierarchy may become over 20 times more expensive than the cost of coordination within a department. Thus it is no surprise that such coordination does not take place or is more likely to be handled by the informal organization. However, in case of clear company goals, transparent policies and management incentives at department level that are aligned with the company goals, the coordination cost model reduces to a linear model. For this model, the coordination cost between departments four steps apart is between 9 and 12 times more expensive. The coordination cost model represents another argument for limiting the coordination needs between departments, through building departments that are largely self-supporting for the services they provide and through organizing the coordinated through a transparent and coherent set of boundary conditions that define the decision freedom and authority of each department.

6.4 Future research

A research topic like organisation structures for dealing with complexity is never finished. The issues addressed in this thesis help managers and researchers to make progress with their own ideas. At the same time when working on models and executing simulation studies as well as being involved in organisation design processes in various companies new problems surface and new ideas come forward.

The queue-server simulations provided some practical answers, but at the same time it became clear that the theoretical background for predicting maximum and minimum leadtime values is not complete. Also testing random number generators and studying the influences on statistical properties of conversions between continuous time descriptions and a discrete time simulation were kept outside the scope of this thesis. It was assumed that the standard statistical libraries of MAPLE are sufficient. The simulation results presented in this thesis show the right trends. But the numerical quality of some results especially under high and low loads is still limited. Of cause these are also the areas where numerical stability is often weak.

But it is believed that improvements are possible through improving the statistical modelling and by tests on other simulation platforms such as Tomas [Veeke 2003].

The evolutionary organization for design processes is still a qualitative model. To really put it to practice quantitative models need to be developed as well, to be more specific about the relationship between the number of teams, the number of engineers per team, the complexity of the design problem and the required leadtime per design phase. In addition field tests of the coordination cost model and Glickstein's optimal span theory can improve our ability to build more effective research organizations.

The evolutionary development model could also be expanded to an organization model for large public funded research programs. Without the communication infrastructure of the evolutionary development model, the only cross fertilization and synergy effects that can be expected take place within existing research groups and not between them. Without the cross fertilization and synergy effects between research groups, creating coordinated focused research agenda's for these public funds will prove counter productive. The coordination and focus often kills the diversity of ideas. Rather than breakthrough, such a program will only deliver losely coupled, non-integrated results.
More specifically, European Commission funded Networks of Excellence (NOE) can be expected to become much more effective if the organisation can copy ideas from the concept of the evolutionary development organisation and if the goals for these networks are being defined in terms of results that will sustain and expand the network (e.g. proposals for integrated projects, coordination actions and specific targeted research projects) rather than hours spend, reports written and infrastructure created but not really used because research is not funded under the NOE-contract.

Author Index & References

A

Arbnor I. 21, 238

Arbnor I., Bjerke B., *Methodology for creating business knowledge*, Sage publications, London UK, 1996.

B

Banzhaf W. 164, 165

Banzhaf W., Nordin P., Keller R.E., Francone F.D., *Genetic Programming, an introduction*, Morgan Kaufmann Publishers, 1998, ISBN 1-55860-510-X.

Belbin M. 179

Belbin M.,*Team roles at work*, Butterworth-Heinemann Ltd., Oxford, 1993.

Bikker H. 43, 44, 45, 48, 153, 252

Bikker H. ,*Lecture notes of Wb5413 Analyses and Design of Production Organizations*, Mechanical Engineering, Delft University of Technology, English edition, 2001.

Bikker H., "The challenge of managing operations and innovations", *Seminar: Future Development in Industrial Management*, European Academy for Industrial Management (AIM), Bochum, Germany, October, 2000.

Bikker H., Till R.H. van, "Productfamilies; the paradigm of the order entry matrix", *4th WDK Workshop on product structuring*, Delft University of Technology, October 1998.

Bikker H.,*Lecture notes of Wb5413 Analyzes and Design of Production Organizations (in Dutch)*, Mechanical Engineering, Delft University of Technology, 1994.

Bikker, H., Dekkers R., "The order entry matrix: Keystone for productivity of design and manufacturing", *2nd Intl. product development management conf. on new approaches to development and engineering*, Gothenburg, Sweden, 1994b.

Blommestijn M. 112

Blommestijn M., *A probability distribution sensitivety study of the information theoretic complexity measure*, Small Research Assignment, Industrial Organisation, Delft University of Technology, 2002.

Boswijk H.K., *Complexity in evolutionary and organizational perspective*, University Press, Erasmus University Rotterdam, 1992.

Brooks F.P., *The mythical man-month (anniversary edition)*,Addison Wesley Longman 1995, ISBN 0-201-83595-9.

Bruijn J.A. de, Heuvelhof E.F. ten, in't Veld R.J., *Procesmanagement: about procesmanagement and decision making (in Dutch)*, Kluwer 2nd.edition 2002, ISBN 90-5261-397-4.

Burch R.,"Charles Sanders Peirce", *The Stanford Encyclopedia of Philosophy*, Fall 2001 edition, URL=http://plato.stanford.edu/archives/fall2001/entries/peirce

C

Checkland P.B.,*Systems Thinking, Systems Practice*, John Wiley & Sons 1993, ISBN 0-471-27911-0.

Clark P.A. ,*Action Research & Organizational Change*, Harper&Row, London, 1972, ISBN 06-3180004-9.

Craig S.C., Douglas S.P., "Responding to the challenges of global markets: change complexity, competition and conscience", *IEEE Engineering and Management Review*, Fall 1997, pp.4-14.

D

Daft R.L., *Management*, 4th edition, The Dryden Press, 1997, ISBN 0-03-27989-0.

Degani A., Wiener E., "Procedures in Complex Systems: The Airline Cockpit", *IEEE Transactions on Systems, Man and Cybernetics*, Vol.27 No.3, May 1997.

Dempster A.P. 14

Dempster A.P., "A generalisation of Bayesian inference (with discussion).", *Journal of the Royal Statistical Society B*, 30:205-232, 1968.

DeTardo-Bora K.A. 250

DeTardo-Bora K.A., "Action Research in a world of positivist oriented review boards", *Action Research*, Vol.2,No.3, Sage publications, 2004.

Dijk M.M. 252

Dijk M.M., *Functional Process Benchmarking between product development organisations in Jaguar and Volvo*, Master-thesis WMBT/IO/428-II, Mechanical Engineering, Delft University of Technology, 2001.

Dijk N. van 109

Dijk N. van, "Always in the wrong queue" (in Dutch), *Natuur en Techniek*, Vol.64 No.12, December 1996.

Dill D.D. 148

Dill D.D., Pearson A.W., "The self designing organization: structure, learning, and the management of technical professionals", *IEEE Conference on Technology Management*, 1991, ISBN 0-7803-0161-7, pp.33-36.

E

Easyjet 20

Easyjet, *Annual Report and Accounts 2005*, http://www.easyjet.com/EN/Investor/investorrelations_financialreports.html.

Ellis J.H.M. 250

Ellis J.H.M., Kiely J.A.,"The promise of action inquiry in tackling organisational problems in real time", *Action Research International, paper 5.* Available on line: http://www.scu.edu.au/schools/gcm/ar/ari/p-jellis00.html

F

Frey D.D. 24, 254

Frey D.D., Dym C.L., "Validation of design methods: lessons from medicine", *Research in Engineering Design*, (2006) 17: 45-57, Springer.

Frizelle G. 95

Frizelle G., Woodstock E., "Measuring complexity as an aid to developing operational strategy", *International Journal of Operations and Management*, Vol.15,No.5, 1995, pp.26-39.

G

Gershenson C. 161

Gershenson C., Heylighen F., "When can we call a system self organizing", *Lecture notes in computer science*, vol.2801, December 2003, pp.606-614.

Glickstein I.S. 6, 104, 106, 107

Glickstein I.S., *Hierarchy theory: some common properties of competitively selected systems*, PhD-thesis, State University of New York at Binghamton, 1996, UMI-963423.

Glickstein I.S., *Quantifying the "Mythical Man-Month*,rev.4.7 March 2003, http://polaris.umuc.edu/~iglickst/mswe601/optimal_MMM.doc .

Grant R.M. 58

Grant R.M., *Contemporary Strategy Analysis - concepts, techniques, applications.* Fourth edition, Blackwell publishing, 2002.

H

Heylighen F. 161, 222

Heylighen F. Gershenson C.,"The meaning of selforganization in computing", *IEEE Inteligent Systems*, July/August, 2003, pp.72-75.

Heylighen F., Joslyn C. and Turchin V. (editors), *Principia Cybernetica Web (Principia Cybernetica, Brussels)*, 2-12-1993, http://pespmc1.vub.ac.be/ZESUGAM.html .

Hogan N. 173

Hogan N., "Rehabilitation Applications of Robotic Technology", guest editorial, *Journal of Rehabilitation Research and Development*, Nov.-Dec. 2000, Vol.37, No.6.

Hohn H.D. 176

Hohn H.D.,*Playing, Leadership and Team Development in Innovative Teams*, PhD-thesis, Delft University of Technology, 1999.

Holland J.H. 11, 164

Holland J.H., *Adaptation in Natural and Artificial Systems*, University of Michigan Press, June 1975, ISBN 0-472-08460-7.

Holland J.H., *Emergence, from chaos to order*, Helix Books, Cambridge Mass., 1998.

Hopp W.J. 111, 112

Hopp W.J., Spearman M.J., *Factory Physics: foundations of manufacturing management*, McGraw-Hill, New York, 2nd edition, 2001, ISBN 0-256-24795-1.

Horgan J.,"From Complexity to perplexity", *Scientific American*, June 1995, pp.74-79.

Hubka V., Eder E., *Design Science*, London, Springer, 1996, ISBN 3-540-19997-7.

I

In't Veld J., *Organisation Structure and Job definition (in Dutch)*, Elsevier, Amsterdam, 1981, ISBN 90-207-2313-8.

In't Veld J.,*Analyses of Organisation Problems (in Dutch)*, Wolters Noordhoff publishers, 8th revised edition (1st edition 1975), 2002, ISBN 90-207-3065-7.

K

Kaplan R.S., Norton D.P. ,"The Balanced Scorecard - measures that drive performance", *Harvard Business Review*, Jan-Feb 1992, pp.71-79.

Katona G.O.H., Nemetz T.O.H., "Huffman codes and Self-Information", *IEEE Transactions on Information Theory*, Vol. IT-22, No.3, May 1976, pp.337-340.

Kemperman J.E.B., Engelen J.M.L. van, "Operationalizing the customer value concept", *Competitive paper 28th EMAC Conference Marketing and Competition in the information age*, Berlin 11th-14th May 1999.

Koolhaas J.W., *Organization, dissonance and change*, IIB Delft University of Technology, 1980.

Kreyszig E., *Introductory mathematical statistics; principles and methods*, John Wiley&Sons, 1970, ISBN 0-471-50735-0.

L

Lambert D., Slater S.,"First, Fast and On-time. The path to success, or is it ?", *Proceedings of IEEE-IEMC 1996*, Vancouver, 1996, ISBN 0-7803-3552-X , pp.275-280.

Lewin K., *Field theory in Social Science: selected theoretical papers*, New York, Harper&Brothers, 1951.

Lubbe J.C.A. van der, *Information theory*, Cambridge University Press, 1997, ISBN 0-521-46760-8.

M

Mackor J.A.,*Development and application of a process model for Supplier Relationship Management for Siemens AG*, Master-thesis IO 03-063, Mechanical Engineering, Delft University of Technology, 2003.

Malotaux P.Ch.A., *Job Evaluation: Complexity the dominant factor*, Malotaux & Partners b.v., Bilthoven the Netherlands, 1977.

Maslow A.,*Maslow on management*, John Wiley & Sons Inc., 1998, ISBN 0-471-24780-4.

Maturana H.R.,Varela F.J., "Autopoiesis and Cognition", *Boston studies in the philosophy of science*, Vol.42, Reidel Publishing Company, 1980, ISBN 90-277-1015-5.

Meijer B.R. "From reducing complexity to adaptive organizations", *Proceedings of IEEE-IEMC 2002*, Cambridge UK, 18-20 August 2002, ISBN 0-7803-7385-5, pp.661-666.

Meijer B.R. e.a.,"Communicating context and strategy for collaborative design in networks and corporations",*Meth. and Tools for co-operative and integrated design*, edt. S. Tichkiewitch e.a., Kluwer 2004, ISBN 1-4020-1889-4

Meijer B.R., "To manage or not to manage complexity", *Proceedings of IEEE-IEMC 1998*, San Juan Puerto Rico, 1998, ISBN 0-7803-5-82-0, pp.279-284.

N

O

Osborn A.F. 177

Osborn A.F.,*Applied imagination, principles and procedures of creative problem-solving*, Charles Scribner's Sons., New York, 1963.

P

Pahl G. 46, 147

Pahl G., Beitz W., *Engineering Design; a systematic approach*, 2nd edt. Translated by Wallace K., London, Springer, 1996, ISBN 3-540-19917-9.

Pall G.A. 52

Pall G.A., *Process-Centered Enterprise; the power of commitments*, St. Lucie Press, 2000, ISBN 1-57444-239-2.

Parsons S. 15

Parsons S., Hunter A., "A review of uncertainty handling formalisms", *Lecture notes in computer science*, Vol. 1455, Springer-Verlag, Heidelberg, 1998.

Peirce C.S. 234

Peirce C.S.,"Lectures on Pragmatism", in *collected papers CP 5.14-5.212*, edt. by C. Hartshorne, P. Weiss, Harvard Univ.Press, 1931-35.

Peppel F. van de 77

Peppel F. van de, *Production Organisation and Control*, Master thesis, Industrial Organisation, WBMT/IO/97/106, Delft University of Technology, 1997.

Prahalad C.K. 1

Prahalad C.K., Hamal G.,"The core competence in the corporation", *Harvard Business Review*, 1990. (May-June),pp.79-91.

Prigogine I. 11

Prigogine I.,*The end of certainty*, Free Press, 1997, ISBN 0-684-83705-6.

R

Rhodes M. 177

Rhodes M.,"An analysis of creativity", *Phi Delta Kappan*, Vol.42, 1961, pp.305-310.

Rijn D. van 58

Rijn D. van, *Corporate Value Creation as a company goal*, Literature thesis, Production Technology and Organisation, Mechanical Engineering, Delft University of Technology, 1999.

Royal Swedish Academy of Science 243, 249

Royal Swedish Academy of Science, "Foundations of behavioral and experimental economics:Daniel Kahneman and Vernon Smith", *advanced information on the Nobel prize in economy 2002*, www.kva.se .

Royal Swedish Academy of Science, "The scientific contributions of James Heckman and Daniel McFadden", *advanced information on the Nobel prize in economy 2000*, www.kva.se .

S

Sanders M.S. 173

Sanders M.S., McCormick E.J.,*Human Factors in Engineering and Design*, McGraw-Hill, 7th edt.,1993, ISBN 0-07-054901-X.

Saridis G.N. 66

Saridis G.N., "Analytical formulation of the principle of Increasing Precision with Decreasing Intelligence for intelligent machines.", *Automatica*, Vol 25, No.3, May 1989, pp. 461-467.

Saridis G.N., "Analytical formulation of the principle of increasing precision with decreasing intelligence for intelligent machines.", *Proceedings of the IFAC 1988 International Symposium on Robot Control*, Karlsruhe 1988.

Saunders M. 21, 241, 244

Saunders M. et al., *Research methods for business students*, Pearson Education, London, 2nd edition, 2000.

Schafer G. 14

Schafer G., *A mathematical theory of evidence*, Princeton University Press, Princeton NJ, 1976.

Schurz G. 234, 235, 237, 238

Schurz G.,"Laws of Nature versus System Laws", *PE Preprints*, Annual 2001, No.5, ed. von der Pfordten D. & Schurz G., Philosophical prepublication series at the University of Erfurt.

Schurz G.,"Laws of Nature versus System Laws", *PE Preprints*, Annual 2001, no.5, ed. von der Pfordten D. & Schurz G., Philosophical prepublication series at the University of Erfurt.

Schurz G.,"Models for Abductive reasoning", *TPD Preprints*, Annual 2002, no.1, ed. Schurz G. & Werning M., Theoretische Philosophie an der Universität Düsseldorf.

Schwartz M. 223, 225, 228, 229

Schwartz M., *Telecommunication Networks*, Addison-Wesley, 1987, ISBN 0-201-16423-X.

Shannon C.E."A mathematical theory of communication", *The Bell System Technical Jnl.*, Vol.27,pp.379-423,623-656, July/October 1948.

Shy O., *Industrial Organization* , MIT-press, 1995, ISBN 0-262-69179-5.

Sivadasan S., Efstathiou J., e.a., "Advances on measuring the operational complexity of supplier-customer systems", *European Journal of Operational Research*, Springer, 2004.

Sivadasan S., *Operational Complexity of Supplier Customer Systems*, PhD-thesis Department of Engineering Science, Oxford University, 2001.

Smets P., *Imperfect information : Imprecision - Uncertainty*, http://iridia.ulb.ac.be/~psmets/Imperfect_Data.pdf, 1999.

Smithson M., *Ignorance and Uncertainty: Emerging Paradigms*, Springer-Verlag, New York, 1989.

Sobek II D.K, Ward A.C., "Principles from Toyota’s Set-based Concurrent Engineering Process", *Proceedings of ASME 96-DETC/DTM-1510.*

Stacey R.D., *Complexity and creativity in organisations*, Berrett-Koehler Publishers, 1996., ISBN 1-881052-89-3.

Suh N.P., *Axiomatic Design: Advances and Applications*, Oxford University Press, New York, 2001, ISBN 0-19-513466-4.

Suh N.P., *The principles of Design*, Oxford University Press, New York, 1990, ISBN 0-19-504345-6.

Suh N.P.,*Complexity,Theory and Applications*, Oxford University Press, Oxford, New York, 2005, ISBN 0-19-517876-9.

Swink M.L., Cantalone R., "Design-Manufacturing Integration as a Mediator of Antecedents to New Product Design Quality", *IEEE Transactions on Engineering Management*, Vol.51,No.4, November 2004.

Szurovy G. 20

Szurovy G., "Safety first: FLS Aerospace's Prime Directive", *Aviation Today*, November 1, 1999.

T

Tietge M.A. & Vughs A.W.A.M. 75

Tietge M.A, Vughs A.W.A.M., *Key-Performance Indicators for Productivity Support (in Dutch)*,Internship report, Industrial Organisation, Delft University of Technology, 1995.

Tomiyama T. 236

Tomiyama T., Takeda H, Yoshioka M., Shimomura Y., "Abduction for creative design", *Proceedings of DETC03, DETC2003/DTM-48650*, ASME 2003.

V

Veeke H.P.M. 38, 116, 200

Veeke H.P.M.,*Simulation Integrated Design for Logistics*, PhD-thesis, Delft University of Technology, DUP Science, 2003, ISBN 90-407-2417-2.

Velde L.N.J. van der 252

Velde L.N.J. van der, Meijer B.R., "A system approach to supply chain design with a multinational for colorants & coatings", *Tackling Industrial Complexity*, edt. Frizelle G., Cambridge UK 2002, ISBN 1-902546-24-5.

Velde L.N.J. van der & Schuurman J.H. 252

Velde L.N.J. van der & Schuurman J.H., *Supplychain approach*, Master-thesis WBMT/IO/99/109, Mechanical Engineering, Delft University of Technology, 1999.

Visscher K. 23

Visscher K., *Design Methodology in Management Consulting* , PhD-thesis, University of Twente, 2001, ISBN 90-365-1678-1.

Voûte H.J. 155, 192

Voûte H.J., "Perspectives on Innovation; A new TAO for business renewal & development", *Proceedings of IEEE-IEMC 2000*, Albuquerque NM, 13-15 August 2000.

Voûte H.J., Meijer B.R., "Communicating Innovation, coding and control in teams",*Proceedings of IEEE-IEMC2003*, Albany NY.

W

Walsh S., Merges M., Kirchoff B., Morabito J., "Portfolio management for the commercialization of advanced technologies", *Proceedings of IEEE-IEMC 1998*, Puerto Rico, ISBN 0-7803-5-82-0, pp.452-456.

Ward A.C., Liker J.K., Cristiano J.J., Sobek II D.K., "The second Toyota Paradox: How delaying decisions can make better cars faster", *Sloan Management Review*, Vol.36, No.3, pp.43-61.

Wheelwright S.C., Clark K.B., "Creating project plans to focus product development", *Harvard Business Review*, 1992, Mar-Apr;70(2):70-82.

Wiendahl H.-P., Scholtissek P., "Management and Control of Complexity in Manufacturing", *Keynote paper Annals of the CIRP*, Vol.43/2/1994.

Wijk R.W. van,*Design of the processes to support CRM & data management for the Hewlett-Packard Organization*, Master-thesis WBMT/IO/408-II, Mechanical Engineering, Delft University of Technology, 2000.

Wittgenstein L. (1953), *Philosophical Investigations:* the German text, with a revised English translation, Blackwell publishing, 3rd edition, 2001, ISBN 0-631-23159-5.

Wittgenstein L., *Tractatus logico-philosophicus*, original published in 1922, bilingual Dutch-German edition, Athenaeum, 1976, ISBN 90-253-1534-8.

Womack J.P., Jones D.T., *Lean Thinking*, Simon & Schuster, New York, 1996, ISBN 0-684-81035-2.

Womack J.P., Jones D.T., Roos D., *The machine that changed the world*, Harper, New York, 1990, ISBN 0-06-097417-6.

Z

Zadeh L.A., "Fuzzy sets as a basis for a theory of possibility.", *Fuzzy Sets and Systems*, 1:1-28, 1978.

APPENDIX-A SHANNON'S INFORMATION ENTROPY

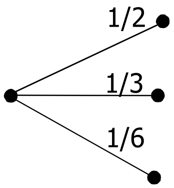
Shannon's information entropy, was originally developed as part of a theory for communication systems. The fundamental problem of communication systems was phrased as reproducing at one point exactly or approximately the message that was selected at the other end of the communication system, regardless of the meaning of that message. Meaning is irrelevant to communication systems since these systems have to accommodate the transmission of messages regardless of their meaning. So the significant aspect is that the actual message is selected from a set of possible messages or in more abstract terms, selecting an event from a set of possible events. This is not a problem, particular to communication systems. Thus a mathematical theory for this class of problems has many more applications, including management problems as selecting a decision from a set of possible decisions. In 1948, Shannon published his paper "A mathematical theory of communication" in the Bell System Technical Journal. This appendix contains a selection of the theorems presented in this paper [Shannon 1948]. Because of clearer explanations, most notations are taken from a textbook on information theory by Jan van der Lubbe [Lubbe 1997].

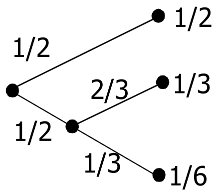
A.1 Choice, uncertainty and entropy

Suppose we have a set of possible events whose probabilities of occurrence are known, $p_1,p_2,.....,p_n$. Can we measure how much "choice" is involved in the selection of the event or how uncertain we are of the outcome? If there is such a measure, $H(p_1,p_2,.....,p_n)$, the following properties are required:

1. *H should be continuous in* p_i.
2. *If all* p_i *are equal,* $p_i = 1/n$, *then* H *should be a monotonic increasing function of n. With equally likely events there is more choice or uncertainty, when there are more possible events.*

3. *If a choice can be broken down into two successive choices, the original H should be the weighted sum of the individual values of H.*





Initially there are 3 possibilities $p_1=1/2$, $p_2=1/3$, $p_3=1/6$, which are decomposed into a choice between possibilities with probability $1/2$, followed by a choice out of two possibilities with probabilities $2/3$ and $1/3$. We require in this case that:
$H(1/2,1/3,1/6)=H(1/2,1/2)+ 1/2*H(2/3,1/3)$

The only H satisfying the three above assumptions is of the form:

$$H = -K\sum_{i=1}^{n} p_i \log p_i$$

where K is a positive constant (unit of measure)[1]. In case of binary units or bits (log base 2), K equals 1. For the remainder of this appendix and in this thesis we assume K equals 1 and log base 2. H is called the entropy of the set of possibilities $(p_1,p_2,p_3,...,p_n)$.

$$H = -\sum_{i=1}^{n} p_i \log p_i \qquad \textbf{(A.1)}$$

A.2 Self-information

A notion not mentioned in the original Shannon paper, but important to this thesis is the notion of self information. Self-information is the unweighted amount of

1. Proof can be found in appendix 2 of the original Shannon paper.

information one receives if an event i occurs that has probability p_i. The self-information of this event is defined as follows [Katona 1976]:

$$Self\text{-}information(p_i) = -\log(p_i) \qquad \textbf{(A.2)}$$

A.3 Marginal probability and Marginal-information

Suppose there are two sets of events X and Y in question with n possibilities for x_i and m possibilities for y_j. Let p_i be the probability of event x_i occurring ($p_i=P(x_i)$) and let q_j be the probability of event y_j occurring ($q_j=P(y_j)$). Let $r(x_i,y_j)$ be the probability of the joint occurrence of x_i and y_j: ($r(x_i,y_j)=r_{ij}=P(x_i,y_j)$). Note in case of independence, $r_{ij}=p_i*q_j$. From the joint probability we can define the marginal probabilities as follows:

$p_i = \sum_{j=1}^{m} r_{ij}$ and $q_j = \sum_{i=1}^{n} r_{ij}$. With (A.1) it now follows that the marginal-information of an event is:

$$Marginal\text{-}information(p_i) = -p_i\log(p_i) \qquad \textbf{(A.3)}$$

Note that the marginal information equals the probability weighted self-information.

A.4 Joint information

With the definition of joint probabilities r_{ij} , the joint information for these sets of events is:

$$H(X, Y) = -\sum_{i,j=1}^{n,m} r(x_i, y_j)\log[r(x_i, y_j)] \qquad \textbf{(A.4)}$$

A.5 Conditional events, conditional-information

Suppose there are two sets of events X and Y, not necessarily independent. Suppose we know that x_i has occurred and we are interested in the probability of y_j. In stead of $q(y_j)$ then have $q(y_j|x_i)$. $\sum_j q\langle y_j|x_i\rangle = 1$ still holds. With (A.1) we can calculate the conditional-information $H\langle Y|x_i\rangle$ as follows:

$H\langle Y|x_i\rangle = -\sum_{j=1}^{m} q\langle y_j|x_i\rangle \log[q\langle y_j|x_i\rangle]$. This result has to be averaged over all values x_i to get the conditional information of set Y, regardless of X.

$$H\langle Y|X\rangle = -\sum_{i=1}^{n} p(x_i)\left(\sum_{j=1}^{m} q\langle y_j|x_i\rangle \log[q\langle y_j|x_i\rangle]\right) =$$

$$-\sum_{i,j=1}^{n,m} p(x_i)q\langle y_j|x_i\rangle \log[q\langle y_j|x_i\rangle]$$

So we can rewrite this result as:

$$H\langle Y|X\rangle = -\sum_{i,j=1}^{n,m} r(x_i, y_j)\log[q\langle y_j|x_i\rangle] \qquad \textbf{(A.5)}$$

and like wise

$$H\langle X|Y\rangle = -\sum_{i,j=1}^{n,m} r(x_i, y_j)\log[p\langle x_i|y_j\rangle] \qquad \textbf{(A.6)}$$

If we substitute the expressions for $p\langle x_i|y_j\rangle = r(x_i, y_j)/q(y_j)$ and $q\langle y_j|x_i\rangle = r(x_i, y_j)/(p(x_i))$, combined with the definition of the marginal probability we get:

$$H\langle Y|X\rangle = -\sum_{i,j} r(x_i, y_j)\log r(x_i, y_j) + \sum_{i,j} r(x_i, y_j)\log q(y_j) = H(X, Y) - H(X)$$

and

$$H\langle X|Y\rangle = -\sum_{i,j} r(x_i, y_j)\log r(x_i, y_j) + \sum_{i,j} r(x_i, y_j)\log p(x_i) = H(X, Y) - H(Y)$$

The uncertainty over Y is never increased by knowledge over X and vice versa. It will be decreased unless X and Y are completely independent, in which case it will stay unchanged. This result leads to two theorems that are used often:

$$H\langle X|Y\rangle = H(X, Y) - H(Y) \qquad \textbf{(A.7)}$$

$$H\langle Y|X\rangle = H(X, Y) - H(X) \qquad \textbf{(A.8)}$$

We may rephrase (A.7) and (A.8) as follows:

$$H(X, Y) = H\langle X|Y\rangle + H(Y) = H\langle Y|X\rangle + H(X) \qquad \textbf{(A.9)}$$

In case of complete independence of X and Y, knowledge of Y does not imply any knowledge on X ($H\langle X|Y\rangle = H(X)$) and it follows that $H(X, Y) = H(X) + H(Y)$. In case of complete dependence of X on Y, X is completely known as soon as Y is known. Knowledge of X does not add to knowledge of Y. Thus the joint information is equal to the marginal information of Y: $H(X, Y) = H(Y)$ and the conditional information is zero ($H\langle X|Y\rangle = 0$). The joint information always satisfies ineqality (A.10), with equality in the case of independence.

$$H(X, Y) \le H(X) + H(Y) \qquad \textbf{(A.10)}$$

APPENDIX-B GAME THEORY

Game theory is quite popular in economics as well as management science for quantitatively modelling and predicting business decisions and business behaviour. In this appendix, the basic principles of game theory are presented.

B.1 Normal form game

In game theory a **normal form game** is defined as follows [Shy 1995, p12-25]:

1. *A set of players whose names are listed in the set* $I \equiv \{1, 2, 3, ... ,N\}$
2. *Each player* $i, i \in I$*, has an action set* A^i *which is the set of all actions available to player* i*. Let* $a^i \in A^i$ *denote a particular action taken by player* i*. Thus player* $i's$ *action set is a list of all actions available to player* i *and hence,* $A^i = \{a^i_1, a^i_2, a^i_3, \ldots, a^i_{k_i}\}$ *where* k_i *is the number of actions available to player* i*.*
   Let $a \equiv (a^1, a^2, \ldots, a^i, \ldots, a^N)$ *be a list of the actions chosen by each player. We call this list of actions chosen by each player* i *an* ***outcome*** *of the game.*
3. *Each player* i *has a payoff function,* π^i*, which assigns a real number,* $\pi^i(a)$*, to every outcome of the game. Formally, each payoff function* π^i *maps an N-dimensional vector,* $a = (a^1, \ldots, a^N)$ *(the action chosen by each player), and assigns it a real number* $\pi^i(a)$*.*

With these definitions, it is possible to span a space of players, their possible actions and the utility function or payoff players may receive if they choose a particular action. In a normal form game the payoff is specified in a matrix.

B.1.1 War-peace game

The war-peace game is a well known example of a normal form game. Suppose a war-peace game is played between two countries, the payoff-functions for both countries can be specified according to table B.1.

		COUNTRY 2			
		WAR		PEACE	
COUNTRY 1	WAR	1	1	3	0
	PEACE	0	3	2	2

Table B.1: War-peace game

This payoff example indicates the following. If both countries play war, the payoff equals 1. If both countries play peace, the payoff equals 2. If one country plays war and the other plays peace, the payoff for the country that plays war equals 3.

B.1.2 Equilibrium outcomes

A game can have many outcomes, but economic theory claims that outcomes that maximize payoff are more likely then outcomes that do not support payoff. In order to make predictions about outcomes, equilibrium concepts are introduced. Equilibrium concepts are conditions or behaviours under which the total set of outcomes can be reduced to one or a few outcomes. To achieve this, one approach is to choose the perspective of one particular player and consider actions that would maximise this player's utility within the context of actions of the other players. One of such equilibrium concepts is the **equilibrium in dominant actions**:

A particular action $\tilde{a}^i \in A^i$ is said to be a dominant action for player i if no matter what all other players are playing, playing $\tilde{a}^i$ always maximizes player $i's$ payoff.

Formally, for every choice of actions by all players except i, $a^{\neg i}$,

$\pi^i(\tilde{a}^i, a^{\neg i}) \geq \pi^i(a^i, a^{\neg i})$, for every $(a^i \in A^i)$. An outcome $(\tilde{a}^1, \ldots, \tilde{a}^N)$ is said to be an equilibrium in dominant actions if $\tilde{a}^i$ is a dominant action for each player i.

In the formal description the following notation has been used. Suppose $a = (a^1, \ldots, a^N)$ is the outcome of a particular game. We can pick a player i and remove from the set of outcomes the action played by this player, thus splitting the set a into two mutually exclusive sets a^i and $a^{\neg i}$. Formally

$$a^{\neg i} \equiv (a^1, \ldots, a^{i-1}, a^{i+1}, \ldots, a^N) \quad \text{and } a = (a^i, a^{\neg i}).$$

A game with an equilibrium in dominant actions for each player constitutes a realistic prediction of how players may interact in the real world. Unfortunately such win-win games do not often occur. The war-peace game of table B.1 has an equilibrium in dominant actions, which is $\tilde{a}^i$ = war. For both countries to play war, results in the highest payoff regardless the move of the other country.

B.2 Nash equilibrium

Another well known class of solutions in game theory is the Nash equilibrium. The Nash equilibrium represents a (locally) optimum strategy for all players.

*An outcome $\hat{a} = (\hat{a}^1, \hat{a}^2, \ldots, \hat{a}^N)$ (where $\hat{a}^i \in A^i$ for every $i = (1, 2, \ldots, N)$) is said to be a **Nash equilibrium (NE)** if no player would find it beneficial to deviate provided that all other players do not deviate from their strategies played at the Nash outcome. Formally for every player i, $i = (1, 2, \ldots, N)$,*

$\pi^i(\hat{a}^i, \hat{a}^{\neg i}) \geq \pi^i(a^i, \hat{a}^{\neg i})$, for every $(a^i \in A^i)$.

Note that for Nash equilibria to occur, the requirement *if no player would find it beneficial to deviate provide that all other players do no deviate* means transparancy. All players have full knowledge about all players, their payoffs and strategies to play.

Another distinction often made in game theory is that of **zero-sum** and **non-zero-sum** games. In zero-sum games, the total amount of payback that can be gained is fixed. Whatever is won by one player is lost by the other players and thus zero-sum. In non-zero-sum games that total amount is variable, which also implies that more players may win or loose at the same time [Heylighen 1993].

Shy presents in his book a large collection of business and economics decisions that are modelled as a game. These samples include theories on pricing, branding, intellectual property and also standards. An example on choices of companies to either comply or not-comply with standards is given in table B.2. A two-firm industry

		FIRM B			
		Standard α		Standard β	
FIRM A	Standard α	a	b	c	d
	Standard β	d	c	b	a

Table B.2: Two standards game [Shy 1995, p.255]

are making a product that could operate on standard α or on standard β. The profit levels of the two firms are given by non-negative numbers a,b,c, and d. What are the Nash equilibrium solutions to this game? In fact there are two outcomes Nash equilibria for this game:

If a, b > max{c,d} then the industry produces on a single standard, that is, (α,α) and (β,β) are Nash equilibria.

If c,d > max{a,b} then the industry produces on two different standards, that is (α,β) and (β,α) are Nash equilibria.

APPENDIX-C QUEUING THEORY

This appendix contains an introduction to the basic notions and concepts of Poisson processes and queuing theory. Queuing theory is often used in modelling the behaviour of queue-server systems and logistic networks.

C.1 Poisson and Binomial distribution

Poisson processes are widely used in modelling the behaviour of queues. A queue or a buffer is one of the basic building blocks of logistic networks. Poisson arrival process are defined as follows [Schwartz 1987, p25-43]. Consider a small time interval Δt $(\Delta t \to 0)$, separating times t and $t + \Delta t$, then:

1. *The probability of one arrival in the interval* Δt *is defined to be* $\lambda \Delta t + o(\Delta t)$, $\lambda \Delta t \ll 1$, *and* λ *is a specified proportionality constant.*
2. *The probability of zero arrivals in* Δt *is* $1 - \lambda \Delta t - o(\Delta t)$
3. *Arrivals are memoryless: an arrival (event) in one time interval of length* Δt *is independent of events in previous or future arrivals.*

The first two statements in the definition show that the starting point of the Poisson probability distribution is also related to the binomial distribution (with $p(1) = \lambda \Delta t$). The possibility of more than one event is ruled out by taking the interval Δt sufficiently small. Suppose $p = \lambda \Delta t$ and $q = 1 - p = 1 - \lambda \Delta t$, then the probability of k events in an interval $T = m \Delta t$, is given by the binomial distribution: $p(k) = \binom{m}{k} p^k q^{m-k}$. Now let $\Delta t \rightarrow 0, m \rightarrow \infty$ while maintaining $T = m\Delta t$ at a fixed, finite time interval, the binomial distribution limits to the Poisson distribution:

$$p(k) = ((\lambda T)^k e^{-\lambda T})/k! \textbf{\textit{ with }} k = 0, 1, 2, \ldots \qquad \textbf{\textit{(C.1)}}$$

For the Poisson distribution it can be shown that the mean and the variance both equal λT:

$$\textbf{\textit{Mean: }} E(k) = \sum_{k=0}^{\infty} kp(k) = \sum_{k=0}^{\infty} k\frac{(\lambda T)k}{k!}e^{-\lambda T} = \lambda T \qquad \textbf{\textit{(C.2)}}$$

$$\textbf{\textit{Var: }} E((k-E(k))^2) = E(k^2) - E^2(k) = \sum_{k=0}^{\infty} k^2 p(k) - (\lambda T)^2 = \lambda T \qquad \textbf{\textit{(C.3)}}$$

This means that $\lambda = (E(k))/T$, is a rate parameter. The combination of Poisson processes is also a Poisson process.

For Poisson processes the time between arrivals τ is an exponentially distributed random variable; its probability density function is given by:

$$f_\tau(\tau) = \lambda e^{-\lambda\tau} \textbf{\textit{ for }} \tau \geq 0 \qquad \textbf{\textit{(C.4)}}$$

The mean value of this distribution is:

$$E(\tau) = \int_0^{\infty} \tau f_\tau(\tau) d\tau = 1/\lambda \qquad \textbf{\textit{(C.5)}}$$

This was to be expected for a process that generates events at a rate λ.

C.1.1 M/M/1 and M/M/2 Queue

The performance of a simple queue-server model (M/M/1) can be modelled as follows. In this notation an M/M/1 queue is subject to a Poisson arrival time process with exponential arrival time distribution and there is one server that has an exponential service time distribution. Likewise, an M/M/2 queue has two servers, each with exponential service time distribution. An M/G/1 queue is subject to exponential arrival times, but has a general service time distribution and one server.

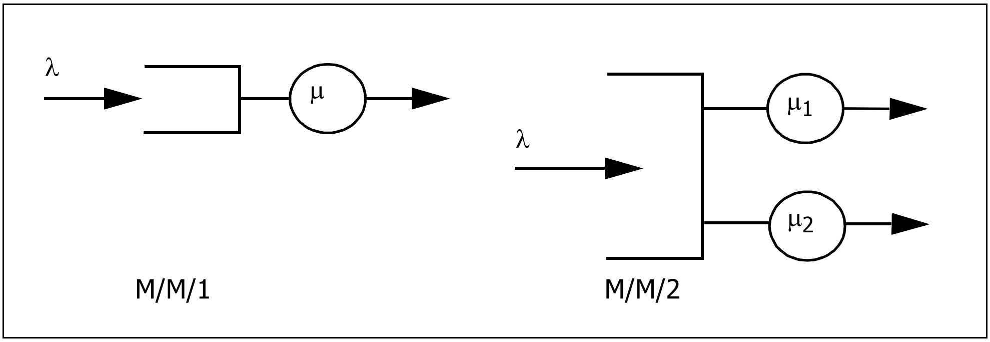


Figure C.1: Queue models with rate parameters λ and μ_i

The state of a queue-server system is often defined as the number of packages in the system. So for the M/M/1 system we can build the following state diagram.

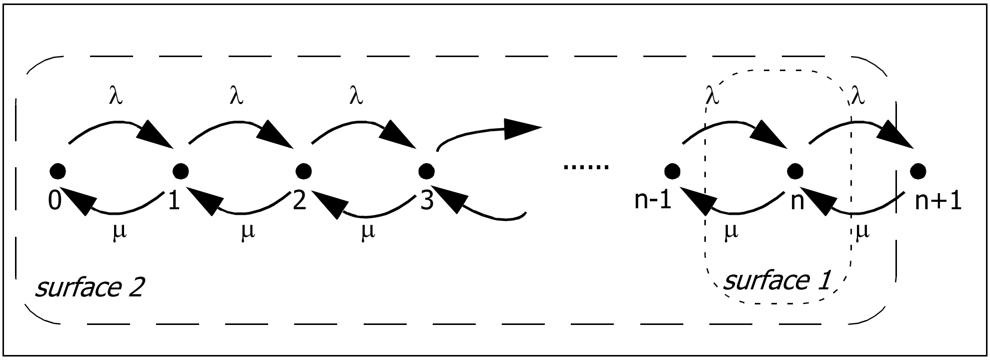


Figure C.2: State diagram for M/M/1 queue [Schwartz 1987, p34]

Now consider 2 closed surfaces, surface 1 around just 1 state n and surface 2 around all states up to n. For both surfaces we can formulate balance equations. For a stationary solution for state n (surface 1),

$(\lambda + \mu)p_n = \lambda p_{n-1} + \mu p_{n+1} \qquad n \geq 1$, should hold. For surface 2, the equilibrium solution requires, $\mu p_{n+1} = \lambda p_n \qquad n \geq 0$. Repeating this equation n-times, and defining $\rho \equiv \lambda / \mu$, being the load factor for this queue, we find: $p_n = \rho^n p_0 \qquad \rho < 1$. Together with the probability normalisation requirement $\sum_{n=0}^{\infty} p_n = 1$, we find:

$$p_n = \rho^n (1 - \rho) \qquad \rho < 1 \qquad n \geq 0 \tag{C.6}$$

and $p_0 = 1 - \rho$ for an infinite M/M/1 queue.

p_0 is the probability that the queue-server system is empty. The load factor ρ has several important interpretations for a queue-server system:

- *ρ is the probability that the queue-server system is not empty.*
- *ρ is the utilization of the server.*
- *ρ is the average number of jobs on the server.*

The average number of customers or jobs for a non-congested M/M/1 queue-server system is:

$$E[n] = \sum_{n=0}^{\infty} n p_n = \sum_{n=0}^{\infty} n \rho^n (1 - \rho) = \frac{\rho}{1 - \rho} \tag{C.7}$$

The behaviour of $E[n]$ is depicted in figure C.3.

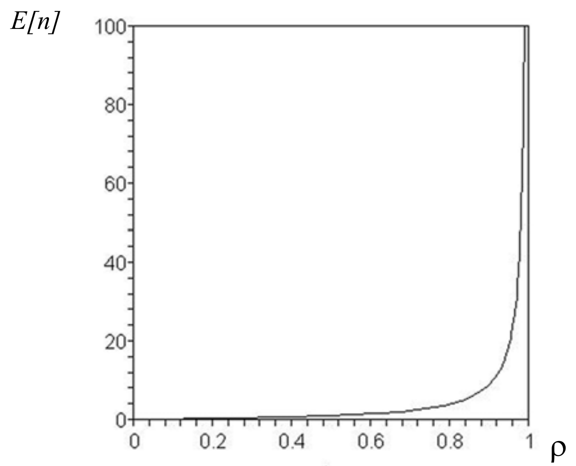


Figure C.3: Average queue size for M/M/1 queue.

So if n_q is the number of jobs in the queue waiting for service, then it follows that:

$$E[n_q] = \frac{\rho}{1-\rho} - \rho = \frac{\rho^2}{1-\rho} \tag{C.8}$$

With Little's law $E[n] = \lambda E[T]$, where T is the total leadtime it follows for the waiting time in the queue Tw that:

$$E[Tw] = \frac{\rho}{\mu(1-\rho)} \tag{C.9}$$

For a finite M/M/1 queue-server, which can hold a maximum of N packages, one finds:

$$p_n = (1-\rho)\rho^n / (1-\rho^{N+1}) \tag{C.10}$$

The probability that this queue-server system is full p_N now becomes

$p_N = (1-\rho)\rho^N/(1-\rho^{N+1})$. p_N is also known as the blocking probability P_B, the probability that packages are turned away because the queue-server system is full. For large N and small $\rho < 1$ the formula for the blocking probability can be reduced to $P_B = \rho^N(1-\rho)$. The behaviour of the blocking probability as a function of ρ is shown in figure C.4.

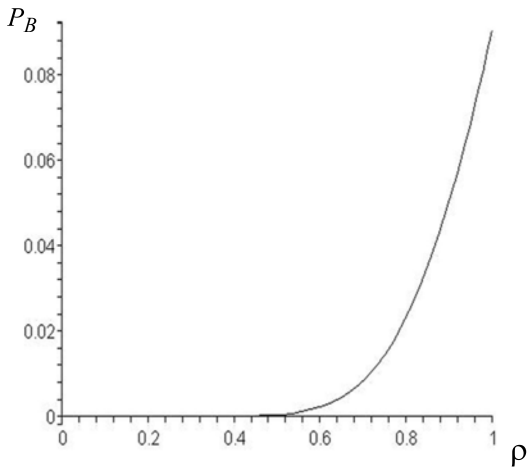


Figure C.4: Blocking probability for a finite (N=10) M/M/1 queue.

From figure C.3 and from figure C.4, it becomes clear that for $\rho > 0.8$, both the blocking probability and the average queue size rapidly increase. The same applies for the average waiting time as well as the total time in the queue-server system [Schwartz 1987, p41-43]. The (server rate) normalized delay for the M/M/1 queue-server system becomes: $\mu E[T] = \frac{(E[n])}{\rho} = \frac{1}{1-\rho}$.

C.1.2 M/M/2 Queue

In case of a queue with 2 servers with equal service rate μ, the M/M/1 theory needs to be expanded to a state dependent queue model. This is a model where the arrival

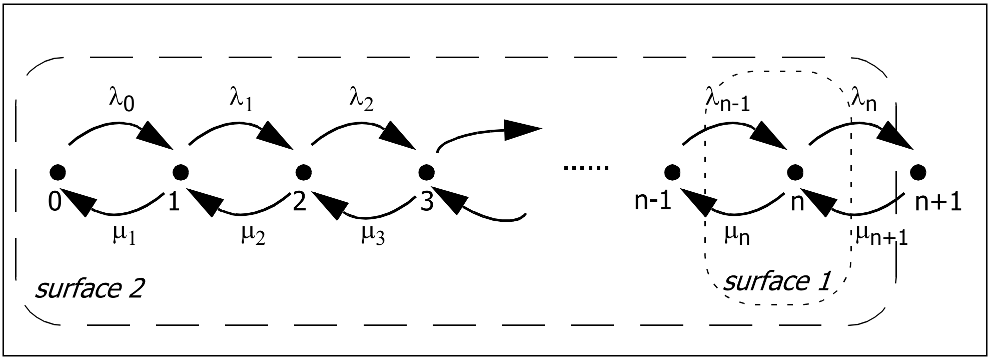


Figure C.5: State diagram for state-dependent queue [Schwartz 1987, p49]

rate and the server rate, also depends on the number of customers waiting in the queue. Having two servers means the state equations for $n = 1$ and for $n \geq 2$ are different.

For the state dependent queue of figure C.5, the following general state-equation applies: $\frac{p_n}{p_0} = \prod_{i=0}^{n-1} \lambda_i / \prod_{i=1}^{n} \mu_i$. For an M/M/2 queue with two servers with equal service rate μ, we have: $\lambda_n = \lambda, \mu_1 = \mu, \mu_n = 2\mu, n \geq 2$ and $p_n/p_0 = \left(\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{2\mu}\right)^{n-1} = 2\rho^n$ with $\rho \equiv \frac{\lambda}{2\mu}, n \geq 1$. Together with the normalization condition, this leads to: $p_n = 2\rho^n \frac{1-\rho}{1+\rho}$. The average queue size now becomes: $E[n] = \frac{2\rho}{1-\rho^2}$ with $\rho \equiv \frac{\lambda}{2\mu}$. The normalized leadtime now becomes: $\mu E[T] = \frac{(E[n])}{\rho} = \frac{1}{1-\rho^2}$ with $\rho \equiv \frac{\lambda}{2\mu}$. Note that we have multiplied the leadtime with the capacity of serving only 1 job here. These results are also presented in figure C.6.

From this figure it is clear that the M/M/2,μ queue has twice the capacity of the M/M/1,μ queue, but that the M/M/1,2μ outperforms the M/M/2,μ queue for smaller loads. This is no surprise since a single server running at twice the rate, has only half the service time.

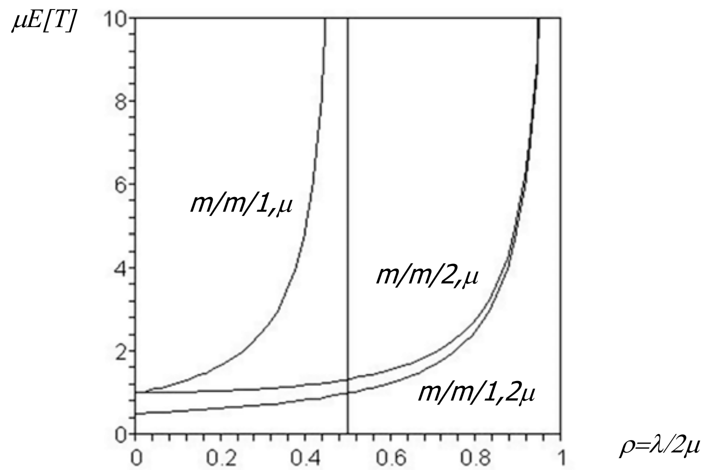


Figure C.6: Normalize time delays compared M/M/1 and M/M/2

The practical consequence of this theoretical result would be that a single server with a high capacity is preferred over multiple-server solution with smaller capacities that sum up to the same total capacity. However there is a downside to a single queue solution that is not considered in the results above. A multiple-server solution is less vulnerable to breakdowns and offers more flexibility in maintaining queue leadtime performance under high loads. The reason is the influence of variability. A load pattern with a high volume variability may be segmented into multiple streams with lower variability per stream. A multiple queue-server system can offer dedicated capacity to each stream, operating closer to the leadtime requirements of the distinct markets, than a single queue server solution can offer. This will be demonstrated in the simulation experiments in chapter 3.

C.2 M/G/1 queue

For queue systems, incoming events are mostly modelled as poisson processes. For the server, responsible for the outgoing rate, an exponential service time distribution is not always a good model. For manufacturing systems, the variance of the service times is usually much smaller than would be expected if the service times were exponentially distributed with the same average rate. The Pollaczek-Khinchine formula's provide a more general result for the expectation of the queue occupancy and the expected leadtime. For a service time distribution with service rate μ and variance of the service time distribution σ^2 operating under a load ratio $\rho = \lambda/\mu$:

$$E[n] = \left(\frac{\rho}{1-\rho}\right)\left[1 - \frac{\rho}{2}(1-\mu^2\sigma^2)\right] \qquad \textbf{\textit{(C.11)}}$$

and

$$E[T] = \frac{E[n]}{\lambda} = \left(\frac{1/\mu}{1-\rho}\right)\left[1 - \frac{\rho}{2}(1-\mu^2\sigma^2)\right] \qquad \textbf{\textit{(C.12)}}$$

Note that for an exponential service time distribution, $\sigma^2 = 1/\mu^2$, which when substituted in (C.11) and (C.12) yield the result (C.7) already known for the M/M/1 queue. Substituting $\sigma^2 = 0$, yields the result for the M/D/1 queue with deterministic service times. Similar to (C.8) and (C.9) we can find expressions for the number in the queue n_q and the waiting time in the queue Tw.

$$E[n_q] = E[n] - \rho \qquad E[Tw] = \frac{E[n_q]}{\lambda} \qquad \textbf{\textit{(C.13)}}$$

APPENDIX-D RESEARCH METHODOLOGY IN MANAGEMENT SCIENCE.

Doing research in the area of business or management science and validating results requires careful consideration on the applied research methodology. The research in this thesis is largely based on systems modelling, inspired by well established theories that are being transposed to a different context to provide answers and new theories for problems observed in practice. Validating results from these tools has been done in two ways. The first validation is checking the correct use of existing theories and checking the logic of the reasoning. The second method applied for validation is the prediction of some artifacts in practice and find them. This is the action research model. The down side of this method is that these practical cases may be difficult to find or to get access to.

For the study of complexity in organisations, model building and action research is considered the only way. The analysis of complex systems is only possible through hypothesizing systems models responsible for the observed behaviour. The use of surveys is limited here. Statistics tent to drown the causes for complex behaviour in noise and even if a survey would be possible, the systems model is indispensable as a reference for the design as well as for the analysis of the results of the survey.

D.1 Deduction, Induction and Abduction

Before Pierce, reasoning logic only knew two basic classes of arguments: deductive arguments (necessary inferences) and inductive arguments (probable inferences). Pierce realised that in scientific methods there was also a third class, different from induction, which was named abduction. Deduction, induction and abduction together are the three basic reasoning modes of science. ***Abduction*** became defined as inference to and provisional acceptance of an explanatory hypothesis for the purpose of testing it. ***Deduction*** came to mean the drawing of conclusions as to what observable phenomena should be expected if the hypothesis is correct. ***Induction*** became to mean the entire process of experimentation performed in service of hypothesis testing [Burch 2001]. In other words, the hypothesis infers predictions about observations that could further support the hypothesis, where the hypothesis has taken the form of generalized observed regularities.

$$Facts \cup Theorems \xrightarrow{Induction} Axioms$$
$$Axioms \cup Facts \xrightarrow{Deduction} Theorems$$
$$Theorems \cup Axioms \xrightarrow{Abduction} Facts$$

Figure D.1: Induction, Deduction and Abduction

Schurz notes that inductive generalizations are not creative, whereas abduction does require a creative step that puts an observed fact that needs explanation in a new hypothetical context. The most general form of abduction is as follows [Peirce 1903]:

F: an observed (singular or general) fact[1] *which is in need of explanation.*
B: Background knowledge which implies that the hypothetical situation R would be a reason (cause or explanation) of F.
==
R: Conjecture that the hypothetical reason R is in fact the case.

Thus abduction is put forward as a more creative reasoning mode that is essential to the progress of science. From the perspective of strategies of irrogation or what-if type of reasoning as a basis for abduction, Schurz has developed a classification scheme for different kinds of abduction in which three main categories of abduction are distinguished: factual abduction as a basis for understanding (non-innovative) design, law abduction as a basis for understanding science and 2nd order existential abduction as a basis for understanding discoveries in science [Schurz 2002].

1. Pierce used the term fact as a reference to the observation. However the “nature” of this fact at the start of the what-if type of abduction processes is different from the fact that is the end result of this process. The observation is also influenced by the hypothesized combination of axioms and theorems. Thus at the start of abductive processes facts are closer to theorems than to undisputable facts.

KIND OF ABDUCTION	EVIDENCE TO BE EXPLAINED	ABDUCTION PRODUCES	ABDUCTION IS DRIVEN BY
Factual abduction:	Singular empirical facts	New facts (reasons, causes)	Known laws or theories
Observable-fact-Abduction	Singular empirical facts	Factual reasons	Known laws
Unobservable-fact-Abduction	Singular empirical facts	Unobservable reasons	Known laws
Historical-fact-Abduction	Singular empirical facts	Facts in the past	Known laws
Theoretical-fact-Abduction	Singular empirical facts	New initial or boundary conditions	Known theories
1st order existential Abduction	Singular empirical facts	Factual reasons postulating new unknown individuals	Known laws
Law abduction:	Empirical laws	New laws	Known laws
2nd order existential abduction:	Empirical laws	New laws/theories with new concept	Theoretical background knowledge
Micro-Part-Abduction	Empirical laws	Microscopic composition	Extrapolation of background knowledge
Analogical Abduction	Empirical laws	New laws / theories with analogical concepts	Analogy with background knowledge
Missing-link Com.cause-Abduction	Empirical laws	Hidden common causes	Causal background knowledge
Fundamental Common cause Abduction	Empirical laws	New unobservable properties and laws	Unification of background knowledge
Theoretical property Abduction	Empirical laws	New theoretical entities	Unification of background knowledge
Abduction to reality	Introspect. laws	External entities	Unification of background knowledge

Table D.1: Classification of kinds of abduction [Schurz 2002]

Tomiyama among others discussed the use of abduction in design. In the framework of Schurz, design could be compared to the class of factual abduction. The created design is a fact within the reasoning frame spanned by the axioms and theorems (requirements). Yet, this type of factual abduction is not capable of generating creative or innovative designs. A framework of axioms and known theorems merely deduces the desired artifact. It is not exactly the same as deduction, since experimental testing is still needed for validation. The theorem specifying the requirements may not be complete or the set of axioms supporting the new artifact may not be complete. Creative design is more likely if the new set of requirements (theorem) requires additional theorems to achieve a fit to the axioms. Sometimes even axioms and theorems of a different theory are needed to be integrated for the purpose of building an integrated design (abduction for integrating theories, part of 2nd order existential abduction) [Tomiyama 2003].

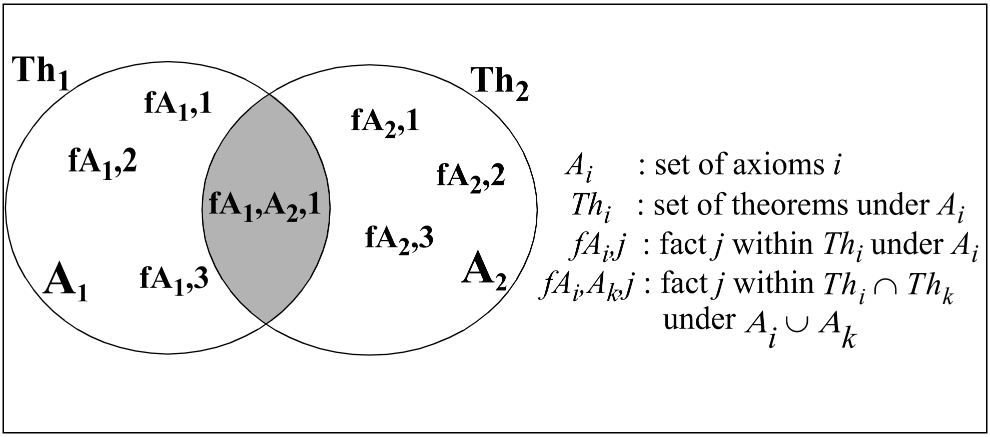


Figure D.2: Abduction for integrating theories

Thus, despite many reasons for distinguishing several kinds of abduction, there is not a fundamental difference between the practice of science and the practice of design. Science and Design need one another for making progress. Both require creativity and an effective reasoning strategy to find a robust design or to discover a new truth. Note that in neither case the reasoning mode is limited to abduction only. As far as there exists a perceived difference in reasoning modes between design and science, this difference maybe attributed to over emphasizing the importance of creativity in design versus too much risk averse analytical or deductive thinking in science.

D.2 Research Methods in management science

The framework of the reasoning modes deduction, induction and abduction is also relevant for correctly applying research methods. The applied reasoning mode drives the type of validation needed for the research. But this is not all there is to it. The nature of the theorems and axioms used as premises in the reasoning also influence the validation. In an earlier paper, Schurz proposed the following classificatory schema to distinguish Laws of nature from System laws [Schurz 2001].

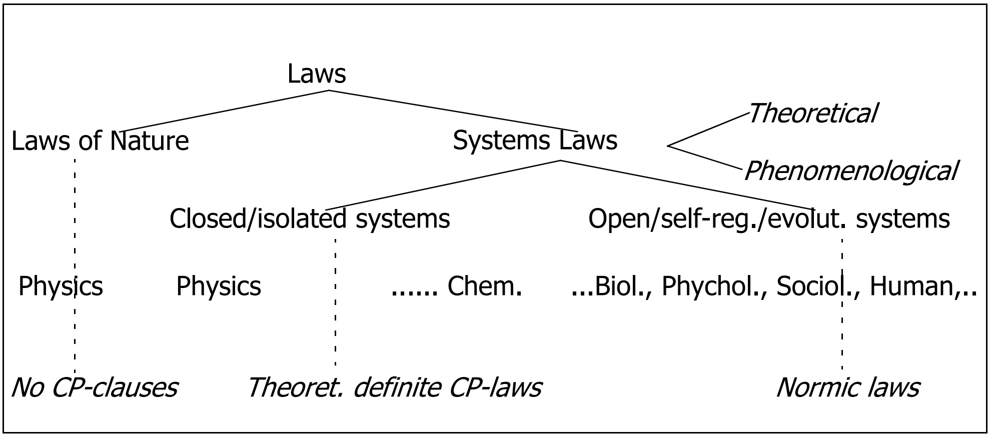


Figure D.3: Laws of Nature versus System Laws [Schurz 2001]

Laws of Nature are fundamental laws of physics which hold everywhere and any time. These laws are strictly true; that is without any ceteris paribus (CP) clauses. *System laws* however, speak about concrete systems of a particular kind, operating under certain so-called boundary conditions present at a certain time. A further distinction made by Schurz is that of *Theoretical* vs. *Phenomenological* system laws. *Theoretical system laws* usually state a set of special differential equations for the system under consideration, operating under certain boundary conditions. *Phenomenological system laws* describe the temporal behaviour in a pre-theoretical vocabulary under certain initial and boundary conditions. The next major distinction is that of system laws for *closed/isolated systems* versus laws for *open, self regulatory or evolutionary systems.* The CP-clauses that are indispensable for formulating system laws for closed systems in fact demand that no further possibly disturbing factors are present within the system boundaries for which the system laws apply. As argued in chapter 2 of this thesis, the quasi-static boundary conditions are such CP-clauses for being able to use closed system laws for process and organisation design for business systems.

For open systems the absence of possibly disturbing factors can no longer be guaranteed, which has a profound effect on the nature of the laws applicable to open systems. Very generally, systems are physical ensembles which preserve a relatively strict identity over time. For closed systems this identity is preserved through their isolation provided by the boundaries. For open- or living systems, preservation of their identity requires the characteristic of self-regulation aimed at maintaining that identity through compensating for disturbing influences from their environment. Schurz states that *normic laws are the phenomenological laws of self-regulatory systems* [Schurz 2001]. As will be discussed in this thesis, the concept of *normic laws* is in agreement with the concept of autopoietic systems of Maturana [Maturana 1980].

Arbnor and Bjerke present in their book three general classes of methods developing business knowledge: the analytical approach, the system theory approach and the actors approach [Arbnor 1996]. With some overlap, these three approaches to developing business knowledge relate to different paradigms or different ways of looking a reality (figure D.4). One could also say these paradigms relate to different axioms and theorems about the reality of business problems and business knowledge.

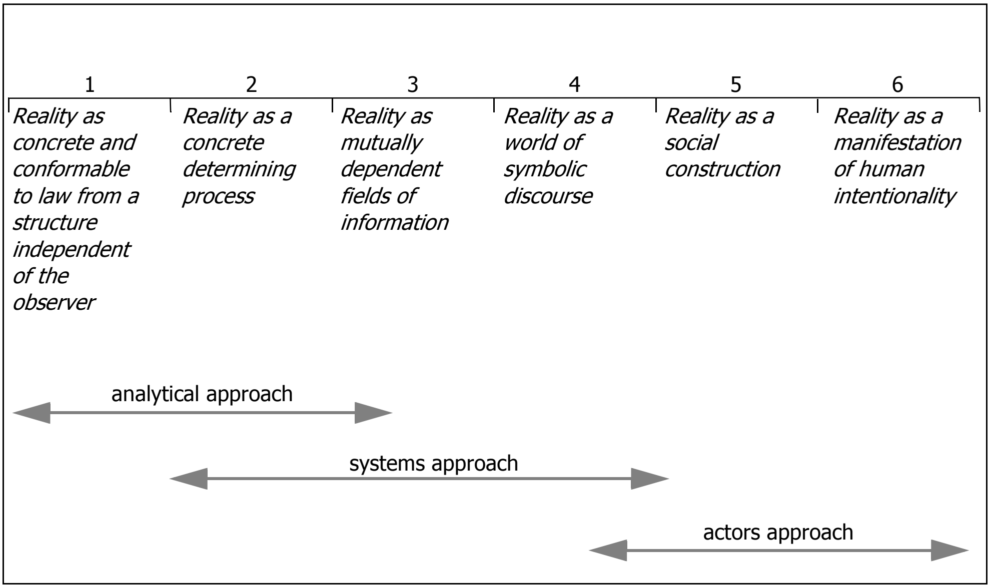


Figure D.4: Three methodological approaches and their paradigmatic categories (from [Arbnor 1996])

The analytical approach and the actors approach are like opposite sides of the same coin, which may be bound by a systems approach.

The analytical approach sees reality as conformable to law and observable without interfering with the observed. The analytical approach also assumes reality has a summative character, that is, the whole equals the sum of its parts. Thus, parts can be studied in isolation. Relations between parts hardly matter. A closed world, as well as distinct boundaries between parts are assumed for practical reasons. Knowledge advances through formal logic based on specific judgments that are independent of individual subjective experience. These judgements consist of hypotheses that can be either verified or falsified. Often surveys and statistical analysis are the practical tools used for that purpose. The logic of combining judgements is mainly based on deduction, verified in practice through statistical analysis of observations.

The actors approach sees reality as a social construction and as a manifestation of human intentionality. This means the research focuses on observation of social interaction and communication and tries to build theories that explain the logic of events from building theories about interest, intent and culture of interacting humans. The primary aim of this type of research is not so much to explain, but to understand the social construct.

The systems approach is based on the assumption that reality is arranged in such a way that the whole is more or less than its parts. Thus the relations do matter because they account for the plus or minus effects between the parts. Boundaries between the parts under study as well as a boundary on the whole now also has a determining effect on the outcome of research according to the systems approach. The system approach may also include or adopt aspects and methods from the actors approach, if the effects of relations between parts are significantly effected by human intent or culture.

In brief, the analytical approach assumes reality is objective and can be studied independent of the researcher. The systems approach assumes that reality is objectively accessible, but realizes that the goals and intentions of the researcher have a dominant impact on the outcome. This is in line with the systems approach of In't Veld as briefly discussed in paragraph 2.3. The actors approach assumes that reality is a social construct that can only be researched through assessing the actors finite provinces of meaning. The intentions of the researcher are an inseparable part of this social construct.

Regardless of the approach, survey methods and statistical analysis are widely used in business administration research. As will be argued in paragraph D.3, the use of survey methods is not without risk. Without models and theories as reference and context for survey experiments it is not possible to design and execute sound experiments with trustworthy results.

D.3 Survey methods

Survey methods are often applied as the process of validation or experimentation. Survey methods in general deliver substantial amounts of data that can be processed and analysed. This is in line with Peirce's concept of induction as long as there are specific axioms that could support the theorems or hypothesis to be tested. This type of analysis can become quite slippery if the axioms are too general or only based on "belief in statistics".

According to Saunders, induction means that theory should follow data [Saunders 2000]. This argument is often used for justifying the use of survey methods and statistical analysis, where hypothesis could be tested against answers of survey questions.

The survey model of doing research is often based on the following experimental model (figure D.5). A company uses resources (material, energy, human labour and capital,....) to fulfil the needs of a collection of customers and is rewarded for this fulfilment in money. Successful companies are those that earn enough money to increase their value. This means, the returns received over the capital deployed in this company exceed those of other companies with a similar risk profile. The

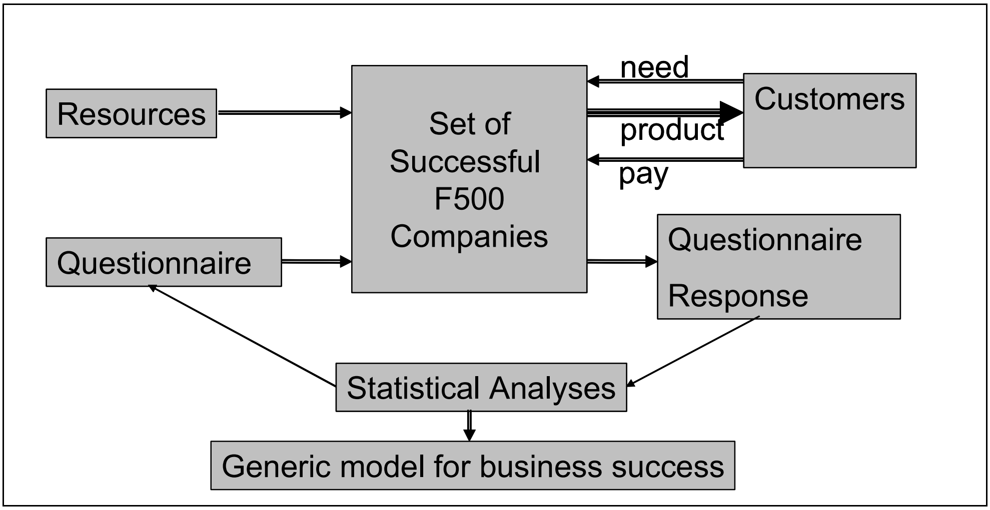


Figure D.5: The survey research model

intriguing question, both for other businesses and for business researchers is: how did they do this? What is their (secret) recipe? In order to find these recipes for success, researchers pose hypotheses and send out questionnaires to these companies to test these hypotheses. The responses to these questionnaires are analysed statistically for the purpose of either accepting or rejecting the hypotheses. Often these questionnaires are being sent to more companies, sometimes active in different markets, in order to find commonalities that could represent a more general truth for business success.

There are many pitfalls with this research model that fall into three major categories: experiment design, questionnaire design and statistical analysis.

D.3.1 Experiment design

Executing a survey is like doing an experiment. In order to get decisive results from an experiment, the researchers need to control the circumstances for the experiment and they need to use accurate response measurement instruments. Depending on the goal for the experiment, the stimulus that triggers the responses needs to be specific or selective. In case of the survey type of experiment, it is not always clear what the stimulus actually is. Ideally, the hypotheses should contain the stimulus and we are interested in the response of various companies to that stimulus. But it is very well possible that the questionnaire itself is becoming the stimulus and the responses are largely opinions and speculations about possible responses.

A special problem in experiment design is sample selection. Ideally the sample should be unbiased. This means that the questionnaires are sent out to a large number of companies, which may share some characteristics necessary for controlling the context of the experiment, but which do not share the same response to the stimulus under study a-priori. The a-priori probabilities of acceptance or rejection should be equal. The same should apply to the response population. However, in practice not all questionnaires are returned. Quite often, questionnaires that are returned are a subset that do not exactly share all characteristics of the entire population. In reporting the results this should be made clear through comparison of known "descriptive" data over the population targeted for the survey, with the specific "descriptive" data over the population that responded. But even then, respondents may have a self-interest to return the questionnaire which may not appear in the return population characteristics. Another reason for self selection may be due to the experiment set up. The time intervals of observation and the categorization over that time interval may lead to biased populations of respondents. Nobel prize laureate James Heckman worked on correction methods for correcting the bias that could enter the results from self selection phenomena. Note however that these corrections rely on a model of the drivers behind self selection [Royal Swedish Academy of Science, 2000].

A special problem may be sample size. For a multi-purpose survey the minimum response sample size may be determined as follows. Select the issue where there is likely to be the greatest diversity in the sample. Now there are two categories: one for which the issue holds (proportion p) and the other one for which the issue does not hold (proportion $q=1-p$). To observe the correct proportions from experiment with an error margin of e and a confidence level of 95% a minimum sample size n is required (1), where c follows from the χ-square distribution with 1 degree of

$$n > \frac{p \bullet q}{e^2} \bullet c \qquad \textbf{\textit{(D.1)}}$$

freedom and confidence level 95%. For a worse case of $p=q=0.5$ and error margin $e=0.05$, a minimum of 385 samples are required in case of an infinite (larger than 10000) population. In practice this is a large number requiring a lot of work. The total number of questionnaires send out often needs to be even bigger, since a 100% return rate is not to be expected for large populations. For a smaller population N,

$$n_{adjusted} = \frac{n}{1 + \frac{n}{N}} \qquad \textbf{\textit{(D.2)}}$$

the sample size may be adjusted (2), but in business research the total population size (N) is not always known [Saunders 2000]. Combining (1) and (2) means that for small population sizes, 100% coverage is needed tot achieve significant results.

D.3.2 Questionnaire design

The questionnaire that is sent out to the companies is effectively a communication tool for which the communication model as presented in this thesis also applies (figure 5.3). This means the researcher's intent should take into account the respondents context when designing the questionnaire. It also means that the questionnaire needs to gather contextual information about the respondents to be able to verify whether the response can be used, to categorize or to correct for a population bias that can be discovered through analysing context information.

Respondents use their own knowledge and their own value system to answer to the questions. Primarily the answers contain opinions and perceptions, not facts. This is not a problem if the survey was designed to gather opinions. However if the researcher wants to gather facts then the questionnaire must provide unambiguous scales of reference for the answers. Even then, there may be incompetent respondents and respondents that have their own reasons for participating in the survey. Through their answers they try to influence the outcome based on their speculation on the use of the results. Additional test questions are needed for checking consistency and for filtering incompetence.

Often researchers assume explicit awareness over all questions asked. Yet sometimes the questionnaire itself triggers the thought process and induces awareness. As a result the distinction between "what is" and "what ought to be" often cannot be made from these answers.

Gut feeling, luck or none of the above, are not often included as possible answers to difficult questions on drivers for business success, whereas entrepreneur often mention their instinct as a factor which guides their attention to certain developments. Open questions can be much more informative in such cases, but they take more time from the respondent and are almost impossible to process statistically.

D.3.3 Statistical analysis

Statistical analysis of the survey response is possible with a large collection of tools and tests. Well known are the χ-squared test for testing the significance of hypothesis for data observed and the student-t test for testing comparability between two populations or two sets of observations. Correlation or co-variance analysis can be used to establish linear relationships between variables. Auto correlation analysis of time-series can be used to detect short term memory or storage effects of systems. Beyond a certain time-span, the auto correlation values become small and observations separated beyond this time-span can be treated as independent.

Often a questionnaire is built over a number of hypothesis that should be accepted or rejected after the collection of data. However be aware that the power of statistical tests like χ-squared, the t-test or any statistical method is *to reject or not to reject* certain hypothesis in relation to observed data. To accept a hypothesis is a different issue. Acceptance beyond any doubt requires proof that the set of factors studied and the set of hypothesis related to these factors is complete. This means there are no other explanations possible for the observations to be analysed. Such a proof of completeness cannot be given for open systems. Only within the boundaries of a model that describes all relevant factors and relations we try to investigate and only under the assumption of quasi-static boundary conditions, a closed environment exist for which a completeness proof may be possible. This demonstrates that modelling is vital before questionnaires can be designed and before proper statistical analysis is possible. The intent of the researcher should include sufficiency conditions that are based on the intended use of the results for creating both the model and the model boundaries.

Unless all observations are time-stamped and the analysis includes methods for the analysis of time-variant models, statistical analysis assumes time-invariance of a system. This is not necessarily a problem for studying humans. Because our rate of change through evolution is rather low, general aspects of human behaviour can be studied safely under the assumption of time-invariance. However if behaviours of humans under market conditions are investigated then the assumption of time-invariance is no longer valid and the experiment design should cater for the expected dynamics that affect the observations. Humans learn and adapt their behaviour in response to changing market conditions. The non-linearity and dynamics of market conditions are also responsible for the poor repeatability of experiments in business research. Only large databases with timed records of many factors describing macro- as well as micro-economic behaviours, can be reused for different experiments improving our understanding of models and mechanisms that have occured in the past.

Statistical tools are good at filtering noise through averaging over all responses. Yet the question is, are all responses subject to the same type of errors that can be characterised as noise? Especially if complex issues are addressed, summing up all answers into a statistical mean, destroys much of the information that was available in the individual answers. Dominant consent is very capable of drowning isolated opinions of respondents that are more advanced in their understanding. In fact the information theory[1] as well as control theory teach us that the odd response from a system driven to its limit, provides more information about the internal mechanism then the background noise that is always there from the steady state of operations. Again physical or systems modelling is essential to shape statistical analysis in such a way that these odd signals are detected and can be treated correctly in the analysis.

1. Self information, paragraph A.2

D.3.4 Surveys, a typical example

Even when a survey experiment is done by the book, one could question whether the laborious work of designing the survey, gathering the data and performing the statistical analysis are needed. A typical example of such work is presented in a recent paper by Morgan Swink and Roger Calantone [Swink 2004]. In their paper titled "Design-Manufacturing Integration as a Mediator of Antecedents to New Product Design Quality", they argue that a high level of design quality is not a direct result of high technological novelty and project organization complexity, but it is only achieved if Design-Manufacturing Integration (DMI) is deployed as an organizational means to deal with the increased level of complexity that originated in uncertainties from new technologies and a multitude of interfaces in a multi-disciplinary project organization. This was concluded from fitting three different relational models over data gathered from a survey. The question here is, is the experiment necessary to support their conclusion? Were there reasonable doubts about the established relation before the experiment was done? Did the authors discover a hidden variable or a new mechanism? I don't think so. The concept of DMI existed before this experiment. The survey explicitly tested for the level of DMI. A more interesting question could have been how DMI is implemented by these companies or how managers create conditions for successful deployment of DMI? For such questions, the diversity of solutions is bigger but the same diversity may cause problems for the statistical analysis.

Nonetheless, the survey as well as the paper describing the results were done "by the book". The population has been described in the paper as well as the population selection method. Tests have been done whether the (sub)population of early responders differed significantly from those that responded later or only after several reminders. The function level of the persons that answered the questionnaire was recorded. Reasons for not returning the questionnaire were also investigated. Theory development and measurement techniques used in the questionnaires were documented and have been related to methods used by researchers that faced similar problems. Where quantitative data mattered, the answering scales where anchored to absolute scales representing these quantitative date. Most of the questions asked are presented in an appendix to the paper. The models used for fitting the data have been described as well as the method used for fitting the data. Even limitations of this study were addressed. Yet the discussion on the results is somewhat more speculative then the hypothesis and the results of model fitting really support. The most interesting questions are actually raised as recommendations for further research, as if the authors felt that the real issue of how to combine technical novelty and high project complexity, while still achieving a high product and project quality is still unsolved.

D.4 Complex systems

In case of complexity and business processes, the value of survey experiments is limited. Partly because a comprehensive set of models and theories does not exist yet, but even more important, individual businesses, living organisms alike, shape their environment. This means that the complexity faced by one business may be different in any aspect from the complexity faced by another, even though they may be in the same branch. Individual companies shape their environment and make explicit choices about the complexity they are facing. This means, being successful, requires the ability to identify and shape a defendable market niche. Within the market niche, quasi static boundary conditions apply and "closed" system models and laws support the development and implementation of efficient business processes. Being successful does not require complete understanding of the complexities of the world market. The world market influence is accounted for in the quasi static boundary conditions. What is required is recognition of conditions that could cause changes in the boundary conditions for this particular niche. A change in these conditions requires companies to respond with appropriate changes in their strategy and operations.

What response is most effective cannot be deduced from axioms driving past performance and the changes themselves. This is the dilemma facing many economic studies. Macro economy is the result of superposition of many micro-economic behaviours as operated by individual businesses and individual customers. The dilemma is twofold; first, how to relate the results of macro-economic studies to studies of micro-economic business behaviours and second, how to manipulate macro-economic variables to stimulate micro-economic behaviours that will lead to improved macro-economic performance. The creativity of individuals and supposedly irrational behaviour at micro economic level stand between achieving a macro economically desired state through promoting desired behaviours at micro level. Statistical analysis of surveys are only suited as scientific instruments for finding (partial) solutions that can solve this problem, if models for behaviours at micro level exist and are taken into account. The success of this approach has been demonstrated by four Nobel prize laureates in economy in the recent past: James Heckman and Daniel McFadden for their work on the effect of self selection and discrete choices in economical studies in 2000 and Vernon Smith and Daniel Kahneman for their work on behavioural and experimental economics in 2002 [Royal Swedish Academy of Science, 2000, 2002].

D.5 Action Research

A relatively new school of research-methods is Action Research. Action research received recognition for some time especially in social sciences. Recently there seams to be a bit of a revival. Since 2000, two new journals have appeared: *Action Research* and *Action Research International*. In his book on action research, Peter Clark identified three trends in the area of organizational change that triggered a rethinking of research methodology. The first observation was that practice shows only few successful innovation, yet the literature is full of case studies of success stories. Second is the predominant explanation of failure, that is based on the assumed resistance of people to organizational change, yet there are many instances where failure did occur in cases where initial resistance was low or even zero. Third, there are many individual case studies, a limited number of useful readings presenting particular schools of thought, and even a few samples of systemic research, yet there is no single text that relates the case studies, the readings and the theories. Consequently, there is confusion about organizational change and the uses to which behavioural sciences can be put to facilitate change. Action research is considered an essential strategy for facilitating as well as learning about organizational change [Clark 1972].

Clark adopts a definition on action research from R. Rapoport:
Action research aims to contribute both to the practical concerns of people in an immediate problematic situation and to the goals of social science by joint collaboration within a mutually acceptable ethical framework.

To this definition M. Foster added the following: *...and the intention of the parties is to be involved in change (which must be) change involving the properties of the system itself.*

According to F.Heller in the book of Clark this means that there are two principle ways to be directly involved in organizational change and to simultaneously provide an increase in knowledge: *First, when there can be an attempt to test out and/or replicate theory in a new setting by immediate involvement in the implementation phase of a situation which constitutes a valid test of the theory. Second, by joining a situation in which a decision about change has been taken by members of the sponsoring system*[1]*. The role of the practitioner is to document the change process and provide feedback on the organizational change process so that it might be stopped, modified or accelerated.*

The difference between traditional research and action research is also put forward firmly by Kimberly DeTardo-Bora [DeTardo-Bora 2004]: *In traditional research, change is often unforeseen, whereas in action research, change is not only the impetus for, but the goal of the project.* John Ellis and Julia Kiely state [Ellis 2000]: *Business Organisations tend to assume the dominant logic is correct and fail to question whether or not the values on which they are based remain appropriate or desirable. Are the right issues being tackled? Are the right questions being asked?*

1. Industry or society

Ellis and Kiely also state: *the ideology of action research is concerned with individual and organizational effectiveness.* Both, DeTardo-Bora and Ellis question the effectiveness of research strategies based on the survey paradigm and put action research forward as an alternative strategy that is much closer to the real needs of society and business for scientific knowledge; the ability to change and improve.

Although the term action research was never used to earmark the research practice used often in the Delft School of Organization Design, it seems that our approach has much in common with the statements above. First our method shares the intent to change and improve the situation of particular companies and in doing so to add to the more generally applicable knowledge of design methods and modelling practices of business processes. This intent is also supported by the engineers attitude to solving problems. Teaching the survey paradigm, industry creates, the market validates and academia tries to understand afterwards, does not support the training of engineers, who we expect to develop skills to identify and solve problems from scratch, preferably well before these problems surface in reality. Second, our method is participative and is most desirably deployed in situation which match the first category mentioned by Heller; that is to test and evaluate theory in a new setting through immediate participation in the implementation phase of a situation that constitutes a valid test for that theory.

The following scheme presents our implementation of action research.

1. Identify problems and improvement opportunities for business processes.
   - Market trends.
   - Impact of new technologies.
   - Benchmarks.

   } ⟹ Scenarios
2. Design sustainable processes and structures that can deal with (new) opportunities.
3. Select relevant companies.
4. Verify the applicability of the design.

Figure D.6: Action Research Model of the Delft School of Organisation Design

Along the lines of the action research model, Bikker's business process redesign model has been developed and the order entry point theory as well as the order specification decoupling point theory, were developed and tested in the past [Bikker 1994b, 1998]. More resent examples, include models for organisation aspects of ERP implementation [Molenaar 1999], functional process benchmarks for product development organisations in the automotive industry [Dijk 2001] and models developed for customer relationship management (CRM) and supplier relationship management (SRM) on a company wide scale [Wijk 2000, Mackor 2003]. The work of Leo van der Velde and Andy Schuurman on building models for supply chain management as a reference for a concurrent ERP implementation could be classified as an example of the 2nd category of Action Research as mentioned by Heller. The decision to implement ERP was already taken by the company. Our research group was invited to join the project to build a reference model concurrently with building the ERP prototype as a benchmark for testing future prospects for the organisation [Velde&Schuurman 1999][Velde 2002].

D.6 Conclusions

For engineers involved in business research, the use of survey methods as the starting point for research is against their nature. Engineers are trained to design useful theories and artifacts from scratch. The survey research model (figure D.5) puts the researcher only in the analysis role. The industry creates, the market validates and researchers try to understand afterwards what happened. Conclusions from such research are often either not specific enough to support better designs and better practices or the market conditions that provided the window of opportunity for industry is long gone when research finally understands the mechanism. Henry Mintzberg noticed this problem long ago:
"Most of the contemporary literature fails to relate the description of structure with that of the functioning of an organisation ... All of this is to say that the conclusions of the research often lack "context"- the type of organization and the part of it to which they apply, as well as the relationships between the structure and the functioning of the organization. As a result, these conclusions often come across to the

reader as detached from reality, devoid of real substance" [Mintzberg 1979, p.12-13].

In other words; one needs to be a designer or an engineer with the ambition to write a prescriptive theory to be fully aware of what part of the context is essential to do the research and to correctly frame the scope of a resulting theory. The action research model is closer to the nature of engineers since it allows to design solutions pro-actively for problems that are identified well in advance.

Placed in the more general framework of various kinds of abduction as presented by Schurz (table D.1), survey methods merely produce results that belong to the category of factual abduction. The known laws and models that drive factual abduction are all the result of reasoning schemes that fall in the category of either law-abduction or 2nd order existential abduction. Being primarily problem solvers, engineers in their role of scientists as well as innovative designers mostly benefits from 2nd order existential abduction reasoning schemes. The action research model in general, as well as the variant we have used ourselves also belongs in this category.

The methodological approaches defined by Arbnor (figure D.4) represent a categorisation of Theorems and Axioms that support the abduction scheme. Depending on the research goal, all categories of abduction may still be used.

D.6.1 This thesis

In Schurz classification of kinds of abduction, 2nd order existential abduction is mostly used in this thesis; especially analogical abduction. In the categorization of Arbnor, the work in this thesis is based on methods and tools from the systems approach, but not without taking into consideration theories and meta-theories that originate from the actors approach. For example, the chapter on boundary conditions in this thesis has been written because of robustness considerations from systems theory. The models and theories presented in the boundary conditions chapter are largely the result of research following the actors approach.

The focus in this thesis has been on developing models and theories that support individual businesses to design their specific processes and structures that could effectively and efficiently face the complexity of their own choice. This calls for a validation practice similar to that of validation of design methods. Validation of the tools and methods in this thesis is done through description of some realized designs, through making explicit the logic behind the ideas and through using a quantitative meta theory on complexity based on Shannon's entropy measure. In a recent article on the validation of design methods, validation practices are compared to methods common to medicine. As in medicine, the use of various models, logic and reasoning is advocated in the validation process before double blind field tests (like surveys) may be considered as the final step in a validation process [Frey 2006].

The type of understanding and systems modelling as developed in this thesis is a necessary precondition for doing surveys and compensating for the many side effects and biases that may be present in the answers. But even then, given the influence of specific strategic choices, past performance and experience on the complexity experienced by a specific company, it is doubtful whether a survey over a number of companies can produce comparable observations that may result in statistically significant results which could not have been found using other methods.

About the author

Bart Ruurd Meijer was born on August 14th, 1962 in IJmuiden, the Netherlands. After finishing his primary education at the Ichthus College in Driehuis in 1980, he started his study in Applied Physics at Delft University of Technology. He received his engineering degree in Applied Physics in 1987. His graduation project was on real time hardware for computer vision, a project that was done in cooperation with the Philips Research Laboratory and the Philips Centre for Manufacturing Technology.

After his graduation he joined the Control Laboratory at the Electrical Engineering Faculty at Delft University of Technology to work on building robot controllers for real time vision based tracking. This project soon became part of the DIAC (Delft Intelligent Assembly Cell) project. DIAC was a government sponsored project in which four faculties of Delft University of Technology participated, aimed at designing and building an automated assembly cell. From 1989 on he was the deputy project manager, responsible for integration and together with the project manager responsible for the project management and administration.

At the end of the DIAC project in 1992, he was appointed as assistant professor with the Flexible Production Automation group at the Mechanical Engineering faculty. The DIAC project taught two lessons that were determined for his work at the Mechanical Engineering faculty. First; despite the marvels of technology developed for DIAC, without hard figures on the business economics of DIAC, industry is not prepared to invest in flexible automation technology like DIAC. Second; all successful automation applications, including demonstrations done with DIAC, were based on well understood, massive scale, repetitive processes. The scale of operations is essential for the funding of engineering and testing to get the automation right. In 1994 he asked and got permission to work on a 50/50 basis for both the Flexible Production Automation group and the Industrial Organisation group that worked on business process and organisation design.

Between 1996 and 1999, his position as manager of the CIM centre of Delft University of Technology provided the platform to work with industry on the same integration issues. It soon became clear that the problems in industry with business process and organisation design were much bigger then the problems with automation technology. This provided the starting point for the research presented in this thesis.

Although until today his career has been in academics, the many opportunities to work with industry have provided him with the experience necessary to bridge the worlds of academics and industry. It is his ambition to continue this type of work especially in the domain of Research, Development and Innovation processes.