Cranfield University

Walter Van Dyck

Predictive Performance of Front-Loaded Experimentation Strategies in Pharmaceutical Discovery: A Bayesian Perspective

School of Management

Doctor of Business Administration

Cranfield University

School of Management

DBA Thesis

2004

Walter Van Dyck

Predictive Performance of Front-Loaded Experimentation Strategies in Pharmaceutical Discovery: A Bayesian Perspective

Supervisor: Prof Peter Allen

Academic Year 2003 to 2004

ABSTRACT

Experimentation is a significant innovation process activity and its design is fundamental to the learning and knowledge build-up process. Front-loaded experimentation is known as a strategy seeking to improve innovation process performance; by exploiting early information to spot and solve problems as upstream as possible, costly overruns in subsequent product development are avoided. Although the value of search through front-loaded experimentation in complex and novel environments is recognized, the phenomenon has not been studied in the highly relevant pharmaceutical R&D context, where typically lots of drug candidates get killed very late in the innovation process when potential problems are insufficiently anticipated upfront.

In pharmaceutical research the initial problem is to discover a "drug-like" complex biological or chemical system that has the potential to affect a biological target on a disease pathway. My case study evidence found that the discovery process is managed through a front-loaded experimentation strategy. The research team gradually builds a mental model of the drug's action in which the solution of critical design problems can be initiated at various moments in the innovation process.

The purpose of this research was to evaluate the predictive performance of front-loaded experimentation strategies in the discovery process. Because predictive performance necessitates conditional probability thinking, a Bayesian methodology is proposed and a rationale is given to develop research propositions using Monte Carlo simulation. An adaptive system paradigm, then, is the basis for designing the simulation model used for top-down theory development.

My simulation results indicate that front-loaded strategies in a pharmaceutical discovery context outperform other strategies on positive predictive performance. Front-loaded strategies therefore increase the odds for compounds succeeding subsequent development testing, provided they were found positive in discovery. Also, increasing the number of parallel concept explorations in discovery influences significantly the negative predictive performance of experimentation strategies, reducing the probability of missed opportunities in development. These results are shown to be robust for varying degrees of predictability of the discovery process.

The counterintuitive business implication of my research findings is that the key to further reduce spend and overruns in pharmaceutical development is to be found in discovery, where efforts to better understand drug candidates lead to higher success rates later in the innovation process.

ACKNOWLEDGEMENTS

I confess that completing a doctoral dissertation is a challenging but quite rewarding endeavour. Fortunately, before starting the voyage I was not really aware of the pains and pitfalls the process would entail. However, in retrospect they did not add up to the satisfaction of finishing a piece of research in an area which is pretty much my daily (and often nightly) life.

Needless to say that this work could not have been completed without the moral support, the contributions and critical questioning of many people that crossed my road in this endeavour.

First and foremost I would like to acknowledge the contributions made by my supervisor Professor Peter Allen, Head of the Complex Systems Management Centre at Cranfield University. His scholarship and directions in a domain as ambiguous and confusing as the Complexity Sciences were not just instrumental in bringing this piece of work to being. I experienced it as a fundamental necessity. Without Peter's support there would not have been a thesis to read.

Also, each of the other members of my panel made their own invaluable contribution to this work. Professor David Tranfield, Director of Research at Cranfield School of Management, opened my mind to the necessity of striving for both contributions to theory and practice by stressing the importance of Mode 2 research and inviting me to reflect on management theory as a design science. Dr Palie Smart, Senior Research Fellow, stressed the importance of looking at the problem in a holistic way, not just focussing on the quantitative decision-making aspects of the problem, and Dr Marek Szwejczewski, Senior Research Fellow, relentlessly reminded me of the methodological rigour required of a piece of academic research.

Other members of Cranfield University School of Management faculty I owe much to is Dr David Partington who taught me how to be a scrupulous proof-reader, critically questioning your own line of argumentation and thought. Also, a special thanks to Dr Nada Kakabadse. Her research philosophy course helped me structure my thinking.

Amongst the many PharmaCo scientists in Discovery, pre-clinical research, Cheminformatics and Chemical-Pharmaceutical Development I spoke to I would specially thank Dr Marcus Brewster, Dr Tina Arien, Dr Marc Vanstockem, Marc Deweer, Dr Marc François, Dr Marcel Michiels, Dr Jan Hoflack, Dr Claire Macky, Dr Michael Engels, and Vic Maes for providing me with deep insight into the inner working of the pharmaceutical discovery and development process as conducted in a world class research-based operation. A management researcher cannot dream of a better team to provide him with this level of empirical insight.

A warm and special thanks also to Frédéric Rolland, a young enthusiastic fellow consultant at the Life Sciences Strategy & Change practice of IBM Business Consulting Services. As a nuclear physicist and trained engineer Fred had no challenge in writing the Excel VBA program that underlies the simulation part of this thesis. Fred not only provided me with program code and endless hours of simulation run for my analysis and interpretation. He also acted as an ideal sparring partner to bounce of ideas and to challenge my emerging lines of thought based on the data the simulation model generated.

Combining the requirements of doctoral level management research with a daily life as a management consultant is not obvious, if not impossible without senior management support. Therefore, I would specially like to thank Guy Lefever, EMEA Industry Leader Life Sciences/Pharma of IBM Business Consulting Services who backed my weird research plans four years ago without hesitation, and supported me throughout the process.

Finally, last but not least I want to express my sincere gratitude for the moral support my wife Annemie continuously gave me to conduct this addictive work that has taken more than a fraction of our lives. Her understanding and encouragement made it possible that this piece of management research, frustrating at some times, could be brought to a fruitful end. My final thanks go to our two sons, Dominique and Jonathan, who both did not see their father playing much with them during the many study weekends of the last four years. I hope that when the time will have come that they will be able and willing to read this work, it may become a source of inspiration for their later professional and academic ambitions.

Affligem, May 2004
Walter M.J. Van Dyck

Notes on dissemination to date

The ideas expressed in this thesis are in the early process of being disseminated in the academic and practitioner literature. Drawing upon an Interpretive view of complexity, a first path to the innovation literature initiates theory building on the emergence of the solution to an innovation problem (Van Dyck, 2004b). A second path discusses the Bayesian view of the innovation process modelled as a complex adaptive system (Van Dyck, 2004c).

Furthermore, in an effort to join academic and practitioner thinking on the use of good practices to manage the innovation process, the concepts developed in this research have been connected to a recent discussion on the process of 'networking' as an underpinning organizational routine and dynamic capability (Smart et al. 2004). Finally, in a practitioner-oriented contribution, innovation process management ideas developed in this work were used to apply the real options view of valuating advanced technology to the world of pharmaceutical R&D (Van Dyck, 2004a).

Smart, P., Gupta, A., Bessant, J., Peakman, T. and Van Dyck, W. (2004) Networking for innovation: Exploring a dynamic capabilities perspective. *R&D Management* **Special Issue on Innovation and Intangible Assets**, Extended abstract submitted.

Van Dyck, W. (2004a) Introducing advanced technology in pharmaceutical R&D: The need for a new paradigm? In: Pacl, H., (Ed.) *The future of pharmaceutical R&D: Challenges and trends*, Frankfurt: Festel Capital, forthcoming.

Van Dyck, W. (2004b) Managing complexity in fuzzy front-end innovation: The need for a paradigm shift? *Cranfield School of Management Working Paper Series*, paper submitted.

Van Dyck, W. (2004c) Predictive performance of front-loaded experimentation strategies in pharmaceutical discovery: A Bayesian perspective. *Cranfield School of Management Working Paper Series*, paper submitted.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

NOTATION

α	Specificity or true negative rate
ADME-T	Absorption, Distribution, Metabolism, Excretion, and Toxicology
API	Active Product Ingredient; the active therapeutic agent in a drug product
B	Bio-availability or drug-likeness of an active drug substance
β	Sensitivity or true positive rate
χ	Test criteria measuring experimentation strategy quality
CVS	Cardiovascular safety
d^*	Optimal experimentation strategy
$\delta_i(t)$	Information value received by an adaptive system at time t
DE	PharmaCo pre-clinical department
DOE	Design Of Experiments
DSM	Design Structure Matrix
E	Environment against which a drug compound must be developed
FD	PharmaCo full (clinical and chemical-pharmaceutical) development department
FDA	Food and Drug Administration
FDP	Fraction of Dose Absorbed; a surrogate marker used in the pharmaceutical discovery research process
FIH	First In Human
h_i^+	Candidate drug compound structure hypothesized to be an innovative solution against a biological target
h_i^-	Candidate drug compound structure hypothesized to be inactive against a biological target
$\eta(t)$	Domain of action of an adaptive system represented as a set of candidate chemical structures at time t
H	Universe of potential compounds
H_i	A chemical class of compounds, a subset of H
HTS	High Throughput Screening
HTL	See H2L
H2L	Hit-to-Lead pharmaceutical discovery phase, also called HTL
I	Total range of signals receivable by an adaptive system
IND	Investigational New Drug
LO	Lead optimization pharmaceutical discovery phase
μ	Measure of performance against test criteria
NCE	New Chemical Entity
NDA	New Drug Application
NME	New Molecular Entity
P	Potency or biological activity of an active drug substance
$\Pi(t)$	Performance function of an adaptive system
π	Prevalence or fraction of really active compounds in the universe of potential compounds H
π^+	Positive predictive value

π^-	Fraction of really active compounds, given they were declared inactive y the experimentation strategy. (1-π^-) is defined as the negative predictive value
PharmaCo	Research-oriented global top-10 pharmaceutical R&D organization serving as an empirical base for this research
PD	Pharmacodynamic properties of a drug product
PK	Pharmacokinetic properties of a drug product
SAR	Structure-Activity Relationship
τ	Adaptive plan modifying a chemical candidate compound structure to become an active lead compound
T	Toxicity of an active drug substance
TI	Target identification
TV	Target validation
U(d)	Expected utility of an experimentation strategy
Ω	Set of operators used by an adaptive system to modify a candidate compound structure
ω	Operator part of a set Ω

1 Exploring experimentation strategies in Pharmaceutical R&D: Business rationale and roadmap for this research

1.1 INTRODUCTION

Experimentation is a significant innovation process activity and its design is fundamental to the learning and knowledge build-up process. Front-loaded experimentation is known as a strategy seeking to improve innovation process performance; by exploiting early information to spot and solve problems as upstream as possible, costly overruns in subsequent product development are avoided. Although the value of search through front-loaded experimentation in complex and novel environments is recognized (Verganti, 1999; Thomke and Fujimoto, 2000; Thomke, 2001; Thomke, 2003), the phenomenon has not been studied in the highly relevant Pharmaceutical R&D context, where typically lots of drug candidates get killed very late in the innovation process while potential problems are insufficiently anticipated upfront.

Therefore, since upfront problem anticipation in research possibly leads to better prediction of subsequent results in product development, the primary aim of this study is to investigate how Discovery Research –the fuzzy front-end of pharmaceutical R&D- can be managed for optimal predictive performance.

Considering the mandatory three-project structure of this doctoral research I started with an exploration of the differences between managing radical and incremental innovation projects leading to a proposed model of complexity-handling in innovation projects, situating the role of uncertainty and ambiguity in choosing for a specific complexity-handling mode. Then, a confirmatory case study was conducted in Discovery Research confirming the exploratory results at least for the front-end of the proposed model. However, to quantitatively explore predictive performance an appropriate problem representation framework needed to be chosen. Taking into account the inferential nature (Pearl, 2000) of the business problem of predictive reasoning and decision-making under uncertainty, a Bayesian framework (Jensen, 2001; Parmigiani, 2002) was constructed relating predictive performance outcome variables to explaining pharmaceutical Discovery Research policy variables derived from the confirmatory case study. Finally, using the developed Bayesian problem representation in a last project a Monte Carlo simulation was conducted to explore predictive performance of front-loaded experimentation strategies in pharmaceutical Discovery.

My research findings contribute to innovation management theory by extending the front-loading concept in scope and breath. More specifically, the theory-building effort of this study supports the view that front-loaded experimentation improves innovation process performance also in a pharmaceutical R&D context. The breath of

the concept is extended by proposing how front-loaded strategies can contribute to improved predictive performance of the innovation process, an outcome variable and causal relationship that has not been studied before. Second, I believe to contribute to the body of innovation management knowledge by proposing a simulation-based Bayesian inference framework to study predictive performance of innovation projects. Finally, the business implication of my research findings is that the key to reduced spend and overruns in Pharmaceutical Development is not only related to time-to-market reduction or efficiency enhancements, but is mainly to be found in Discovery, where efforts to better understand drug candidates are proposed to lead to higher success rates later in the innovation process.

In the remainder of this introductory Chapter I will review and analyse the pharmaceutical R&D business problem leading to this research. Then, the research context and basic literature domains providing a first handle to the business problem will be explored. Finally, after a discussion of my philosophical research positioning, a methodology and roadmap for this project-driven research will be given.

1.2 BACKGROUND AND RATIONALE FOR THE RESEARCH

1.2.1 Situating the business problem in Pharmaceutical R&D

Over the years the pharmaceutical industry has developed a very successful model for making new medicines. The world's leading R&D-based companies have collectively validated about 500 targets – the biological mechanisms (usually receptors or enzymes in human cells) through which drugs work (Drews, 2000). They have created large compound libraries, containing as many as two million molecules apiece. They have evolved a phased development process that includes large-scale clinical trials to establish the safety and efficacy of their products. During the past two decades, sales of new drugs and new formulations of older drugs have consistently outstripped the fall in income from products that have come off patent. However, for a variety of reasons the current model has recently come under growing pressure. Although recognized as a predominantly research-driven industry, one of the main reasons for failure is lack of R&D productivity. Indeed, the biggest drug makers need to produce three or more billion-dollar blockbuster drugs a year, just to maintain their sales growth. Yet they are far from fulfilling this promise (Arlington et al. 2002: 5).

Drug discovery and development is a lengthy and costly process. It takes on average 15 years and US$880 million to generate a successful NCE or new chemical entity (Tollman, 2001). A raise in productivity can be achieved by continuing efforts to speed up development and by eliminating weak projects as soon as possible. While resource usage increases considerably at each development phase, postponement of termination decisions not only wastes resources on redundant projects, it also denies resources to more successful projects.

Global project planning, realistic clinical study protocols, active collaboration with regulatory authorities, the use of project and data management and communication technologies, and project team cohesion and empowerment are cited by the fastest drug

development companies as the five practices that contribute most to reducing drug development cycle time[1] (Getz and De Bruin, 2000). Implementing heavy-weight project teams, 'disintegrated' from the traditional functional hierarchy (Case, 1998), leading the innovation process drawing upon the functional skills in the lines, is considered to be good practice. Also, managing the delicate balance between line and process is considered to be a contributor to project performance. Project performance is proposed to be higher when functional managers have greater influence over go/no-go decisions and project leaders have greater influence over clinical decisions (Basa, 1996).

However, too often the thrust of recommended change has been on process, rather than on science-based risk management of the Drug Development portfolio. Lesko *et al.* (2000) made a distillation of the outcome of two tandem practitioner conferences[2] discussing Drug Development optimization. They conclude that the underlying thesis of both conferences was that, in order to get to better therapeutic agents with lower development risks, the pharmaceutical R&D process needs to move from an essentially empirical mode to a more mechanistic and predictive one. The general goal should be to integrate early knowledge gained during Discovery into the Drug Development decision-making process. This approach would allow finding failures faster, resulting in more economical and informative development programs. Integrating Drug Discovery and Development processes 'provides a better understanding of the mechanism of drug action, suggests improved animal models to evaluate drug targets and drug-disease interactions, and helps to design animal experiments which provide more clinically useful information' (Lesko et al., 2000: 1336).

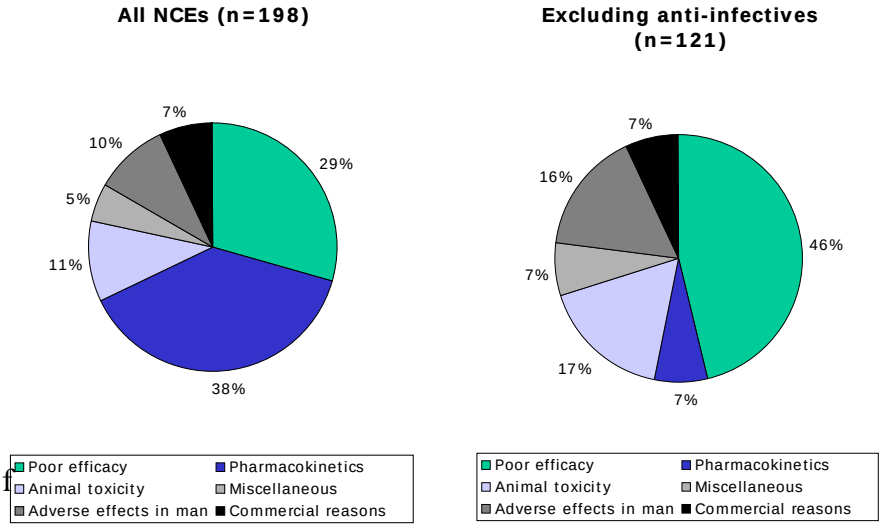


Figure 1-1: Reasons for failure in early development as cited in Kennedy (1997)

[1] Following Getz & de Bruin (2000) development cycle time is defined as the elapsed time from First in Human (FIH) studies to NDA submission.

[2] The two tandem practitioner conferences discussed are; "AAPS, ACCP, ASCPT, FDA Symposium on clinical pharmacology: Optimizing the science of Drug Development," held in September, 1998 in Arlington, Virginia, USA, and the second entitled "5th EUFEPS Conference on optimizing Drug Development: fast tracking into human," held in December, 1998 in Wiesbaden, Germany.

At a recent practitioner conference[3] it was acknowledged that innovation project risk in Pharmaceutical R&D is fundamentally *biological target related risk*. Late development problems are caused by poor target validation. Hitting the right target with sufficient efficacy and specificity is a challenge that can only be met if sufficient knowledge of the disease is built up during Discovery (Hopkins, 2003). During early development – before human testing- the biggest problem is target 'druggability', or the ability to find 'drug-like' chemical or biological structures, i.e. structures that are effective against the target of interest, but not toxic or non-absorbable by the human body. Industry-wide attrition during these early phases is 76%, mainly due –24 to 49% in Figure 1-1 above- to poor or inadequate bioavailability or ADME-T[4] properties (Kennedy, 1997). Hence, the major industry approach towards this involves moving ADME-T evaluations earlier on, before early Development, into Drug Discovery. Some even claim that these evaluations should be conducted in-silico in the early Discovery stages, during lead identification or optimization, in parallel with biological activity investigations (Pickering, 2001; Yu and Adedayo, 2003)

Discovery research practitioners develop a scientific viewpoint on a number of experimentation strategies available to run pharmaceutical discovery research to deal with these problems. To understand we need to discuss what this process looks like. Actually, once a biological target is identified, the Pharmaceutical drug discovery process aims to find a therapeutic agent with positive effect on this scientifically and commercially interesting target. It proceeds in essentially two stages. First, a lead molecule should be found in a diverse compound collection, constituting a chemical library. A typical major drug company will have hundreds of thousands to millions of compounds in its collection typically valued at 50-140 million dollars, or more (Young et al. 1997). Alternatively, combinatorial chemistry has come into recent use to create large collections of candidate compounds *screened* for their effect on the biological target (Thomke et al. 1998). Nowadays, robotized High Throughput Screens (HTS) are capable of handling about 10.000 chemical substances per day (Reiss and Hinze, 2000). The lead compound coming out of this screening process will be modestly potent. Therefore, in a second stage the candidate compound will be *optimized* by synthetically adding or removing parts through medicinal chemistry. This results into a candidate becoming more and more complex in structure throughout the discovery process. Essentially, 'lead-like' structures serve as initial starting points to be optimized into 'drug-like' leads (Oprea et al. 2001).

To run this *screening & optimization* drug discovery process as efficiently and effectively as possible, a number of fundamental questions need to be answered in a comprehensive experimentation strategy. First, as the number of both biological targets and compound libraries increase, there is a need for efficient screening strategies (Young et al. 2002). The optimal size of the screening library should be determined by

[3] Cambridge Health Institute (CHI) Intelligent Drug Discovery & Development 2003, May 28-29 2003, Philadelphia, PA

[4] ADME-T is an industry acronym for Absorption, Distribution, Metabolism, Excretion, and Toxicology. These are all 'bioavailability' properties of candidate drug-like structures indicating their propensity to be absorbed by the human body, and to gain a positive effect on it. A drug candidate can show good efficacy against a biological target –then it is said to show good bio-activity- but show poor bio-availability, hence having poor drug-like 'ADME-T' properties.

inventory costs and the cost of gain from synthetic modification. Typically about 10K compounds are screened for a minimal level of potency. Detection levels are set low enough to provide medicinal chemists with a reasonable choice of compounds to synthetically optimize but high enough not to waste efforts. Obviously, the break-even point between screening and optimization is sensitive to compound cost. To cite typical industry figures; if the cost of synthetic modification is as high as $6000 per compound and screening can be kept as low as $11.5 per screen, then it is economically feasible to screen over 1.4 million compounds before beginning lead optimization. Using combinatorial chemistry as an experimentation strategy can bring the cost of screening even more down making it economically justified to screen even more before switching to the significantly more expensive optimization (Young et al. 1997). Also, an operational choice needs to be made for a sequential versus complete screening effort. Clearly, if a target is valuable, fully validated and the therapeutic area market is large, complete screening will often be considered justified. Conversely, when a target is not tractable leading to large numbers of compounds for any hope of success, a sequential screening strategy is more likely to be chosen, whereby a relatively small set of compounds is initially screened and results are updated in an iterative process. Sequential screening is economically advantageous if the marginal cost of screening a compound is high (Young et al. 2002: 423). Methods for selecting subsets of molecules from large compound libraries are known and can be applied to very large chemical databases (Higgs et al. 1997). More and more, 'wet' screening efforts are complemented with 'in-silico' searches in large chemical computer databases. However, the reality of the complex drug-receptor interaction makes these new technologies having only partially delivered their promises and still leads to heavy debates among experts (Oprea , 2002).

Second, a discovery experimentation strategy should consider the conversion from 'lead-like' to 'drug-like' candidate compounds. Now, the emphasis is on optimization and selection. Less complex molecules coming out of screening are good starting points for optimization and eventual discovery of potential drugs (Hann et al. 2001).

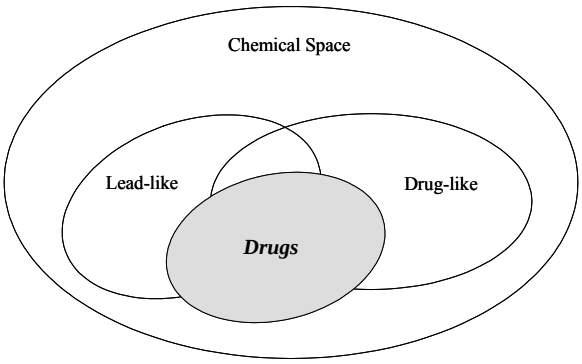


Figure 1-2 Venn diagram of the medicinal chemistry space related to drug discovery[5]

[5] Adapted from figure 8 page 1314 of (Oprea et al. 2001)

To optimize a lead-like structure into a potential new molecular entity (NME) to be transferred to subsequent Clinical development, potency needs to be increased and the structure must show to have good ADME-T properties. Optimization consists of generating chemical analogues within the chemical solution space until a drug-like structure and ultimately a drug candidate appears[6]. An experimentation strategy used for running this optimization part of the discovery process has to answer the fundamental question of the level of knowledge that needs to be built up before transfer into pre-clinical development is considered; in the *Old paradigm*, applied in the industry some time ago, discovery research was only concerned about biological activity or potency and selectivity of a candidate compound. Drug-likeness properties were only taken into account from pre-clinical research on[7].

Nowadays, *Front-loaded discovery* combines in-vitro screening with in-silico data (Pickering, 2001; Coty, 2002) on drug-likeness of candidate compounds and selects the most promising ones based on the fullest multi-factorial –potency and drug-likeness- picture available. Now, drug-likeness problems are identified upfront and can be anticipated during the multi-factorial lead optimization process before transferring the compound to pre-clinical development. Still, a debate prevails around the question how early one can start computer assisted front-loading. Influential experts warn against '…uncritical application of high-throughput [screening] in-silico methods. Structure-based and computer-aided approaches can only be as good as the medicinal chemistry they are based on. The search for new drugs, especially in lead optimization, is an evolutionary process that is only likely to be successful if new methods merge with classical medicinal chemistry knowledge' (Kubinyi, 2003: 665).

Another element of the lead optimization strategy is the number of compounds one promotes to optimization after having passed initial screening. Knowing the chemical cost of getting a developmental compound typically amounts to about $4,500,000 (Young et al. 1997: 893) this is a decision not to be taken light-heartedly. Nevertheless, practitioners argue for more generous promotion of screened compounds for further study and in-vivo optimization. While more expensive, they claim it will lead to a better understanding and hence will lead to better drugs being transferred into pre-clinical development (DeWitte, 2002).

Summarizing, an extensive practitioner literature exists providing viewpoints on experimentation strategies for pharmaceutical Discovery. The business problem identified as the starting point for this thesis is the requirement for more predictive experimentation strategies in Discovery leading to less wasted efforts in subsequent Development. As exemplified above, practitioners have started a debate about the benefits of front-loaded experimentation and more extensive lead optimization efforts. Both strategies are hypothesized to increase predictive performance of discovery

[6] Unfortunately, the complexity of nature sees to it that some drugs are part neither of the drug-like nor of the lead-like medicinal chemistry space (see Figure 1-2); 'compounds derived from natural products, and proteins are currently marketed as drugs but are not representative for the types of drugs that are expected to come out of medicinal chemistry efforts' (Oprea et al. 2001: 1314). This explains part of the drug-space overlapping both lead- and drug-like spaces.

[7] For an idealised *'Old paradigm'* project flow chart in a typical Pharmaceutical Discovery operation see Cavalla (1997)

experimentation since they increase the level of knowledge on the candidate compound before its transfer into subsequent pre-clinical development, maximizing avoidance of unanticipated problems. However, all of these viewpoints are case-based and lack the rigour of management and decision sciences research methods to make them more conclusive about business risk or predictive performance. Therefore, this should be the focus and contribution of my research.

In the following sections I will review two relevant literature domains needed to transform the business problem into a researchable question; experimentation behaviour, and the Bayesian-oriented literature on predictive reasoning under conditions of high uncertainty. By applying the latter to the stated business problem I will develop notation and an argument as to how one can systematically and quantitatively reflect on the predictive performance of decision-making following pharmaceutical Discovery experimentation efforts.

1.2.2 Exploring experimentation behaviour

Experimental design or experimentation strategy is the general approach to planning and conducting experiments (Montgomery, 1997). Experimentation is a significant innovation process activity and its design is fundamental to the learning and knowledge build-up process. Experimentation, a form of problem-solving, lies at the heart of every company's ability to innovate. Therefore, 'enlightened' and profitable experimenters organize for rapid experimentation using computer simulation, experiment with many diverse ideas and fail often, and they use front-loaded development, exploiting early information to spot and solve problems as upstream as possible (Thomke, 2001). Also, West and Iansiti (2003) acknowledge the vital role and value of search through experimentation in complex and novel environments. Their study of the semiconductor industry provides evidence that generation of knowledge through experimentation and retention of knowledge through experience are significantly correlated with performance, whereas other -previously cited- measures of R&D commitment and organization were not.

Experimentation is an iterative process with product developers going through a series of design-test cycles, called experiments, homing in to a solution to their innovation problem (e.g. Whitney, 1990). A subset of pre-screening experiments typically precedes a more formal "design of experiments" (DoE) set to exhaustively test all explaining factors and their impact on the objective variables (e.g. Montgomery, 1997; Yang and El-Haik, 2003). Concurrent engineering is the predominant paradigm used to compress the experimentation process. Formal models have been designed to optimize concurrency schemes (e.g. Krishnan et al. 1997; Loch and Terwiesch, 1998; Roemer et al. 2000; Mihm et al. 2003). A Design Structure Matrix (DSM) describes sequential iteration by computing the expected duration of the iterative solution process, and suggests an initial ordering of the coupled design tasks to minimize expected duration (Smith and Eppinger, 1997a). The work transformation matrix model is an extension to DSM, which can be used to predict slow or rapid convergence of iteration while experimenting (Smith and Eppinger, 1997b).

Experimentation behaviour research focuses mainly on strategies accelerating product development lead-time and enhancing efficiency. Examples include survey-based research showing that experiments can be conducted in different modes and that managers can determine optimal switching points between modes to reduce development time and costs (Thomke, 1998a). Empirical results in automobile industry tests for crashworthiness showed that developers can increase the frequency of problem-solving cycles while reducing the total R&D budget by speeding up and simultaneously reducing cost of design iterations through computer simulation and rapid prototyping (Thomke, 1998b). Or, in a study exploring the impact of different learning strategies on development performance in pharmaceutical and biotech development, results indicate that learning by doing is essential for efficiency in biotech while, in contrast, in pharmaceuticals learning before doing is preferred to model future production experience (Pisano, 1994; 1996; 1997). Thomke *et al.* (1998) offer a detailed case study of the impact of combinatorial chemistry on the pharmaco-economics of the drug discovery process. A recent simulation-based study built a mathematical model of a complex product development project requiring intensive communication among its many interdependent actors (Mihm et al. 2003). Their model highlights that during development design oscillates through numerous design solutions before converging to a final solution. As the problem size grows, time to conversion also grows exponentially, negatively impacting development lead time. Mitigating experimentation strategies countering the oscillation problems are proposed including modularization, immediate communication, and exchanging preliminary information.

Front-loading is a key theme in the experimentation literature. Verganti (1999) acknowledges the central role of the early phases of product development. Based on in-depth case studies of 18 Italian and Swedish companies active in helicopter, vehicle, and white goods sectors, he sees early analysis and problem solving as a difficult task because detailed insights are not available until one gets into detailed design. His data show that companies are locked in a dilemma between problem anticipation and reaction, the latter being defined as delaying decisions to downstream phases where information will come available. He argues that neither anticipation nor reaction should be considered as best practices. Rather, they should strongly interact in a mechanism he calls ‘planned flexibility’ i.e. the capability to build flexibility into the development process due to decisions taken early on. Planned flexibility deals with uncertainty through early identification of specific critical areas of a given project and early planning for reaction measures. Triggered by Boeing’s and Chrysler’s experience with the use of digital mock-ups identifying interference problems that are very costly to solve if only identified further downstream the development process, Thomke and Fujimoto (2000) conducted a field study at Toyota. Their case evidence showed how systematic efforts to front-load the development process have shifted problem identification and problem solving to the earlier stages. It led to efficiency and lead time improvements, giving field support for front-loading as an important methodology to accelerate and improve development performance. Survey-based research of 29 Internet software development projects (MacCormack and Verganti, 2003) also provides support for the front-loading benefits argument. Here, for projects facing greater uncertainty, early technical and market feedback had a stronger association with performance. While greater uncertainty was associated with making later changes to the product design, this practice was not associated with performance. Finally, in the late nineties also

pharmaceutical and biotech companies discovered the benefits of finding potential failure modes as early as possible in the development process. In a case study of Millenium Pharmaceuticals it is shown how new technologies for experimentation can form the basis for fundamentally rethinking the innovation process by shifting failures to earlier phases (Thomke, 2003).

Parallelism or the number of alternative approaches explored to solve a problem is known to be related to the quality of the solution. Allen (1966) found that development groups producing higher rated solutions generated fewer new approaches during the course of the project. However, later research (Abernathy and Rosenbloom, 1968) showed it is common in technological development to explore several approaches so that the best approach can be chosen. Since the outcome of any approach is uncertain it was found to be difficult to choose the best one early in the process. The authors argued that, based upon 14 case studies of development projects, a parallel experimentation strategy provides tangible economic benefits including the value of providing information on the choice of an approach, its value as a hedge against failure, and its value as a means to enhance useful competition. More recently, problem-solving efficiency in complex and novel environments is associated with a broad exploration and framing of the solution space, reaching across multiple knowledge domains (Schrader et al. 1992; McDonough and Barczak, 1992; Iansiti, 1998). Toyota tries not to converge too quickly on a "best guess" solution for their new designs. Instead, Toyota has a high regard for learning on multiple ideas in parallel. 'Toyota seems to value highly the reassurance that the chosen solution truly is the best and deems it worthwhile to spend the resources for that assurance' (Sobek II et al. 1999: 75). This results in engineers talking about sets of ideas and regions of the design space, not about "the" idea. In fact, the authors claim to increasingly hear of US companies successfully using this 'set-based' approach to product development. The process enabled innovation teams to avoid rework and meet aggressive time lines.

It is acknowledged that increasing parallelism also increases the need for enhanced coordination and preliminary exchange of information between concurrent development tasks. In an effort to study the fundamental drivers of parallel and sequential testing strategies, Loch *et al.* (2001) develop optimal experimentation policies. Their model analysis shows that the optimal mix of parallel and sequential testing depends on the ratio of the financial cost and the cost of testing. More expensive tests make sequential testing more economical. Second, imperfect tests decrease the attractiveness of parallel testing strategies. Third, they show that modularising product architecture can radically reduce testing costs. Based on case studies of five engineering problems, recent research concludes that set-based coordination requires the absence of ambiguity, and should be emphasized if either starvation costs or the cost of carrying multiple design options in parallel are low (Terwiesch et al. 2002). A recent analysis of NPD decision-making based on a pharmaceutical product development case, developed closed-form solutions for the optimal number of concept tests to be conducted under profit uncertainty (Dahan and Mendelson, 2001). Their key finding was that the optimal number of concept tests depends not only on the cost of testing and the scale of uncertainty, but also on the distribution shape of that uncertainty. Also, these authors value the option of retaining flexibility at the concept selection stage.

Finally, also simulation is used to study the complexities of experimentation. Mena *et al.* (2001) use a Genetic Algorithm-based model to analyze the validity of the evolutionary analogy for product development, and to understand how the concepts of evolution and learning can be used to improve the product development process. They found that diversity of designs explored and process duration are by far the most important parameters in finding a satisfactory solution while other factors were found to be negligible.

Concluding, the literature on experimentation behavior complements –a.o. simulation-based- formal models with empirical research to advance theory. With the exception of Thomke *et al.* (Thomke et al. 1998; Thomke, 2003) empirical case-based research is conducted in mainly discrete manufacturing technology-intensive industries, not in pharmaceutical or biotech companies. Hence, to replicate techniques and conclusions of present experimentation research to answer the business problems stated above, they will need to be tested for their applicability in the pharmaceutical sector.

Second, from the above and confirmed by the results of a recent literature review and meta-analysis on relationships between integrated product development (IPD) characteristics and project performance (Gerwin and Barrowman, 2002), I conclude that present experimentation behavior research emphasizes strongly the study of the impact of various experimentation strategies on development lead time and efficiency, performance variables that are typical for product development and innovation research. No study was found to contribute to the exploration of *predictive* performance of experimentation strategies. Since the latter will be essential to frame the business problem and to convert it into a research question, this will be the topic of the next section.

1.2.3 Predictive performance of front-loaded experimentation strategies: The need for Bayesian thinking

As discussed above, pharmaceutical Discovery can be thought of as a screening and optimization experimentation process. To explore the business problem of finding ways to increase the predictive performance of this process, a quantitative framework is needed modelling the relationship between experimentation results and ensuing decisions. In deciding whether a candidate compound is presumed to be active or not, scientists reason under high uncertainty and ambiguity. There are several approaches to model reasoning under uncertainty. The approach I will adopt in this thesis draws upon probability theory as applied to decision sciences; decisions are made "given" the result of an intervention -a set of experiments-, which makes it *conditional* probability thinking. 'Probabilistic relationships, such as marginal and conditional dependencies, are helpful in hypothesizing initial causal structures from observation' (Pearl, 2000: 25), which will be the subject of my research.

To introduce the decision sciences part of my thesis, in the remainder of this section I will model a scientist's decision-making process as a normative system, maximizing expected utility, excluding subjectivity. The aim of a normative system is to take decisions by exploiting accumulated knowledge and by exploring experiences.

Typically, then, 'actors will be engaged in the following types of actions; using observation to interpret a situation, focusing a search for more information, deciding for intervening actions, adapting to changing environments, learning from experience' (Jensen, 2001: vi). Now, to analyze the quality of an experimentation strategy to discover a really active[8] chemical structure in the Drug solution space (the "Drugs" Venn diagram in Figure 1-2) I propose to do the following thought experiment. If we indicate by π the fraction of really active compounds in this universe of potential compounds, the odds[9] of finding at random an active chemical structure equal $\pi/(1-\pi)$. Since the universe of potential active structures –found through a process of screening and optimization- approaches infinity, this is a very small number approximating zero. Any experimentation strategy replacing the random draw from Nature by a screening & optimization effort should do better. To measure how much better, the following relationship can be defined where π^* is the fraction of really active compounds, delivered by the experimentation strategy;

$$\frac{\pi}{1-\pi} = k.\frac{\pi^*}{1-\pi^*} \qquad \textbf{(1-1)}$$

Clearly, k indicates how much the experimentation strategy is better than a random draw from the drug solution space. To operationalize this measuring concept we draw upon Bayesian inference logic. In Bayesian terms the above equation would read that the prior odds of finding an active structure equal the posterior odds times a weight of evidence or Bayes factor. This way, equation 1-1 can be converted in Bayes' celebrated inversion formula (see a.o. Pearl, 2000; Parmigiani, 2002);

$$p(H \mid e) = \frac{p(H)\,p(e|H)}{P(e)} \qquad \textbf{(1-2)}$$

which states that belief we accord a hypothesis H upon obtaining evidence e can be computed by multiplying our previous belief $p(H)$ by the likelihood $p(e \mid H)$ that e will materialize if H is true. Thus, applying this thinking to our situation, it models the experimentation strategy as a learning process that modifies one's initial probabilistic belief about the prevalence π prior to observing experimentation outcomes to updated or posterior knowledge incorporating both prior knowledge and the data at hand (Congdon, 2001). The data at hand or evidence is the decision made after Discovery experimentation; a chemical structure can be declared to be active, which I will denote as H_i^+, or inactive, denoted H_j^-. Only after clinical development will it be confirmed whether the hypothesized structure is really active, denoted C_i^+, or not, denoted C_j^-. Using this notation, to evaluate predictive performance of a Discovery experimentation strategy, the question must be answered: What is the probability that a structure H_i^+ hypothesized to be active, is *really* active denoted as C_i^+? The *positive predictive value*, then, denoted $p(C^+ \mid H^+)$ or π^+ is read as the probability that a compound will pass

[8] '*Really active'* is being defined as passing Discovery and subsequent Clinical development testing

[9] The odds of an event are defined as the ratio of the probability of the event happening to the probability of it not happening. Odds can be any positive number. Odds of 1 or 'even odds' correspond to a probability of 0,5 (see a.o. Parmigiani, 2002: 11)

pre-clinical and clinical tests $p(C^+)$, given it has been declared active $p(H^+)$ by the Discovery experimentation strategy. Similarly, the probability that a compound will not pass pre-clinical and clinical testing $p(C^-)$, given it has been declared inactive $p(H^-)$ by the experimentation strategy, is called the *negative predictive value*, and denoted $p(C^- \mid H^-)$. Knowing that the fraction of really active compounds, given they were declared inactive by the experimentation strategy is called π^- or $p(C^+ \mid H^-)$, the negative predictive value is denoted as $1-\pi^-$ (Parmigiani, 2002).

Conversely, an experimentation strategy featuring high positive and negative predictive performance, then, has a high value for π^+ and a low π^-. The transition from π to π^+ and π^- models the learning made by the experimentation strategy about the true status of the universe of potential compounds H. Using the Bayesian logic set out above, it quantifies how inferences about the universe of potential compounds H are updated in the light of new evidence, provided by the experimentation strategy.

Now, to formally state the problem of an R&D manager acting as a normative system having to select between experimentation strategies, the optimal decision d^* can be modelled as follows (see a.o. Müller, 1999);

$$d^* = \arg\max_{d \in D} U(d) \quad \text{where} \quad U(d) = \int u(d,\theta,y)\, p(\theta) p_d(y \mid \theta) d\theta dy \qquad \textbf{(1-3)}$$

U(d) is the expected utility of an experimentation strategy d, an element of the universe of possible experimentation strategies D. The utility function $u(d,\theta,y)$ is in our case specified by solving a decision tree of the outcomes of the various (H,C) combinations. Then, in Chapter 3 it will be shown how ordering these combinations in a decision tree, and based on cost assumptions for each experimentation strategy a financial outcome can be calculated and used for comparison to make an optimal decision d^*. To solve U(d) a statistical experiment needs to be set up to calculate a model $p_d(y \mid \theta)$, a distribution of observables y conditional on prior distribution $p(\theta)$. In our case, this model is not analytically solvable implying the need for numerical solution strategies (Müller, 1999). This is one of the reasons why in Chapter 4 Monte Carlo simulation (see a.o. Critchfield and Willard, 1986; Müller, 1999) will be used to explore the research questions and build theory on predictive performance of experimentation strategies. The other reason why simulation is used as a research methodology is given by the fact that in our case parts of the outcomes of the decision tree are not observable in practice, preventing the use of empirical methods like surveys.

Simulation for theory-building has been used before as a research strategy to study 'garbage can' decision making processes (Cohen et al. 1972; Masuch and Lapotin, 1989) Applying Bayesian inferential thinking in a context of problem-solving, experimentation, and ensuing decision-making is a very recent trend in the literature. Pearl (2000) presents and unifies the probabilistic, manipulative, counterfactual, and structural approaches, and he devises mathematical tools for studying the relationships

between causal connections and statistical associations. In their seminal work on causation, prediction, and search, Spirtes *et al.* (2000) axiomatize the connection between causal structure and probabilistic independence and explore several varieties of causal indistinguishability. Jensen (2001) suggests Bayesian networks and decision graphs to model causal impacts between events. He distinguishes between test decisions to look for more evidence to be entered into the model, and intervening and non-intervening action decisions, which force a change of state for some variables in the model, or not. Recent work (Pelikan and Goldberg, 2003) proposes an optimization method called the hierarchical Bayesian optimization algorithm incorporating three important features for robust and scalable optimization of complex problems; proper decomposition, chunking, and preservation of alternative solutions along the problem-solving and experimentation process. Their computational method, called the hierarchy machine, uses several concepts from genetic and evolutionary computation like population, selection, exploration via recombination (Goldberg, 1989; 2000; Mitchell, 2001), combined with Bayesian or 'belief' networks. Finally, other recent work (Callan, 2003) applies causal thinking to automate belief network-based problem solving in the context of Artificial Intelligence (AI).

Summarizing, this preliminary literature review gave us the handles to convert the business problem of the requirement for more predictive pharmaceutical Discovery experimentation strategies, into a researchable question. Advocated experimentation strategies like front-loading or parallel concept exploration, as applied in other technology-intensive sectors, need to be tested for their predictive performance in pharmaceutical Discovery. The literature review showed us this was not done before. This section introduced us to a Bayesian simulation-based probabilistic framework designed to explore this phenomenon in a quantitative way.

1.2.4 Research question and purpose

To guide my research efforts I transformed the pharmaceutical R&D business problem into the following more generally applicable research question: *How to increase predictive performance of the fuzzy front-end innovation process?* With Kim and Wilemon I define the fuzzy front-end of the innovation process as 'the period between when an opportunity is first considered and when an idea is judged ready for development' (Kim and Wilemon, 2002). In the fuzzy front-end fuzziness or ambiguity about the performance of the idea prevails, preventing it from being transferred to development. Resolving this ambiguity by clarifying the product concept and by identifying and solving problems upfront is a prerequisite for a cost-efficient development phase, where it is costly to rework or kill non-performing product ideas. The literature review above shows the relevance of researching experimentation behaviour in this front-end innovation process phase. Only, until now the focus has been on cost-efficiency and time-to-market performance, not on the systematic study of predictive performance.

The purpose of my research, then, is to advance new process theory explaining the mechanisms –or the *How?* and *Why?*- of fuzzy front-end experimentation behaviour leading to increased predictive and business performance. Also, my theory-building

effort should fit the purpose of mode 2 management research (Gibbons et al. 1994) to design theories which have the potential to change R&D managers' behaviour. The product of this research, then, will be a propositional model and a technological rule, the latter being defined as '[a] chunk[s] of general knowledge, linking an intervention or artefact with a desired outcome or performance in a certain field of application' (van Aken, 2004b: 228). In other words, the main purpose of this research is not to generate universal laws but a general prescription for a class of problems being limited to a certain field of application. Technological rules resulting from this research should stand the test of descriptive and goal relevance, and be operationally valid (van Aken, 2004b). In effect, external validity, the extent to which the defined rules really matter for practice, and the extent to which a practitioner can effectively control the independent variables in the model will be the true test of practitioner relevance.

As depicted in Figure 1-3 below, Van de Ven and Poole *et al.* (Van de Ven and Poole, 1995; Poole et al. 2000) provide four archetypal theories explaining processes of change like innovation; life cycle, teleology, dialectics, and evolution.

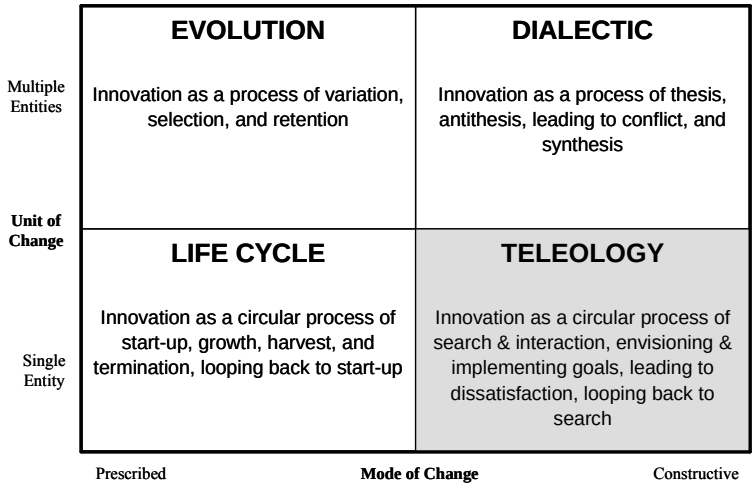


Figure 1-3 Archetypal theories explaining innovation processes[10]

These four ideal types are composed of distinctive event sequences and generative mechanisms the authors call "motors", that explain how and why changes unfold, which makes them an ideal starting point for my research. They introduce two analytical dimensions to classify the four ideal types of theories; the unit and mode of change. Evolutionary and dialectical theories operate on multiple entities like populations. Conversely, since the level of analysis of my research is the innovation team carrying out actions like experimentation and decision making, I will focus on a *single entity* as the unit of change studied. The mode of change studied distinguishes between change processes unfolding following a prescribed or constructed emerging pattern. Evolutionary and life cycle theories operate on prescribed modes of change. Dialectic and teleological theories follow a constructive mode of change.

[10] Adapted from (Poole et al., 2000); Figure 3.1 - 'Typology of organizational change and development theories', p 66, emphasis added.

Since the unit of analysis of my research is the experimentation and decision making carried out in the fuzzy front-end part of the innovation process, which tends to be *constructive* in nature, and since I don't focus on the dialectics in the innovation team, the advanced theory will be *teleological* in nature. A teleological theory assumes that the innovation process proceeds towards a goal, is carried out by a purposeful and adaptive entity –the innovation team-, 'by itself or in interaction with others, constructing an envisioned end state –through a process of experimentation and decision making-, taking action to reach it, and monitoring its progress' (Poole et al. 2000: 61).

1.3 SUMMARY OF THE RESEARCH PROCESS

1.3.1 A Critical Realist ontology

Building teleological process theory in my research context entails probing into scientists' experimentation behaviour, making sense of the complexity they are facing when trying to solve their innovation problem; the goal of their endeavour. However, the sources of complexity I want to address in my research don't reside in the science or technology and its interconnectedness with its constituent components or target application environment as described by a.o. Kim and Wilemon (2003). Instead, I want to explore complex ways of thinking of the complexity of experimentation behaviour. Doing so, I enter the domain of second-order complexity (von Foerster, 1973), or the domain of the 'thinker thinking about complexity' (Hatch and Tsoukas, 1997). Therefore, ontology is needed to accommodate my objective of developing process theory while acknowledging the lack of absolute causal certainty when dealing with 'complexity-handling' processes. Hence, my research will need to be built in a theory of reality allowing for 'explanations of cause and effect which: (1) exist in the form of "mechanisms" which may not be consciously perceived by research subjects or theoretically preconceived by researchers, which therefore may act independently of thought and which are only accessible through the creative speculation by the researcher of plausible alternatives whose "truth" is ultimately dependent on consensual validation by informants; (2) from data which do not necessarily explicitly link the elements of the paradigm model; (3) from data which are not based on direct observation of the researcher' (Partington, 2000: 98).

Two reasons make me conclude that the logico-scientific approach of Critical Rationalism (Popper, 1959) does not provide an adequate ontology to study this second-order complexity domain. First, its critical attitude of inventing theory through a tentative process of conjecture and refutation in pursuit of the absolute truth does not lend itself to the study of the phenomenon since the connection between predictive performance and the innovation team's teleological process is not reducible to absolute causal laws. Instead, I take the position you cannot go further than taking the Conventionalist stance (Kuhn, 1996) that the advanced theory will not be true or false but will 'only [be] useful in solving puzzles which it defines, using criteria which it specifies. Truth becomes a matter of community consensus' (Blaikie, 1993: 108). Also, while no theory exists explaining the impact of experimentation behaviour on predictive performance, theory construction as opposed to theory testing is the most adequate way forward. The approach followed in this thesis, then, will subscribe to the Realist

constructive empiricist perspective that theory construction 'is about constructing hypothetical models of mechanisms and then endeavouring to demonstrate their existence' (Blaikie, 1993: 115). The focus will have to be on the empirical adequacy of the developed theory and its capacity to explain the observed. Now, as the theory will describe more than what is observable considering the second order complexity nature of the phenomenon, I take van Fraassen's position that what matters is empirical adequacy and not 'discovery of truth concerning the unobservable'(van Fraassen, 1980: 5). Then it becomes possible 'to accept a theory as empirically adequate without believing it to be true' (Blaikie, 1993: 115).

Second, in falsifying hypotheses the Critical Rationalist School only considers the observable and takes the socially constructed world for granted, meaning they ignore the sensemaking mechanisms that lead to the observable reality. However, precisely these mechanisms are the subject of my research interest. Therefore, the phenomenological emphasis of Interpretativism that reality is socially constructed admitting that person and world are inextricably related through persons' lived experience of the world (Sandberg, 2000) seems to be better suited to explore the research questions. Within the Interpretative tradition the Realist perspective considers science to be concerned with what kind of things there are and how these things behave. 'It is concerned with a reality that is claimed to exist and act even if it has not yet been observed, and this reality has a life of its own apart from the activities of science' (Bhaskar, 1986: 5). Critical Realist ontology is concerned with a reality that is claimed to consist of three overlapping domains: the experiences of the *empirical*, the events of the *actual* and the mechanisms of the *real*. Hence, this philosophy of science fits my purpose of enquiry since its multi-level ontological perspective acknowledges the existence of the real domain or the generative mechanisms that produce observable events of the empirical world. Researching mechanisms of experimentation behaviour then, is formulating causal relations regarded as powers or tendencies of things that interact with other tendencies in the *'real'* so that an event in the *'actual'* may or may not be produced, and may or may not be observed in the *'empirical'* domain. This can be illustrated by an example in my research; the dimensions of the complexity experienced by the scientists in solving their innovation problem may lead to a specific way of organizing the innovation process which may lead to observable and non-observable predictive performance. Conversely, in the Positivist view causal laws are regarded to be universal constant conjunctions between events always producing the same observable outcome.

1.3.2 An interpretative epistemological positioning

A Critical Realist epistemological positioning entails an Interpretative epistemological positioning taking a view of reality as a social construction, embedded in a context. The task of my research is to go beyond description into explanation. Therefore, *Abduction* is the appropriate research strategy (Blaikie, 1993: 163); it is the process used to construct theory within an Interpretive approach, occurring in a context of ontological, conceptual and theoretic assumptions. Since research on experimentation has been carried out in related environments I will build on existing theory. Hence, I don't want to start from a blank slate as suggested by pure Inductivists to construct theory. Neither

do I want to start from a hypothetical model of actually existing entities and their hypothesized relations to be tested as suggested by a pure Reductionist strategy since innovation process-predictive performance relationships have not been hypothesized before. Then, Retroduction, or initial hypothesis formulation (Blaikie, 1993: 164), will not be the first step of my investigation. Instead, as suggested by Bhaskar (1986) my research will start in the domain of the 'actual', with observed connections between phenomena and causal links between computer-simulated variables. Describing actual activities and meanings derived from social actors will be my starting point. Then the existence of structures and mechanisms of the 'real' will be postulated to explain the actual phenomenon. The last step of my logic of discovery is to prepare for empirical observation of the acclaimed effect of the postulated mechanisms in the 'experienced' by formulating a prescriptive framework and a propositional model derived from computer-based simulation, to be tested by subsequent empirical research.

The experimentation and decision making process conducted by the innovation team, my unit of analysis, is the focal point of my theory-building effort. Considered a constructor of reality in an ambiguous technology-intensive environment, engaged in complexity-handling behaviour, I will model them as a *socio-technical system*. The latter consists of technical and social elements whose interrelation can be analyzed at various levels (Molina, 1990; Klaes, 1997). The focus of the socio-technical system in my research is the experimentation and decision making process it is carrying out. To structure theory-building using an abductive strategy I will use Klaes' (1997) analytical layers, distancing itself from methodological reductionism and providing an elegant categorization to structure the complex interplay of technology, process, and human behaviour within the socio-technical system. The following analytical layers or 'constituency levels' applied to my research context are distinguished: (1) the physical level of the technology matching a biological receptor with a chemical active structure, (2) the level of the basic financial, human, material, time, and space resources directly contributing to the materialisation of the experimentation and decision making process; the innovation team, (3) the intra-institutional level or the specific mix of basic resources together with the interplay of individuals organizing the basic resources for level 2; the R&D organization of which the innovation team is part, (4) the inter-institutional level describing the interaction of various institutions and social groups of level 3 fostering the development of the innovative solution to the problem, (5) the inter-constituency level, at which the socio-technical system under analysis interacts with other socio-technical systems carrying out the same task but on a different innovation problem, (6) the level of the historical context in which the socio-technical system is situated. Taken together, the analytical layers of the socio-technical system serve as a descriptive device. Any claims for explanatory powers rely on the addition of propositional frameworks relating elements within or across these layers.

To probe into the generative mechanisms connecting the socio-technical system's innovation process to predictive performance a three-phased research strategy theorizing from process data was designed featuring process grounding, computer-based simulation, and theory development, depicted in Figure 1-4.

In a first *process grounding* phase, empirical adequacy will be obtained by taking recollections of scientists' experiences involved in experimentation and decision

making activities as the starting point of my enquiry. This will be followed by reconstruction of their sensemaking behaviour into social scientific language. This means scientists will be interviewed to catch their process of interpretation by which they make their choices for this or that experimentation approach, or make this or that decision. The interview data will be based on recollected accounts of pharmaceutical innovation projects, which implies I will act as a detached observer. Also, in order to improve my understanding of the process and to capture the full richness of the data I will use different participant recollections (Schwenk, 1985) of each innovation project. Visual mapping and temporal bracketing (Langley, 1999) will be used as research methods to structure the first-order lay accounts. Visual mapping will provide a first step between the raw process data and a first attempt to abstract conceptualization of the innovation processes of all studied innovation cases. Temporal bracketing will summarize the different case visual maps in one structure of more discrete and connected blocks that can form the basis for replication and explanation. Grounded Theory (Glaser and Strauss, 1967), although with the potential to provide very accurate rendering of activities and meanings since firmly grounded in the raw data, was not chosen. The aim of this research is to generate formal theory at a higher level of generality, involving concepts applicable to a number of substantive areas. Therefore, Grounded Theory was not an option since staying very close to the process data can make it very difficult to move from a substantive theory generated in the case-specific context of my research phenomenon into more general formal theory. Instead, I will start interviewing using a minimal theoretical framework, building upon results from previous work as suggested by Eisenhardt (1989). Formal theory generated will encompass all but layer 1 of the socio-technical system descriptive framework.

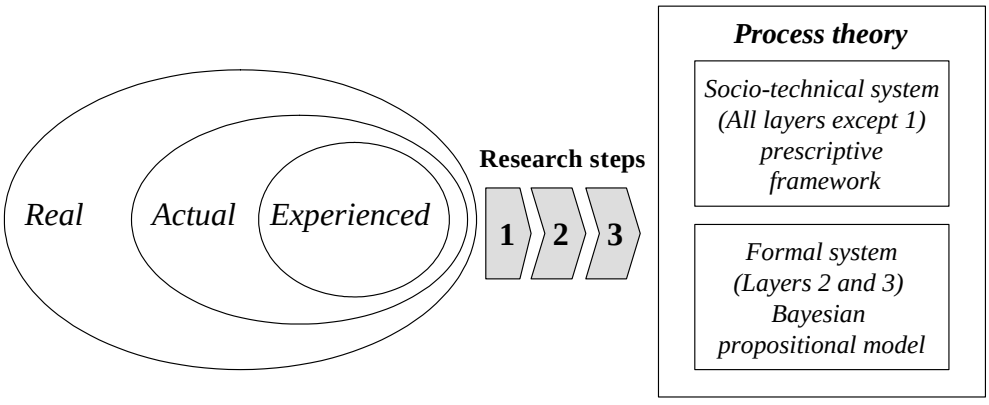


Figure 1-4 A three-step theory-building process resulting from a Critical Realist ontology and epistemological positioning[11]

In a second phase, *computer-based simulation* will be used to generate a propositional model relating innovation -experimentation and decision making- process variables with outcome variables measuring predictive performance. Now, in terms of the socio-technical system descriptive framework the focus is narrowed down to layers 2 and 3. The pharmaceutical Discovery innovation process, carried out by its team members, is now modelled as a reduced form of the socio-technical system; a formal

[11] Bhaskar's ontology representation using Venn diagrams adapted from Partington (Partington, 2000)

system as described above. This represents a Positivist view of reality, now reduced to a fully observable, measurable and objective phenomenon with the innovation team seen as 'a machine moving through the organizational decision process' (Gilbert and Doran, 1994: 37). Clearly, simulation trades accuracy and completeness of reality representation for simplicity and generality (Langley, 1999). However, the strength of a computer-based simulation model emulating a real process is threefold; it allows for risk-free experimentation, it may allow for detection of inconsistencies in existing theoretical frameworks, and above all it allows going beyond explanation into prediction. Knowing that prediction follows from premises, established theory, and available data, in the Popperian sense, making a prediction in these terms I must be prepared to accept that incorrect predictions mean my premises, theory, or data were flawed (Johnson, 2000). However, in Chapter 3 it will be shown further that simulation is the only way of explaining and predicting a proposed quantitative impact of possible pharmaceutical Discovery experimentation strategy options, taking simplifying assumptions as starting premises.

Finally, during *theory development* both perspectives will be integrated into a proposed formal theory, at least partially –for all 'experienced' parts- subject to empirical testing. The developed teleological process theory will show how the case study-based empirical work can be informed and complemented by simulation-based formal reasoning. The process grounding phase, probing into the *meaning* of the innovation process for the team, will lead to a prescriptive framework holistically explaining the *mechanisms* used by the socio-technical system to face the complexity of its mission to solve the innovation problem. The computer-based simulation phase will lead to a propositional model, connected to the prescriptive framework, explaining the impact of front-end experimentation strategy options on the predictive performance of the outcome of the decision making resulting from the experimentation process. Congruent with the requirements of mode 2 research (Gibbons et al. 1994) the constructed theory should stand the test of reality by adequately representing the tendencies recorded in the lay accounts of reality in such a way that social actors should be able to recognize themselves and others in the foreign concepts of the proposed formal theory explaining their experimentation behaviour.

1.3.3 A roadmap for this research

This introductory chapter provided the business rationale and the philosophical and methodological underpinnings of my research. In the next three Chapters I will focus on the results of my empirical and computer-based simulation research, conducted in three consecutive projects. To conclude, in the last Chapter I will integrate the findings of these research projects into a proposed formal theory explaining experimentation behaviour during fuzzy front-end innovation processes. Drawing upon concepts from the body of knowledge of the management of innovation, decision and complexity theory, theory will be built in a number of discrete steps along the roadmap for this research.

Readers who are not interested in the detailed discussion of the methodology and results of the three consecutive research projects can skip the following three Chapters and turn their attention immediately to the final Chapter integrating the findings into a proposed formal theory.

Chapter 2 starts from the radical versus incremental dichotomy proposed in the innovation literature, developing an argument for looking at the research question from the perspective of complexity theory. Then, results of the first exploratory research project will be discussed. To develop a more fine-grained understanding of how and why radical innovation experimentation differs from incremental innovation management practice, six exploratory case studies will be conducted in pharmaceutical development. This will lead to a proposed prescriptive framework linking the dimensions of the complexity experienced by the innovation team to their choice for a 'complexity-handling mode', a specific way of organising the innovation process.

The second research project, described in Chapter 3, will go through a confirmatory case study analysis to answer the research question whether the pharmaceutical discovery process can be used for literal replication of the prescriptive framework that emerged from the exploratory study. Various front-loaded discovery experimentation strategies will be mapped and discussed. Then, a rationale and Bayesian methodology will be proposed to evaluate predictive performance of these front-loaded experimentation strategies using Monte Carlo simulation. Finally, research conjectures will be formulated to guide subsequent process simulations, linking experimentation policies found in pharmaceutical discovery to predictive and business performance.

The third research project, described in Chapter 4, will discuss the results of top-down simulation-based theory development on predictive performance of front-loaded experimentation strategies. Following a review of theoretical models representing the complexity of dynamic experimentation and decision-making processes, it will be argued that an annotated *adaptive system paradigm* is the best choice to emulate fuzzy front-end experimentation behaviour as conducted in pharmaceutical discovery.

Finally, in Chapter 5 formal teleological process theory will be developed, inductively grounded in the research results described above. The formal process theory will integrate the case study-based empirical work with the simulation-based theorizing effort. Process grounding conducted in the first and second research projects, probing into the *meaning* of the innovation process for the team, will lead to a prescriptive framework, holistically explaining the *mechanisms* used by the team as a socio-technical system handling the complexity of its mission to solve the innovation problem. Conversely, computer-based simulation, prepared for in the second project and conducted in the third research project will lead to a propositional and predictive model, explaining the impact of various front-loaded experimentation strategies on predictive and business performance.

2 Managing complexity in radical innovation projects: The need for a paradigm shift?

2.1 INTRODUCTION

The purpose of my research is to build a theory of experimentation strategy required to manage radical innovation projects for performance. As a first step, in this Chapter I explore radical versus incremental innovation project management practice. I define a radical innovation "as a product, process, or service with either unprecedented performance features or familiar features that offer potential for significant improvements in performance or cost, using non-existing or non-proven technologies that did not fully exist at the start of the project *with* high market uncertainty (Leifer et al. 2000; Lynn and Akgun, 2001). It is opposed to an incremental or continuous innovation that refers to adaptation within a particular technological paradigm (McKee , 1992; Baker and Sinkula, 2002) An experimentation strategy is the general approach to planning and conducting experiments (Montgomery, 1997).

Recent previous research acknowledges the dichotomy of incremental versus radical innovations. A radical innovation project is known to require a flexible trial-and-error approach to manage the high levels of uncertainty and ambiguity (Chandy and Tellis, 1998; Veryzer, 1998), as opposed to the stage-gate driven 'compression model' project management style for incremental innovation projects (Wheelwright and Clark, 1992). Finally, all agree that the key difference between radical and incremental projects lies in the project management approach being focused in the former on learning and on planning in the latter (Cheng and Van de Ven, 1996; Leifer et al. 2000; De Meyer et al. 2001).

However, whilst the competitively differentiating role of the radical innovation process is recognised (Iansiti and Clark, 1994; Leifer et al. 2000; Baker and Sinkula, 2002), its actual detailed operation is far less understood (Veryzer, 1998). Therefore, the purpose of this Chapter is to develop a more fine-grained understanding of how radical innovation experimentation management differs from incremental innovation management practice.

This Chapter is organised around two research questions focussing on the strategic perspective of experimentation conducted in radical innovation projects: *'What does the radical innovation process look like?'* and *'How and why is it different from incremental innovation project management practice?'*

Previous innovation research has largely focused on diverse discrete manufacturing based industries as automobiles, computers, mainframes, domestic appliances and telecom. Instead, I will use technology-intensive non-clinical pharmaceutical development as the context to explore the research questions. This science-based development environment, characterised by high technological

uncertainty and complexity, is ideally suited to study radical innovation project teams struggling to reduce the complexity they are facing.

As a result, my evidence suggests that exogenous project characteristics like innovation outcome –'radical versus incremental'- or technological uncertainty –'low-tech' versus 'super high-tech'- are bad predictors for explaining the experimentation approach chosen to manage radical innovation projects, or for explaining the difference between managing incremental innovation projects. Instead, I propose that the complexity experienced by the innovation project team endogenously drives the choice for a particular experimentation approach. Using distinctive experimentation approaches, a mental model of the innovation problem solution depicting variables, their interrelationships, and their value ranges is proposed to emerge and applied during the innovation process. Experienced complexity, characterized by the level of uncertainty and ambiguity facing the innovation team is proposed to drive their choice for specific types of experimentation approaches I will further refer to as complexity-handling modes.

2.2 LITERATURE REVIEW

Recent literature distinguishes between sustaining and disruptive innovations. Sustaining innovations have as a target to improve the most valued attributes of the most demanding customers of the present companies' value network (Christensen and Rosenbloom, 1995; Christensen, 1997). By targeting these most attractive customers in mainstream markets sustaining innovations 'improve or maintain profit margins by exploiting existing processes and cost structures and by making better use of current competitive advantages' (Christensen and Raynor, 2003: 51). Disruptive innovations, in contrast, don't attempt to bring better products to established customers in existing markets. Rather, they disrupt and redefine that trajectory by introducing products and services that are not as good as currently available products.

But disruptive technologies offer other benefits –typically, they are simpler, more convenient, and less expensive products to appeal to new or less-demanding customers'(Christensen and Raynor, 2003: 34). Two strategies for creating new disruptive growth business are distinguished; 'Low-End' disruptions disrupt the prevailing business from the low end using a new financial and/or operational approach. 'New-Market' disruptions target non-consumption by improving performance in new attributes, typically simplicity and convenience (Christensen et al. 2002; Christensen and Raynor, 2003).

While excellent at providing a managerial framework to think about and organising for both types of innovations (Christensen and Overdorf, 2000) Christensen *et al.*'s analysis does add to the confusion in definition of types of innovation when stating that sustained innovations can have an incremental or breakthrough character. Also, the unit of analysis of their research is the innovative end-product and its contribution to competitive advantage and far less the process leading to the radical innovation, which is the subject of my research.

Therefore, since my research perspective looks at innovation as a team-based problem solving process, which builds on previous knowledge and explores a solution space through experimentation and learning, I will further review recent academic literature concerned with disciplined problem solving, coordination and management of interdependence between team members, and adaptive learning.

2.2.1 Disciplined problem solving

Eisenhardt and Tabrizi (1995) contrasted a compression and experiential problem-solving strategy as two theoretical models to enable fast adaptation through product innovation. Both accelerate development provided the first is applied to stable, mature environments and the latter to environments characterized by high market or technical uncertainty. A *compression strategy* assumes the product development process consists of a predictable series of steps that can be accelerated by rationalizing and squeezing them together. An *experiential strategy* emphasizes the use of improvisation and ways to deal with environments characterized by high levels of complexity and chaos. Key is rapid intuition building and creating flexible options to learn as quickly as possible about technical and market uncertainty. In some conceptual papers (Barrett, 1998) jazz players are used as a metaphor for understanding organizational learning and innovation. Seven characteristics that allow jazz bands to improvise coherently are identified and mapped to the management environment. Lynn *et al.* (1996) perform four in-depth historical case studies in different sectors to investigate the phenomenon of reducing market uncertainty during radical innovation. The pattern they observe is the use of a "Probe-and-Learn" process to continuously reduce market and technical uncertainty.

More recent studies acknowledge the inherent difficulty of managing the 'fuzzy front end' and suggest a holistic approach to project management that keeps the balance between creativity and discipline (Khurana and Rosenthal, 1997; Khurana and Rosenthal, 1998). To steer this balance a 'product champion' is needed with the drive to advance the project, and with a vision to frame serendipity (Cox, 1989; Lee and Na, 1994; Veryzer, 1998). In an empirical study of 61 problem-solving attempts in environments involving radical technological change Iansiti (1998) concludes that the more effective organisations applied a 'system focused' approach to integrate novel technological concepts into their designs. From a problem-solving perspective this approach involves emphasizing as early as possible the systemic impacts of novel technological concepts on product functionality and production system performance. Based on a detailed survey of the early development phases of 18 projects Verganti (1999) operationalizes anticipative and reactive capabilities and discusses the 'planned flexibility' concept that implies early identification of critical areas and early planning of reaction measures. The same line of thinking leads to 'enlightened experimentation' (Thomke, 2001) meaning organizing for quick inexpensive or 'massive' (Iansiti, 1998) experimentation allowing to fail early and often to spot and solve problems as upstream as possible.

2.2.2 Coordination and interdependence

Allen *et al*. (1979) already found that the manner in which research and development projects acquire new technologies varies with the nature of the project. Research projects were found to perform best when all project members maintain high levels of communication with the outside, where development projects showed higher performance when one or a few gatekeepers monopolized external communications.

More recently Adler (1995), based on an inductive analysis of 9 printed circuit board development projects in the design/manufacturing interface, concludes that increasing technological novelty requires the use of more interactive coordination mechanisms like cross-functional teams as opposed to standards, plans and mutual adjustments that seem to fit more incremental projects. Grandori (1997) shares this conclusion. In her organizational assessment of inter-firm coordination modes she concludes that in the case of high cognitional complexity, entailing highly innovative joint problem solving, characterized by many serendipities, implying the discovery of cause-effect relations, the main bureaucratic coordination mechanisms are expected to fail. A recent study of project coordination in 37 technology-intensive multinational companies (Gassmann and von Zedtwiz, 2003) proposed that a.o. the type of innovation determines the spatial distribution of project teams and their coordination behaviour. Radical innovation projects are proposed to be preferably run in centralized venture teams since they require a high degree of face-to-face communication and exchange of tacit knowledge. Incremental innovation projects can be run in a decentralized self-coordinated mode of operation since used technologies are known.

Finally, field research conducted by Chesbrough (2003a; 2003b; 2004) suggests business models for managing innovation become more and more open and that innovating companies should distinguish between bringing innovations to current or new markets. Targeting incremental technology to current business is equated to playing chess where one can think several moves ahead. In new markets featuring high technical and market uncertainty, the guiding metaphor to be used is playing poker (Chesbrough, 2004). Where in chess resources are well defined, competitor's resources are well understood, and no new information arrives during the game, in the poker game one must adapt and adjust as new information arrives, and resources and competitors emerge over time. An Open Innovation business model must cope with both situations and rely not only on internal resources but also on more open structures like new ventures to manage new technologies collaboratively with external innovation partners.

2.2.3 Adaptive learning

The learning view on innovation distinguishes between two general situations of learning, involving the accumulation and use of knowledge, skills, and technological capabilities in organisations. '*Exploration* includes activities like search, variation, risk taking, experimentation, play, discovery, innovation. *Exploitation* includes refinement, choice, production, efficiency, selection, implementation, execution'(March, 1991). His findings suggest that maintaining an appropriate balance between exploration and exploitation is a primary factor for system survival. Later, Ekvall acknowledges these

findings and recognizes that the problem of innovation is rooted in the nature of creative processes and creative persons and that radical, revolutionary innovation, as opposed to adaptive and confirmatory innovation is facilitated through different organizational conditions, which creates organizational dilemmas. These dilemmas must be managed since the company must have the capacity of radical innovation to survive in the long run, while at the same time it has to be capable to run effective operations. This requires an act of balance (Ekvall, 1997). In essence, both studies focus on the context of the adaptive learning process. However, neither of them discusses the actual process of learning and how it differs for both situations.

Learning in highly novel and uncertain situations is done through an adaptive trial-and-error process since prior knowledge is not or poorly available to guide the innovation team. Trial-and-error learning involves people taking a course of action, interpreting the outcome response of their action, and is followed by a continuation or adaptation of their course of action based on the results of their interpretation. Hence, it is path-dependent. Van de Ven and Polley (1992) tested this adaptive 'goals-actions-outcomes' model based on a 5-year real-time longitudinal study of the development of a biomedical innovation and observed different patterns of learning in different periods of innovation development. Their findings suggest a faulty learning process of action persistence despite the occurrence of negative outcomes suggesting a random and unpredictable process of learning during the beginning ambiguous period of development. Later, during the still uncertain but less ambiguous concluding period, strong evidence was found for the adaptive learning model cited above. These findings were confirmed in another radical biomedical innovation project where Garud and Van de Ven (1992) speculated that trial-and-error learning guides innovation development under conditions of uncertainty, but action persistence occurs when the developmental process is ambiguous, meaning when it is not clear what specific ends are worth pursuing (Cheng and Van de Ven, 1996).

Chaos theoretical analysis of the innovation process indicates that it is neither an orderly progression of phases, nor a random sequence of blind events. Instead, based on empirical findings from a process research study of the two radical biomedical innovation projects cited above, Cheng and Van de Ven (1996) conclude that learning in chaotic conditions is an expanding and diverging process of discovery, while learning during more stable and periodic conditions is viewed as a converging process of testing. The innovation process is proposed to begin in chaos and to end in order. They argue that the innovation process consists of a non-linear dynamic system, which is neither stable nor stochastic. However, their simulation results have not been successful identifying the parameter values where the innovation process shifts from chaotic into periodic behaviour.

2.2.4 Conclusions

From the above I conclude that the different views in the literature generally acknowledge that radical innovation projects require management practices that are different from incremental innovation project management (Eisenhardt and Tabrizi, 1995; Veryzer, 1998; Iansiti, 1998; Leifer et al. 2000; Lynn and Akgun, 2001). Radical

innovation projects require specific tools that the project manager may find appropriate to the specific situation the project faces (Leifer et al. 2000).

However, what constitutes a radical innovation project? Abernathy and Clark (1985) already agreed upon the binary classification problem and suggested a continuum for technological change defined by 'polar extremes' but failed to indicate what falls in between and where. Abetti (2000) acknowledges that the classification of technological innovations in incremental or radical categories does show different shades of grey since they are not defined according to clear criteria. Therefore, he proposes a five-level scale to classify the radicalness of innovations and cites critical success factors for radical technological innovations. However, he fails to give a clear indication of the need for specific management approaches for the different levels of radicalness identified. Lynn and Akgun (2001) distinguish between radical-, evolutionary-, and incremental innovations and associate success with vision clarity, support, and stability for each of these project types. Based on a cross-industry case study of 16 projects De Meyer *et al.* (2001) propose a framework to adapt project management tools and approaches to the type of uncertainty the project is confronted with.

Summarizing, so far incremental or radical innovations are still intuitive, ill-defined concepts (Durand, 1992). Therefore, to come to a more fine-grained understanding of how to manage radical innovation experimentation for performance I suggest shifting paradigms and focusing on the collaborative problem-solving behaviour of the innovation team. I propose to take an interpretative view, which takes into account the endogenous nature of the technical problem framing and solving process taking place in the innovation project. Problem framing and solving are central methods used by the team members to match the complexity they are facing in innovation projects. I will further refer to this type of complexity as '*experienced' complexity*.

2.3 COMPLEX SYSTEMS THEORY AS A CONCEPTUAL FRAMEWORK

'The concepts related to complex systems may function as unifying cross-disciplinary scientific themes that are essential to understanding emerging interdisciplinary perspectives in the natural and social sciences' (Jacobson, 2001). In the following I distinguish between a logico-scientific and an Interpretative view of complexity. The first view characterizes complex systems by the interactions of their numerous elements and possibly shows emergent behavior not exhibited by individual elements. The latter view takes the perspective of the problem solver thinking about the complexity she/he is facing leading to problem-solving behavior.

2.3.1 Logico-scientific view

The concept of complexity is both related to systems and people (Flood, 1987). Defined as an exogenous characteristic it says for example that complex technologies are systemic, have multiple interactions between their individual components called system architecture (Henderson and Clark, 1990), and are non-decomposable, meaning it cannot be separated into its components without seriously degrading the performance of the whole (Singh, 1997). Within this view, others distinguish between different degrees of technological, logistical, organizational, and environmental complexity and try to show that understanding the degree and type of complexity through a learning-based paradigm can help managers decide how to organize their factories (Khurana, 1999). Relating project characteristics like technology novelty and complexity to project outcome Tatikonda and Rosenthal (2000) find that technology novelty is strongly associated with poor unit-cost and time-to-market results, while project complexity is only associated with poor unit-cost outcomes. Their findings suggest that future research should investigate detailed project task characteristics. Kim and Wilemon (2003) examine several sources of complexity including technological, market, development, marketing, organizational, and intra-organizational complexity and study the impact of these various dimensions of complexity on reasons why NPD projects are late, over budget, or suffer from performance problems. They construct a template project managers can use to evaluate complexity in their development projects. In effect, the logico-scientific way of studying complexity is based on an analysis of the complexity of the task at hand.

The information processing literature provides a common language both for analysing task complexity attributes and for translating the implications of these attributes into person processes (Campbell, 1988). Within this view, Schroder *et al.* (Schroder et al. 1967) identified three primary properties of a complex task; (1) information load or the number of information dimensions, (2) information diversity or the number of alternatives associated with each dimension, and (3) the degree of uncertainty or change involved. In this rationalistic exogenous view complexity can be defined objectively, independently of the person executing the task. Taking this logic further, complexity increases as each of these dimensions increases. Conversely, the three task attributes 'can capture the cognitive demands experienced by a task-doer in completing a task'(Campbell, 1988). Campbell further distinguishes four basic task characteristics that contribute to an increase of complexity: the presence of multiple paths, multiple desired outcomes, conflicting interdependence among paths to outcomes, uncertain links among paths and outcomes. As another example, using the theory of computation Gell-Mann (1995) equates complexity with the size of the shortest programme describing the regularities of a given task or phenomenon. This 'objective' complexity is distinguished from the subjective complexity experienced by the task-doer. In terms of experienced complexity Schroder *et al.* (1967) noted in their experiments a point of performance disintegration varying from individual to individual leading them to the only conclusion that objective and experienced complexity are not identical.

Hence, I conclude that both logico-scientific classifications may lead to a better understanding of the exogenous characteristics of first-order objective complexity and its repercussions on managerial action. However, it does not answer the question how the problem-solver makes sense of the experienced complexity of the task at hand.

2.3.2 Interpretative view

The interpretative view of complexity takes the perspective of the 'thinker thinking about complexity' (Hatch and Tsoukas, 1997), of the –radical or incremental-innovation project team as an interpretative adaptive system making sense of complexity (Weick, 1995). 'This shift from focusing on the system itself (first-order complexity) to focusing on those who describe the system as complex (second-order complexity) exposes the interpretive dimensions of complexity' (Hatch & Tsoukas, 1997: 12). In responding to the complexity they are facing they have the capacity to adapt or enact the experienced complexity, thus adapting to it or modifying it, rather than merely responding to any objectively given complexity in the environment (Boisot and Child, 1999).

The joint effects of interpretation and enactment have in the past served to distinguish social from natural or biological systems. The recent surge of interest to study social systems through the lens of complex systems theory is motivated by the argument that we can enhance our understanding of social systems by modeling them using the analogon of natural or biological systems (Holland, 1992; Anderson et al. 1999; Stacey et al. 2000). This view is supported by at least five properties held in common by natural, biological, and social systems (Hatch & Tsoukas, 1997: 7) called complex systems; (1) *non-linearity* meaning there is no proportionality between causes and effects, (2) *scale-dependency*, there is no single measurement which will give a true answer, (3) *recursiveness*, tending to repeat a basic structure at several levels, (4) *sensitivity to initial conditions*, small perturbations can lead to chaotic system behavior, (5) *emergence*, the tendency to shift to a new mode of behavior, the description of which is not reducible to the previous description of the system's behavior.

Interpretative research into radical innovation echoes these findings of the complex systems school of thought concluding for example that 'discontinuous [or radical] innovation seems to be an inherently messy process [where] coincidence and fortuitousness play an important role' or 'the lack of formal structure seems to reflect an appropriate and necessary "looseness" rather than poor implementation' (Veryzer, 1998: 318). In a recent grounded theory study on a radical organizational innovation Carrero *et al.* (2000) find the innovative actions indeterminate, non-linear and non-repetitive, showing chaotic patterns. The radical innovation is conducted through a self-managed process. Also, chaos-theoretical analysis of the innovation process tells us that it 'consists of a non-linear dynamical system, which is neither orderly and predictable nor stochastic and random' (Cheng and Van de Ven, 1996), or that organizational decision process level data in the development of a new biomedical device reveal a simple underlying chaotic order (Koput, 1992). And finally, Kiel (1991) suggests that while the tools of the non-linear paradigm may not afford the complete description of social

reality they do appear to represent an incremental step toward greater understanding of the behaviour of complex social systems

Hence, I conclude that studying the innovation process through the lens of complex systems theory can lead to fresh insights and more fine-grained interpretative understanding of the project management approaches required to deal with the serendipity, ambiguity, uncertainty, and chaos which are typical for the radical innovation situation.

2.3.3 Complexity-handling

How then does the innovation project team, acting as a socio-technical system, handle the complexity it is experiencing? In a recent article, Allen (2001a; 2001b; 2001c) describes complexity handling essentially as a trade of 'complexity' of the real world or the problem to be solved for the 'simplicity' of some reduced representation and cites four assumptions through which the complexity reduction is made[12]. This reduced representation abstracts from phenomena those regularities that underpin the form they adopt and is created through a modelling process of *codification* –involving the assignment of data to categories- and *abstraction* –involving a reduction in the number of categories to which data needs to be assigned for a phenomenon to be apprehended. In modelling complexity two complementary types can be identified; *cognitive* and *relational* complexity. The first focuses on the content flow among agents, the second on the structure of the interactions that such flows allow among agents (Boisot and Child, 1999). In this paper I focus on modelling cognitive complexity. The innovation project team acts as a complex adaptive system modelling the cognitive complexity it faces by separating regularities from randomness.

2.3.4 Uncertainty and ambiguity characterizing experienced complexity

To understand how the complex adaptive system models this cognitive complexity I distinguish between the interpretative choice over uncertainty and ambiguity the complex system has in the sense-making situation it faces.

Within the interpretative research tradition *uncertainty* and *ambiguity* are dissimilar concepts (March, 1978; Schrader et al. 1992; Weick, 1995) that should be studied in the context of sense making in organizations (Weick, 1995). A problem is 'an

[12] The assumptions through which the reduction from complexity to simplicity occurs include: (1) the relevant system boundary, excluding the less relevant, (2) the reduction of full heterogeneity to a typology of elements, (3) individuals of average type, (4) processes that run at their average rate. If all four assumptions can be made then we can model the complex situation at hand using a set of deterministic differential equations as used in system dynamics. If only the first three assumptions can be made we have stochastic differential equations that can self-organize as the system may jump between different basins of attraction. With only the first two assumptions the complex system behaviour can be modelled as adaptive evolutionary change in which the system can spontaneously evolve new types of agents, behaviours or new problems and in which case it is impossible to predict the creative response of the system to any particular action taken (Allen, 2001c).

undesirable situation –a gap between the way things are and the way one wants them to be- that is significant to and may be solvable by some agent, although probably with some difficulty' (Smith, 1988). The relevance of this definition for my work is in the word 'undesirable', which implies that the problem is not given by the environment. Instead, '[it] is a relationship of disharmony between reality and one's preferences, and being a relationship, it has no physical existence. Rather, problems are conceptual entities or constructs' (Smith, 1988: 1491) that are designed or discovered (Weick, 1995: 89) in a problem-solving process, which elucidates the inseparability of thought and action.

Within this process the problem solver can dynamically choose for levels of ambiguity and uncertainty to frame the problem acknowledging that both the problem and the set of possible solutions are in a constant flux, taking it for granted that assumptions and solutions are never final but that always improvements can be made, thereby voluntarily reintroducing ambiguity or uncertainty into the innovation process (Schrader et al. 1992).

Uncertainty refers to ignorance (Weick, 1995: 95) or lack of information about variables known to the problem-solver. In the case of uncertainty, the problem-solving process chosen consists of specifying the precise values of the identified problem variables, thus reducing experienced complexity. In contrast, ambiguity refers to confusion (Weick, 1995: 91) by too many interpretations, to lack of clarity about the various variables and their relationships relevant to the problem. 'The problem in ambiguity is not that the real world is imperfectly understood and that more information will remedy that. The problem is that information may not resolve misunderstandings' (Weick, 1995: 92). Here, the definition and nature of the problem or problem structure is in doubt, the variables and their relationships are not known, goals are vague, multiple and often conflicting. In the case of ambiguity, problem solving becomes the creative process of specifying the variables and their relationships relevant to solve the problem through a process of 'model building, negotiation, problem framing, evaluating and reframing, and model testing' (Schrader et al., 1992). 'Ambiguity understood as confusion created by multiple meanings calls for social construction and invention' (Weick, 1995: 95).

Mapping this to the work of complexity theorists the argument is developed that interpretative systems have two distinct ways of handling complexity (Boisot and Child, 1999). They can either reduce it or absorb it. *Complexity reduction* involves getting to understand the problem and acting upon it directly by reducing the level of uncertainty. Scanning for and filling in the values of the known problem variables reduce uncertainty. In contrast, *complexity absorption* involves the creation of options and risk-hedging strategies to explore environmental variety reducing the level of ambiguity.

2.3.5 Experienced complexity as the lens of complex systems theory

I argued above that to study experimentation approaches used in innovation projects one needs to focus on the interpretive dimensions of the mental modelling process. I will use the lens of complex systems theory focusing on second-order or experienced complexity. Experienced complexity is a property of the interaction between problem solver and problem. The levels of uncertainty and ambiguity facing the innovation team are modelled to characterize experienced complexity. These levels are not exogenously given problem variables but two dissimilar components of problem framing whose level the innovation project team dynamically chooses.

2.4 RESEARCH METHODS

Exploring experienced complexity requires a research method capable of capturing the complexities of the sense-making process within the innovation team. Therefore, I used an inductive multiple case comparison methodology to study the phenomenon of joint problem solving through experimentation in its natural setting. I chose this method since it is particularly suited to explore 'How' and 'Why' questions concerning a contemporary phenomenon in a new topic area (Eisenhardt, 1989). I use the lens of complex systems theory as described above as the minimal theoretical framework to start conducting the interpretative exploratory analysis. It allows me to build upon the existing literature base therefore ruling out the risk of reinventing known relationships, while still having a high likelihood to discover novel insights. Multiple case study design is generally regarded as more robust as single case study, since the former provides for the observation of a phenomenon in different settings (Yin, 1989).

2.4.1 Case sample characterisation

The project case sample was drawn from the Chemical-Pharmaceutical Development division (approx. 500 scientists) of a global top-10 pharmaceutical company I will further refer to as 'PharmaCo'. Six cases were selected to represent a continuum from incremental to radical innovation projects within their division. For reasons of confidentiality Greek alphabet denominators were used for the drug product innovation projects involved. To assess projects' suitability for this research project characteristics *technological uncertainty* representing the departure from the present technology base[13] and *technological novelty* -the latter with values 'existing', 'new-to-world', 'new-to-company', 'new-to-(pharmaceutical) industry' and 'new-to-company'-, were taken as proxy measures to facilitate project positioning on the 'incremental-versus-radical' spectrum. Outcomes of all five projects rated 'new-to-company', 'new-to-industry' or

[13] Following Burgelman and Maidique (Burgelman and Maidique, 1988) I define a technology base as 'the set of technologies, which are embodied in a firm's products/services and production/delivery system'.

'new-to-world' fit the definition of a radical innovation[14] used in my research and were rated under *innovation outcome* in case summary Table 1.

Unit of analysis of my research is the problem-solving process within the project team in the context of the drug delivery system innovation projects described below conducted in a chemical-pharmaceutical development setting.

Case Studies[15]	Technological Uncertainty[16]	Technological novelty	Innovation Outcome
Alfa IR Tablet	Low-Tech	Existing	Incremental
Gamma Controlled Release Tablet	Medium-Tech	New to company	Radical
Supercritical Fluids	High-Tech	New to industry	Radical
Nanosuspension	High-Tech	New to world	Radical
Microemulsions	Super High-Tech	New to world	Radical
DNAZyme	Super High-Tech	New to world	Radical

Table 2-1: Exploratory case study projects key exogenous characteristics

Alfa Immediate Release (IR) is a classical tablet formulation project that needed to be executed under time pressure. It drew upon the existing 'Oral Solids Delivery System' technology base for execution. The project consisted of designing a formulation for a broad dose range and pharmaceutical production process for a commercial product with an active ingredient to be released immediately into the body with a daily intake of 2 to 3 tablets. This incremental innovation is based on well-known existing technology and its related development procedures, and is well documented in best practices. Key to project success is fast and efficient execution.

The *Gamma Controlled Release (CR) tablet* project required the design of a tablet formulation to be taken once a day releasing its dose slowly, following a controlled release curve during the whole day without dose dumping, and this for a wide dose range. To get the same effect using IR technology two tablets need to be taken a day. The Gamma project being in the Alzheimer franchise the differentiation potential was based on patient compliance. Also, physicians appreciated the absence of side effects that go with the typical IR dose overshoot peak at intake. Therefore this was classified as a radical innovation at project initiation. At its start the project technology used could draw on basic IR knowledge and skills but it did require a significant departure from this IR technology base. Other companies had been experimenting with

[14] As cited in the introduction to this paper I use the Leifer *et al.* (2000) definition of a radical innovation outcome 'as a product, process, or service with either unprecedented performance features or familiar features that offer potential for significant improvements in performance or cost'.

[15] For reasons of confidentiality Greek alphabet denominators were used for the drug product innovation projects involved.

[16] Following Shenhar and Dvir (1996) proposed project typology *Low-Tech* means implementing familiar technologies relying on the existing and well established technology base, *Medium-Tech* projects involve adaptation of familiar technologies and rest mainly on existing technology base, *High-Tech* implies first use of the new but existing –having been developed prior to the project's initiation- technology, *Super High-Tech* projects are based on new and non-existent technologies at the moment of project initiation.

the technology and basic techniques used were known to work at small lab scale level so I classified it as ‘new-to-company’ and ‘medium-tech’. The chosen design consisted of a multi-particulate system of multiple beads all carrying active substance covered by a controlled release membrane and an immediate release coating filling a capsule. The cost and manufacturing efficiency advantage of this was that to increase the dose range only more of the same complex beads were needed as opposed to a monolithic one-bead system complemented with CR beads where for every dose another bead needed to be manufactured. Also, this latter case complicated manufacturing logistics.

Supercritical Fluids. Melt extrusion is a well known single continuous production process used in pharmaceutical manufacturing to prepare new solid drug dosage forms like tablets, granules or pellets. An optimised extrusion process is known to be simple, cheap, and solvent free. Also, extruded formulations are considered especially useful in enhancing the bioavailability of poorly water-soluble drugs. An important drawback of the technique is related to the high temperatures needed to achieve the molten state of the polymers used in the extruded blend. As such, this formulation approach is limited to thermostable formulations including not only active compounds, but also the excipients, including the polymers themselves. The plasticizing action of a supercritical fluid (SCF) could reduce the temperature of melt extrusion thereby extending its use for thermosensitive polymers and compounds like peptides or proteins. Due to their unique solvent properties, SCF applications were first directed in the fifties in the food industry towards extraction, such as decaffeination of coffee, spice extraction and lipid purification. More recently, SCF have also been used as a medium for polymerisation in the pharmaceutical industry, especially because of the regulatory limits on solvents. It has possibly wide applicability as a platform technology used for micronization, to replace human tissue, as an environment friendly manufacturing technique, or even to accelerate the joining of human bones. Therefore, I classified it as a radical innovation that is ‘new-to-industry’. The case project is the first to explore the application of this technology platform in the pharmaceutical industry with the objective to develop the use of CO_2 SCF to broaden the applicability of melt extrusion, therefore it is classified as ‘high-tech’.

Nanosuspension was a ‘new-to-world’ technology at project initiation in the mid-nineties. Nanoparticle-based delivery technology allows decreasing particle sizes to the 100 to 200-nanometer region. It consists of a sterile milling process facilitated by specially developed beads and additives or excipients. It was a radical innovation at project initiation while nanometer range fineness increases drug formulation surface area available, positively influencing solubility and dissolution rates of active substance. The extremely small particle size prevents blocking of the veins, which is particularly interesting for compounds having their action at the level of the brain or the central nervous system. Also, it can be used in parenteral applications, like for intra-muscular usage, eardrops or nasal administration into the central nervous system. The difficulty at this size range is that hydrophobic particles have the tendency to cluster. This has to be prevented by coating the particles. Unfortunately, the range of excipients usable for coating is rather limited. At the start of the project, nanosuspension was available with an outside technology provider who had shown the first proof of concept on lab-scale. However, the milling process had to be optimised and made sterile throughout. Also, the formulation had to be designed for PharmaCo specific compounds. Heat-resistant

excipients had to be found to make the formulation resistant to the heat required during the manufacturing process to make it sterile. Since the technology already existed in embryonic form at the start of the project I classified it as 'high-tech'.

Micro-emulsions are distinguished from emulsions by their transparency. More fundamentally, it is a technology known to the pharmaceutical industry to be used as formulation adjunct for compounds whose bioavailability is solubility or dissolution time limited. An example situation is found in the central nervous system where this technology could optimise drug uptake or it can be used in general to improve the uptake of poorly water-soluble drugs. Since most micro-emulsions contain a toxic component and cannot be injected, mainly oral micro-emulsions based products are on the market. Only one combination can be used for intravenous use. Unfortunately, it features low solubility of the active substance. This project investigates a radical innovation, which is the 'new-to-world' application of this technology as a platform to cover oral, transdermal, and parenteral applications using biodegradable liquid polymers systems. Key inputs for the project are synthetic absorbable polymers used in the medical device industry -in this case the Johnson & Johnson Corporate Biomaterials Centre (J&J CBC)- to construct for example artificial veins. Through macromolecular engineering these polymers can be modified from liquids, to pastes, waxes, rubbers or solids. Not only these mechanical properties, also the spectrum of absorbability can be regulated to a large extent to cover the whole application spectrum. Since this project focuses on non-existent technologies at project initiation and even at this very moment, I classified it as 'super high-tech'.

DNAZyme. The advent of high-throughput screening has caused a general trend in the pharmaceutical industry for drug candidates to be of higher molecular weight, more lipophilic, of poorer water solubility and generally lacking in oral bioavailability. For more than twenty years scientists have tried, unsuccessfully, to develop safe and efficient delivery systems for these macromolecular therapeutics. Applications include gene therapy and intracellular delivery of DNA fragments to inhibit protein-production that is highly thought about for cancer. In this project one tries several 'new-to-world' approaches to get into the cell and show functional activity, which is non-toxic. It involves the design of synthetic molecules, called RNA-cleaving DNA enzymes or DNAzymes that can select, bind and cleave a therapeutic target thus prohibiting protein production. Physical chemistry and metabolism are the major issues to be handled. More specifically, DNA fragments are made up of nucleotides that don't go through cellular membranes well and human enzymes very rapidly break down DNA. Hence, this project features extreme technological complexity trying out delivery technologies that have never been used before. However, if a solution can be found it will open at least the very large cancer market. Therefore, I classified it as a 'super high-tech' project with a radical innovation outcome.

2.4.2 Forming an interpretative community

To probe into the 'How' and Why' of the experimentation approaches used, I formed an interpretative community consisting of the six project managers and one team member. Interviews lasted between one and two hours, were tape-recorded and transcribed.

Interviews were complemented with reviews of innovation project specific archives and written departmental procedures. A total of 14 interviews were conducted. For each interviewee the first interview session was structured around the process they had been using in their cited projects to solve their innovation problem. Visual mapping and temporal bracketing (Langley, 1999) was used to make sense of the data. In the following sessions interviewees were confronted with their process maps emerging from the first interview round. Maps were checked for accuracy and refined. Also, these sessions focused on the 'Why' of the different phases in their innovation projects and the reasons for transitions between phases.

Data coding was organized by arraying empirical indicators –activities and choices made by the innovation team- on multiple tracks along the problem-solving process and its constituent phases. A retroductive approach (Poole et al. 2000) was used by initially taking the theoretical background discussed above to generate categories to include (1) experimentation approach, (2) learning approach, and (3) coordination approach, followed by a new category I called (4) target setting approach that emerged during interviews. Inconsistencies were verified by using additional sources of data or through verification by the original informants. To ensure reliability a case study database was developed to formally assemble qualitative and quantitative evidence material. A detailed summary of this case study database including the number of interviews conducted, the respondents, empirical indicators, relevant PharmaCo internal document references, and visual maps of all studied cases resulting from the interview sessions, can be found in Appendix A.

2.4.3 Shaping a conceptual framework

To increase internal validity first within-case analysis, then cross-case visual map pattern matching was executed. I looked for within-group similarities coupled with inter-group differences. Using temporal bracketing, groups were chosen along the identified project phases cited above, becoming units of analysis for replicating the emerging conceptual framework, which was shaped in two steps.

The first step focused on the *'how'* of the innovation process. Within-case visual mapping analysis led to the identification of distinct 'periods' or temporal brackets. A period typically is terminated by a team decision. Each bracket was given an identifier A, B, or C. Then, cross-case bracket pattern matching led to an emerging frame of across-project bracket types describing what was done by the various teams to solve the innovation problem. Finally, each period consisting of similar bracket types was given a name describing how complexity was handled by the innovation team in this specific period.

A second step searched evidence for the *'why'* of the emerging project phases. This was done by recording empirical indicators for 'uncertainty', 'ambiguity' and 'phase transitions', and mapping them to the emerged bracket types now called complexity-handling modes cited above.

2.5 CASE STUDY RESULTS

Complexity experienced by the project team is driven by its sense making process that consists of finding a solution to the innovation problem at hand. In all cases an innovative drug delivery system[17] had to be designed. Temporal bracketing results indicated three distinct 'periods' emerging from visual mapping and relevant empirical indicators (see Figure 2-1). The three bracket types emerging from the cases focussed on (1) generating and selecting solution concepts, (2) fully characterising the chosen solution concept, and (3) applying the characterised concept to a specific product to bring it to market.

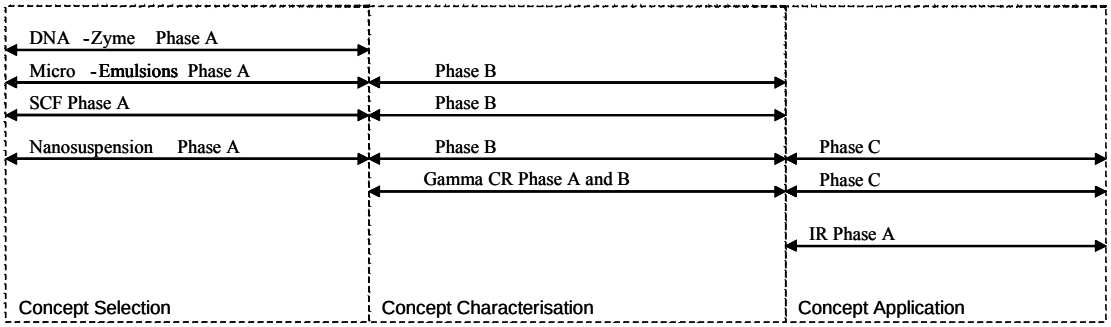


Figure 2-1 Exploratory case study temporal bracketing results

Phase A identified in the DNA-Zyme case (Figure 2-2), the Micro-emulsions case (Figure 2-3), the Supercritical fluids (SCF) case (Figure 2-4), and Nanosuspension (Figure 2-5) constitute a complexity-handling mode I call *'Concept Selection'*.

Phase B identified in the Micro-emulsions case (Figure 2-3) and the SCF case (Figure 2-4), together with Nanosuspension case Phase B (Figure 2-5) and Gamma Controlled Release (CR) Phases A and B (Figure 2-6) constitute a complexity-handling mode I call *'Concept Characterisation'*.

Phase A identified in the Alfa IR case (Figure 2-7) and Nanosuspension (Figure 2-5) and Gamma Controlled Release Phase C (Figure 2-6) constitute a complexity-handling mode I call *'Concept Application'*.

2.5.1 Concept Selection

'Concept Selection' leads to a mental model characterising the innovative drug delivery system's core, depicting the critical variables and their relationships affecting solution proof of concept level performance. Working towards meeting solution critical requirements it gradually resolves ambiguity and brings focus by ruling out as soon as possible in the process candidate solutions that don't work, and by organising work

[17] As defined here a drug delivery system can be a technology (DNA-Zyme, Microemulsions, SCF) or a final product incorporating a technology (Nanosuspension, Controlled Release tablet, IR tablet)

primarily around the key questions to be solved, operationalised in the solution critical requirements. The latter are defined at proof of concept level.

Before starting the experimentation process, *solution critical requirements* are specified as a target against which alternative solutions are tested. As testified in the DNA-Zyme, SCF and Microemulsions cases;

> "What Michel [project manager] tried to do is to set up, even before we started looking at these people, what we like to see for the very first criteria; that was functional DNA-Zyme activity in the cells".

> "…I think the idea was really; let's identify these criteria [solution critical requirements]. We have the opportunity to look at our range of technologies, but having said that, the very first step is to get rid of as many as possible. What would happen, if all of these passed the first test, I don't know. But we knew that the bar was so high, that it wasn't likely to happen."

> "For the meltextrusion application we stated as an objective that the temperature with SCF had to be lower as the temperature we would have had without SCF. And this for a pressure that would not be so high it would endanger the functioning of the meltextruder".

> "…The objective of these first two phases [refers to Microemulsions project process chart] is really to check for entire families of polymers whether it is possible to form micro-emulsions and to encapsulate an active ingredient, or not".

In the micro-emulsions project (Figure 2-3) the proof of concept consists of showing for the biodegradable liquid polymers that they can be made water-soluble, that the active ingredient can be solubilized using the polymers, that the solution is non-toxic and permeable into human tissue. In the Supercritical fluids project (Figure 2-4) both polymers and active substance have to show solubility at certain critical pressure and temperature values. Finally, proof of concept needs to show that it is possible to bring the supercritical fluid-based system in the extrusion process at a temperature that is lower as the one used in classical melt-extrusion.

The experimentation process consists of running several experiments in parallel exploring how critical variables affect response of the innovative delivery system to be designed. Informed by the academic literature on the subject, in the DNA-Zyme project (Figure 2-2) several competing completely different solutions are tried in parallel. Ruling out from the beginning the known non-workable delivery systems documented in the literature, a number of collaboration agreements were set up with outside technology companies and universities testing various embryonic delivery technologies[18] as an overall solution for delivery of the therapeutic agent into the cell. Considering the high costs involved in running intensive reciprocal research collaborations in parallel, the chosen experimentation approach is marked by its step-by-step character. Each time a set of experiments is co-designed to show response that fits one of the system critical requirements. Ones confirmed a next step could be funded to show positive response on the next requirement and so on. This gradual approach, ruling out non-working concepts is advocated by all project managers;

[18] For confidentiality reasons the specific names of these embryonic gene therapy related delivery systems couldn't be mentioned here.

"…Everybody got the same message what was going to be the first milestone, what we wanted to see at that point. For those companies, that were either at the first milestone or that we had a good feeling about we broadened the topic to say, the next area will be this, and than after that it will be…always raising the hurdle…" (DNAZyme)

"..I don't know if everybody does, but we took a stepped approach with each of the candidate solutions…first they had to show you can get it into the cell, then you have to have some in-vivo activity, then…" (DNAZyme).

"…the process is really to check for each family of polymers if micro-emulsions can be formed, then if active substance can be encapsulated. If it doesn't work, go back, if it does, take another family and start all over again…. After this initial phase we will be able to fully characterise the formulation for parenteral and oral applications" [refers to Microemulsions project process chart]

"…In [Phase A] … we tested both concepts on one polymer. Both single and twin screw worked but the latter gave the best blend properties so we chose the last one to show proof of concept … "In [Phase A] you check; is it feasible or not?" (SCF)

Although the micro-emulsions project is carried out in-house it does rely on candidate biodegradable polymers coming from an outside technology provider and also uses this gradual experimentation approach, resolving ambiguity and delivering proof of concept in a step-by-step mode. Each polymer first has to pass the water solubility test, then it is checked whether it could form micro-emulsions, and then it is checked whether active substance could be solubilized in the micro-emulsion. Running test loops this way, of the initial five proposed classes of polymers only two proved to show self-emulsifying properties and were further characterised on their physico-chemical properties using computer modelling.

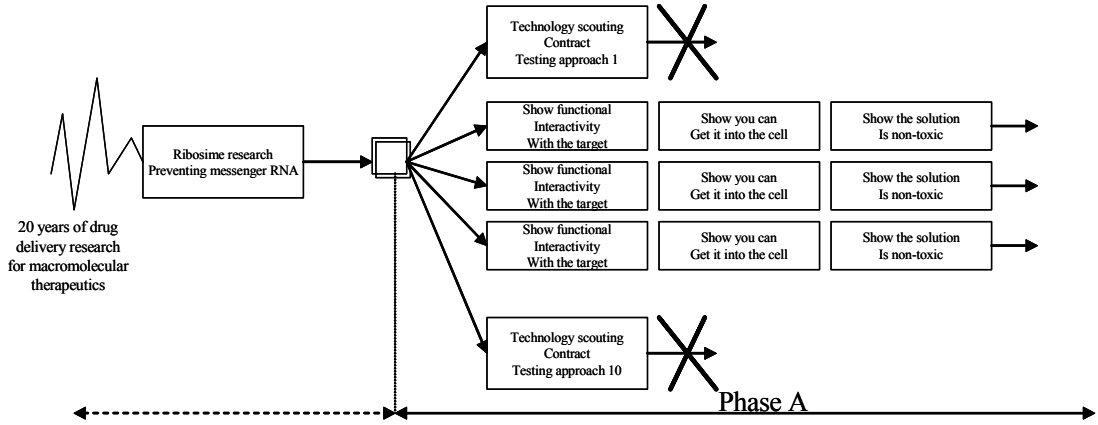


Figure 2-2: DNA-Zyme case process flow

All projects cited above started with explicit external learning by consulting various, not necessarily domain-specific literature. In the SCF project e.g. also food and plastics processing literature was consulted. However, predominantly an external tacit learning approach was applied in all projects. It consists of working with outside

technology providers that can provide part of or the totality of the system solution. Due to the highly uncertain nature of the experimentation process each relationship and its results are closely monitored and adjusted or finished if no further solution potential can be shown to fulfil the critical system requirements. Relationships are reciprocal since technology providers also gain from the learning done in the experiments to fine tune their novel technologies. Microemulsions and DNA-Zyme project managers acknowledge this reciprocal approach;

> "So, there are two big pharma-ideas about this; one is that you sit back, and hope for the best to wait for somebody to give you something off the shelf or you can try to have some impact on their development, by going in early and trying to steer, manipulate the direction which is going to be most interesting and beneficial for you" (DNAZyme).

> "…This is definitely a reciprocal collaboration contract we have with UCL [university] (Microemulsions) "

In the DNA-Zyme case eventually three technology companies and research institutions were chosen to carry out the research. In the micro-emulsions project two classes of polymers were selected for further research and development. Knowledge transfer is an on-going process during the full period of joint research where both parties learn from each other and try collaboratively to solve the disruptive innovation problem.

However, if a sufficient knowledge base exists in the company, relationships with outside technology providers can be structured on a problem-solving basis without any further joint commitments;

> "…Knowledge transfer happened mainly at the beginning… later [Phase B] we used the relationship only for problem solving, and it was not a joint development effort…" (Nanosuspension)

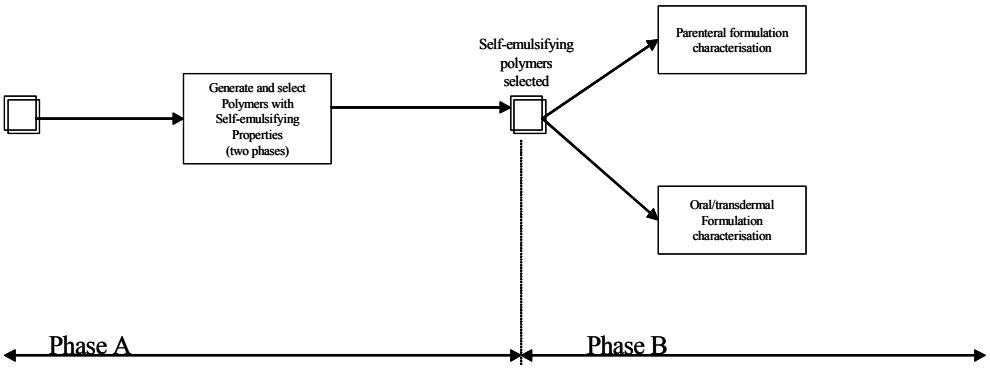


Figure 2-3: Microemulsion case process flow

So, the more ambiguous the situation, the more coordination of the different sets of experiments in each project is done using an intensive reciprocal process where work is distributed amongst specialist organisations and organised to meet a milestone at

which the fulfilment of an up-front specified set of system critical requirements must be shown. If it cannot be shown, the candidate solution is rejected.

Ideally, this work is planned in advance like in the micro-emulsion project where small lab-scale experiments could be conducted in advance to get some insight in time lines required showing candidate polymer performance and plan milestones.

> "…Based upon some first experiments we estimated the time needed to do synthesis and characterisation of the compounds, then, we estimated it would take so many months to synthesize so many polymers… a professor from UCL has also made an estimation, then we compared and saw we came to about the same results (Microemulsions)

In the DNA-Zyme and supercritical fluids projects this could not be done since no solution path was known to exist or, in the latter case no equipment was available to conduct these preliminary tests. Then the project team relied on specifying target milestones delivering system critical requirements and close monitoring of progress towards these milestones, gradually eliminating non-workable candidate solutions, resolving ambiguity and bringing focus into the solution.

2.5.2 Concept Characterisation

'Concept Characterisation' leads to a mental model characterising the full innovative solution's application domain, depicting all relevant variables and their relationships affecting application system performance. Starting from proof of concept the full application domain gets characterised through the definition of uncertainty areas in which a more structured adaptive learning process gradually resolves ambiguity beyond proof of concept level down to all variables and their functional relationships affecting solution performance. Also, critical value ranges get defined for all variables to indicate the boundaries of the application domain, as compared to other drug delivery technologies.

Solution requirements become a moving target starting off from a broad 'minimal' proof of concept level, zooming in to a 'feasible' application domain for which the technology can be used in further application development. The exact target is not known upfront as in the previous complexity-handling mode with the solution critical requirements, but emerges during application domain characterisation through an adaptive learning process.

> "…[refers to project process chart] then we would like to characterise parenteral and oral applications [in parallel] exploring how much active substance can be solved, whether a capsule can be filled, does the capsule break when it gets into contact with the acids in the gut? How about the toxicity and permeability for each of these applications? " (Micro-emulsions)

> "In the first phase the question to be answered is; is it stable or not? In this phase, the experimental plan will be more directed towards exploration of ranges of parameters –is this range a good range to function at production scale? (Microemulsions)

"As soon as we find out if and how we can get the active substance into the SCF we will have to check how good the active can be solubilized, or which volume you can usefully bring into the extruder, what is the stability, bio-availability, of the mixture?"(SCF)

"The target in Phase A and B is evolving in the sense that you're discovering domain limits. While in Phase C you know the limits but now your efforts are directed to developing a product" (CR)

The DNA-Zyme and micro-emulsions projects did not get into this mode yet although for the latter project two application domains have already been identified with a time indication when the respective application domains will be characterised and ready to be used for application development.

Since the SCF project team has passed the proof of concept hurdle mid 2000 it now concentrated on further characterising the meltextrusion application domain. Using active compounds and polymers that were soluble and showed positive interaction with supercritical CO_2 ($SCCO_2$) during Concept Selection, the meltextrusion process and selected polymer interactions were further characterised. Limits of the applicability range were determined through experimentation, thus narrowing the targeted application domain.

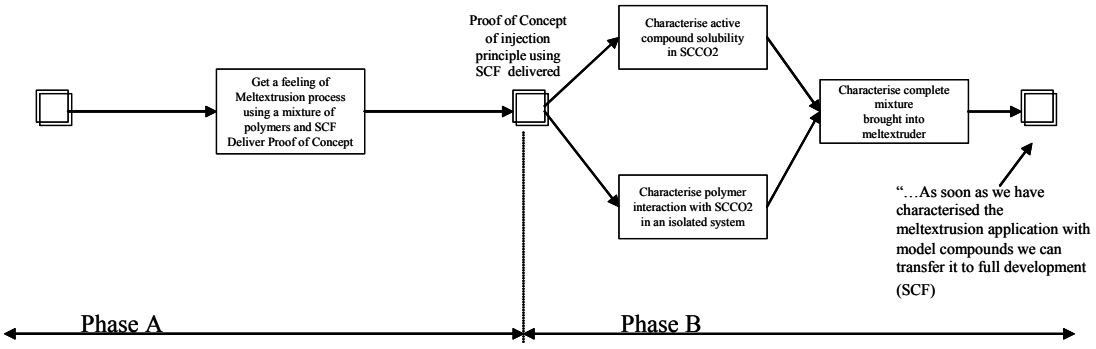


Figure 2-4: Supercritical Fluids case process flow

In the Nanosuspension project an outside technology provider had delivered proof of concept but not on sterile technology. So the team had to concentrate on investigating how to make the process sterile and finding heat resistant formulations required for the sterile production process.

"…finally we had a working formula…this was the end of Concept Selection" (Nanosuspension)

As soon as proof of concept was delivered, the targeted application range was discovered through experimentation. Hence, during Concept Characterisation (Phase B in Figure 2-5) two domains were further characterized; the product formulation using nano-technology, and the sterile manufacturing process. Also in the Controlled Release project the working principle was known at the start but one didn't know how to obtain the exact kinetic release profile of the CR tablet. Application domain requirements emerged gradually by conducting two series of 'test–decide–redirect' cycles leading to

the choice for a multi-particulate controlled release system using solution-based – instead of suspension-based- drug coated sugar spheres with an organic rate –instead of aqueous rate- controlling membrane.

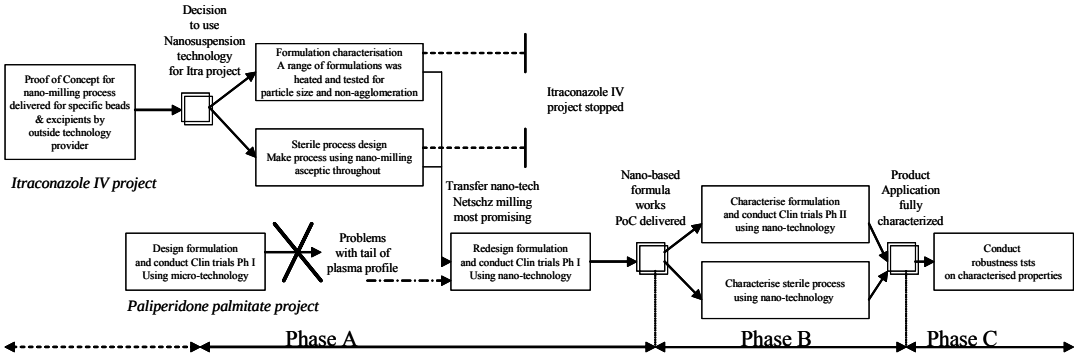


Figure 2-5: Nanosuspension case process flow

Experimentation in Nanosuspension, SCF, and CR projects is conducted within a predefined number of *uncertainty areas* to be explored. For each uncertainty area critical assumptions are defined, and experiments are conducted to test assumptions on limits to applicability ranges.

> "…We split up the work in two coherent design packages [Formulation design and Sterile process design] each exploring and characterizing part of the solution…" (Nanosuspension)

> "Then we started characterising the product…every parameter; particle size as a function of time, sedimentation as a function of time,… we checked for robustness" (Nanosuspension)

> "You split up the concept proven in Phase A into subparts that you will characterise" (SCF)

> "…You start with determining under which conditions you can add SCCO2 to the polymer, then you look for the lowest possible temperature at which this can happen…you characterise parameters like speed of addition, pressure, single versus twin-screw geometry of the extrusion screw…" (SCF)

> "We selected a multiparticulate concept with an IR and CR piece [Uncertainty area 1]. Then, we said we will be working using an aqueous solution and fine-tuned it to get a workable profile. That was Phase B or concept characterisation [Uncertainty area 2]" (CR)

More systematic Design of Experiments (DoE) -based experimental guidelines are used to plan and carry out the experiments.

> "Design of Experiments will also be easier to be used in this phase than in the previous because you now have a good idea which parameters are going to influence more or less strongly the process" (Microemulsions)

> "In this mode the question is much more; is it stable or not? Or, [for this parameter] is this a good working range to have the application work at full production scale? DoE will be easier in this phase [Phase B] while you have an idea which parameters will influence the process more or less. So you know the factors to do a proper DoE". (SCF)

Experimentation results from different uncertainty areas are integrated into a characterised application domain for a limited application range. Unforeseen results may lead to the definition of new uncertainty areas to be characterised or even explored at proof of concept level.

> "…[After parallel characterisation of both areas] then we could bring the characterised supercritical fluid with active substance into the characterised extrusion process, bringing the two together…if it works we can start modelling the nozzle, if it doesn't work we'll have to check whether we could use carbon solvents" (SCF)

For the Nanosuspension project the basic technology available from an outside technology provider, consisting of a nano-milling process using specified excipients and beads, was assumed ready to be made reproducible for clinical studies with company specific compounds. In contrast to the Microemulsions project where model compounds were used to carry out concept characterization, here application domain characterization was immediately carried out on the compound to be taken to market.

In the SCF project, after having characterised active compound solubility in $SCCO_2$ and polymer interactions with $SCCO_2$ in the previous mode, now the key uncertainty area to be characterised are the design of the melt-extruder using $SCCO_2$. Basic process options like single versus twin-screw extrusion need to be tested and characterised for process parameters like flow-rate, temperature, pressure, foaming and shaping behaviour at the die opening. In parallel, two sets of experiments are conducted. Characterisation of this uncertainty area will not only lead to critical value ranges for the cited variables. Also, it will be possible to compare application ranges with other technologies hence locating it into the spectrum of technologies enhancing solubility.

As cited above, the CR project uncertainty areas were sequentially characterised leading to an integrated multiparticulate CR system that was the result of a number of fundamental choices sequentially made after explorations of several options. Finally, proof of concept delivery of the micro-emulsion project is assumed to lead to the application domains cited above.

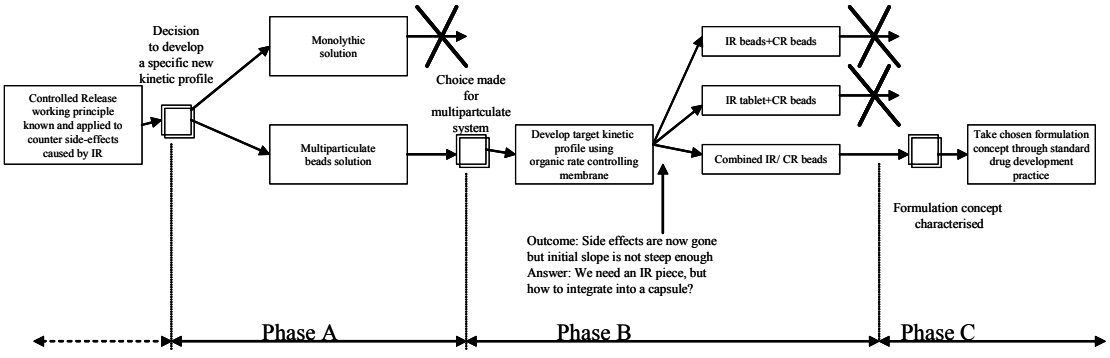


Figure 2-6: Gamma Controlled Release case process flow

External explicit learning is still used in this mode to start the experimentation process or for on-going problem solving. However, although external tacit remains the dominant learning mode the difference with the previous mode is that knowledge transfer occurs at the start of the relationship with the outside technology provider if such relationship is set up. Both Nanosuspension and SCF projects technology providers were used as a technology source and for problem solving.

> …The characterisation of the supercritical fluid-active substance interaction is done in house with problem-solving help of the University of [x], for the extrusion characterisation part we do this in collaboration with the University of [y]…. It is not a real collaboration contract we have with them, it's more consultancy, they give advice to us…no milestone agreements exist with them" (SCF)

> "… The basic technology we bought was the milling process and excipients/beads that gave this type of product this stability. We brought the sterile knowledge to the project" (Nanosuspension)

> "…Knowledge transfer happened mainly at the beginning,… later we used the relationship only for problem solving, and it was not a joint development effort…" (Nanosuspension)

So, unlike the previous mode where partners were selected as quickly as possible using kick-out criteria, after which a collaborative research effort started, here this is hardly the case. Therefore, I classify the learning approach used as sequential and intensive; technology licences are bought at the beginning and the relationship serves for problem solving, not for a joint purpose to find a solution for the problem, as exemplified in the first mode where a reciprocal intensive coordination mode is preferred.

2.5.3 Concept Application

'Concept Application' leads to a mental model characterising the full innovative drug system's application to the level of all relevant variables and their relationships, with a specification of the parameter values delivering an effective application system.

Starting from product requirements the whole application gets optimised for usage by conducting a pre-planned range of experiments. Target product requirements are defined up-front –they can only be formally modified at the gates of a stage-gate process (Cooper, 1990; 1994)- and fall within a characterised application domain.

> "…The product definition is documented in the TCDS [Target Core Data Sheet] and can evolve up to a certain point in the stage-gate process after which no more modifications are allowed" (IR)

The experimentation process consists of parallel testing performance of different candidate standard formulations delivering active substance into the human body. Previous complexity-handling modes characterised an application domain or range. This mode optimises one application within a domain, meaning the exact values of all known application domain variables need to be defined for this particular application. In the IR Tablet project three known concept or platform formulations containing the new active substance were screened in parallel. Testing against IR Tablet application domain

requirements led to the choice of a certain concept, after which finding the exact values for the different variables further optimised the formulation at lab scale.

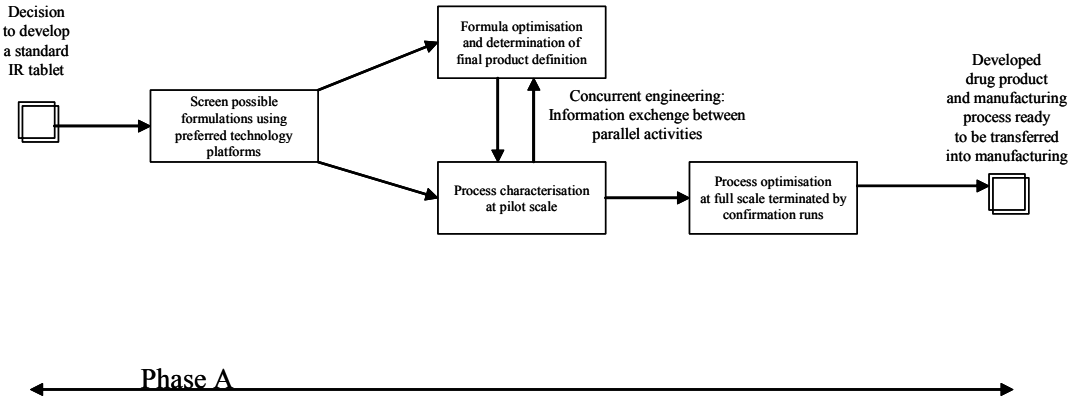


Figure 2-7: Alfa Immediate Release case process flow

> "A standard platform formulation is like a rough formula. We took three platforms; lactose, calciumphosphate, and cellulose. … In parallel we used three standard platform formulations and tested each of them for stability…then we picked one out, the best. …Then you can start refining the formula" (IR)

> "…Development was done using the 'Direct compression formulation development flow chart' [shows this documented procedure]. We use this procedure to run all our standard tablet formulations" (IR)

Further experiments optimised the up-scaled formulation concurrently with the design of the production process. Concurrency is needed while process design possibly affects formulation design, and development in this late part of the process is always under time pressure. All structured design techniques (DoE, QFD, VOC, FMEA, CTQ[19]) are used. Scientists are required to use these tools, which are part of the corporate standard Design Excellence (DEx) toolkit to conduct these experiments.

> "…[refers to ref (3)]After parallel formula concept screening the process gets developed concurrently with formula optimisation because it is possible that you should adapt the formula because of the process" " (IR)

> "…For these type of projects [standard solid tablets and oral dosage forms] we use an integrated set of development techniques under the banner of 'Design Excellence', which consists of techniques like pre-designed DoE, FMEA, VOC, QTC etc." (IR)

Learning in the IR Tablet project is explicitly following formal documented development procedures using to a maximum extent platform technologies. However, internal tacit learning was used in the platform choice and formula optimisation. The team had all the tacit experience on board to handle the innovation problems at hand.

[19] Voice of the Customer (VOC), Quality Function Deployment (QFD), Failure Mode and Effects Analysis (FMEA), defining factors Critical to Quality (CTQ), and Design of Experiments (DoE) are part of the standard DEx JJPRD toolkit whose use is compulsory in late development phases.

"…We use the standard platform formula. Of course, results of formulation work are never black or white…experience, skills are still important…if results are on the low side, which of the parallel-developed solutions should we pick out? (IR)

"…Experience built up over the years, by working with different products, and also experience in interpreting results are crucial in determining which type of filler we will use" (IR)

Detailed GANTT charts, available as templates up-front depicting all activities and their interrelationships coordinate the experimentation process. Follow-up is done through a formally documented stage-gate process where deliverables need to be shown at specific gates. Evidence needs to be given that experiments have been conducted using DEx compliant tools and best practices.

"…At each gate of the stage-gate process we have to be able to show which Design Excellence techniques have been used with what results …GANTT charts [shows project chart] are used to track progress…" (IR)

2.5.4 Complexity-handling modes descriptors

Summarizing, Table 2-2 below describes the three identified complexity-handling modes along the dimensions target setting, experimentation, learning, and coordination approaches followed by the innovation team.

An overview of the process transforming complexity-handling mode empirical indicators into summarizing sets of descriptors is provided in Appendix B.

	Complexity-handling modes		
	Concept Selection	***Concept Characterisation***	***Concept Development***
Target setting approach	• *Target defined up-front* as minimal system critical requirements to pass	• *Target moving* toward feasible application domain requirements to pass	• *Target defined up-front* as product/process requirements to pass. Only modifiable at formal stage-gates
Experimentation approach	• *Run parallel experiments to characterise* critical variables affecting response for different candidate system solutions • System solution selection • Show Proof of Concept by spelling out assumptions about the set of relevant variables and their functional relationships	• *Run parallel/concurrent experiments to characterise* all variables affecting response for different uncertainty areas within proof of concept delivered solution • Integrate uncertainty areas into limited / characterised application domain	• *Run parallel/concurrent experiments to optimise* variables values within a solution • Followed by concurrent engineering driven by QFD derived product definition • FMEA based process design
Learning approach	• *External explicit:* Mainly at project start learning from published science. Later ad-hoc for problem solving. • *External tacit:* On-going knowledge transfer by interaction between teams and external technology suppliers.	• *External explicit:* Mainly at project start learning from published science. Later ad-hoc for problem solving. • *External tacit:* Knowledge-transfer mainly at project start between external technology supplier and team. Later for problem solving. • *Internal tacit:* Use of pockets of previous knowledge	• *Internal explicit:* Learning based on formal procedure-based inquiry • *Internal tacit:* Learning by doing, based on previous internal experience
Coordination approach	• Define milestone targets that are reached if results can be shown • Estimate work package effort/ timeline based on first experiments or expert knowledge • Focus on experiments capable of selecting as quickly as possible solutions that meet all system critical requirements • Through close monitoring of progress: Eliminate as quickly as possible candidate system solutions not meeting one of the system critical requirements	• Define uncertainty areas and assumptions to be tested per area • Define milestone targets for uncertainty areas that are reached if results can be shown. • Use DoE experimental guides to systematize the testing process • Guide progress through real-time coordination of concurrent results of different uncertainty areas • Bring focus through adaptive learning i.e.; assumptions testing, learning, continue/redirect efforts to characterise the feasible application domain • Possible go back to previous mode if application domain cannot be delivered or if new application domain emerges	• Define product/process requirements to be met • Use pre-designed DoE experimental guides to systematize the experimentation process • Use GANTT type plans & schedules for cross-activity programming and tracking task completion • Monitor plan variation and act accordingly by executing contingency plans • Use of standard approaches and documented best practices to problem-solving

Table 2-2: Complexity-handling mode descriptors

2.6 DISCUSSION

2.6.1 Experienced complexity driving the choice for a complexity-handling mode

The three complexity-handling modes emerging from my study provide support for the 'learning-versus-planning' thesis (Cheng and Van de Ven, 1996; Leifer et al. 2000) put forward in the literature. The Alfa IR Tablet 'incremental' project is managed using a complexity-handling mode drawing heavily on planning and control. As hypothesized by the literature this contrasts with all other case projects, tagged 'radical' and using a trial-and-error adaptive learning-based approach.

However, this study provides evidence for a more fine-grained description of the latter learning-based approach since all five projects tagged 'radical' fall into two emergent groups, each group being managed using a different set of complexity-handling modes drawing upon the same but differently applied adaptive learning approach. Important while central to my thesis, it can be verified on Figure 2-1 providing a chronological overview of complexity-handling modes used per project in this case study, that project exogenous characteristics cited in Table 2-1 provide poor explanation for the occurrence of different observed complexity-handling modes.

Clearly, while both DNA-Zyme and Nanosuspension are both classified as 'new-to-world' and 'radical' technologies, they are managed using different complexity-handling modes at various moments in time. Also, Gamma Controlled Release, Supercritical fluids, and Nanosuspension projects have different exogenous characteristics but share the same complexity-handling approach. Therefore, I conclude that my case study evidence suggests project exogenous characteristics to be bad predictors of complexity-handling modes chosen by the innovation team to manage their projects.

Instead, based on my exploratory case study results I propose experienced complexity to explain the complexity-handling mode chosen dynamically through the course of the innovation project.

> *Proposition 1-1: The experienced complexity type facing the innovation team determines the choice for a complexity-handling mode*

The innovation team handles experienced complexity by constructing a *mental model* of the problem-solving situation. Mental models specify the variables, their relationships and their value ranges that are relevant for understanding and describing the problem and provide the solution space within which the problem will be solved. Key to understanding the subjective nature of experienced complexity reduction through mental models is the acknowledgement that ambiguity and uncertainty are not exogenously given characteristics of the problem to be solved. Instead, mental models are constructed through a creative problem framing and solving process in which levels of ambiguity and uncertainty are dynamically chosen and reduced by the innovation

team (Schrader et al. 1992). My case evidence below (see also Appendix A) provides support for this view.

In the following I characterize experienced complexity using levels of ambiguity and uncertainty facing the innovation team. Table 2-3 below provides the definitions for the different constructs emerging from the cases.

Experienced Complexity Type	**PoC Ambiguity-based**	**Ambiguity-based**	**Uncertainty-based**
Ambiguity level	Proof of Concept (PoC) critical variables and their functional relationships unknown to the team	Only PoC critical variables and their functional relationships known to the team	All application domain specific variables and their functional relationships known to the team
Uncertainty level	Value ranges of critical variables unknown to the team	Not all value ranges of critical variables known to the team	Value ranges for all application domain specific variables known to the team

Table 2-3: Experienced complexity typology

Empirical indicators of uncertainty and ambiguity during '*Concept Selection'* evidence that the complexity experienced by the team is Proof of Concept (PoC) ambiguity-based. Ambiguity level is PoC-based since the mental model containing relevant variables to solve the problem at PoC level is not yet formed within the team. The innovation team goes through a creative search process among potential candidate solutions, composed of variable sets, to discover the set of variables that might generate a system that has the required solution critical functionality. Since during this process the winning mental model and its related variable sets is not yet known, by definition neither could the variable value ranges be known. Hence, although uncertainty level is also high this is not relevant for the choice made by the innovation team. First, the winning mental model must be found, showing PoC-level performance on the solution critical requirements. Therefore, I call this experienced complexity type *Proof of Concept ambiguity-based*.

In the DNA-Zyme, SCF, Micro-emulsions, and Nanosuspension projects this mode was chosen while no evidence existed to the team that proof of concept had been delivered and documented before for these new-to-world technologies. Nanosuspension was brought to this level before, but not for the sterile technology-based application domain required for the specific PharmaCo product.

> "We knew from the answers the delivery for that sort of fragments was going to be a problem; we knew that simple techniques won't work. So we already knew that the barrier was pretty high"… The approach that was started together with a group in discovery…was to canvas real novel approaches for transfaction, given the fact that most things had already been tried in that area. (DNAZyme)

> "We also went to the scientific literature and made a special point to specific scientific media. In the beginning there was a relatively long list, we looked at anything that might be

useful; so at one point we must have had 10 or 12 different approaches…You knew that you had to look at some pretty avant-garde therapies for approaches" (DNAZyme)

"…There was so much published on so many failures, that we knew we could not do with the traditional approaches so we started already at the beginning with things that were pretty off the wall ... We knew, that when we had any success, it would be from novel approaches and not from things that have been tried fifty times before". (DNAZyme)

"…Mid 2000 Proof of Concept was delivered in the Supercritical fluids project when polymers in interaction with $SCCO_2$ showed that they can decrease the meltextrusion temperature and that active substance can be made soluble in $SCCO_2$ under certain critical conditions" (SCF)

"[At the start of the project] it was New to world technology" (Microemulsions)

"… The basic technology we bought was the milling process and excipients/beads that gave this type of product this stability. We brought the [internal tacit] sterile knowledge to the project…finally we had a working formula…this was the end of Concept Selection" (Nanosuspension)

The Controlled Release (CR) Tablet project team did not choose this complexity-handling mode while the technology had been characterised before to proof of concept level by an outside technology provider. The concept of delivering a specific kinetic profile was known. Only, it had to be developed for this specific product application domain following this specific curve. This leads me to formulate the following proposition;

Proposition 1-1a: A 'Concept Selection' complexity-handling mode will be chosen if the innovation team experiences Proof of Concept ambiguity-based complexity.

Second, case evidence suggests that the choice to manage the innovation project following the '*Concept Characterization'* complexity-handling mode is made by the team whenever it is facing an ambiguity type sense-making opportunity. Ambiguity is experienced since only a mental model containing relevant variables to solve the problem at proof of concept level has been formed within the team. Since neither problem relevant variables nor their functional relationships are known beyond proof of concept, by definition neither could the variable value ranges be known beyond this level. Although uncertainty level is still high this is not relevant for the choice made by the innovation team. Therefore, I call this experienced complexity type ambiguity-based and propose the following;

Proposition 1-1b: A 'Concept Characterization' complexity-handling mode will be chosen if the innovation team experiences ambiguity-based complexity.

Concept application domain characterisation will gradually resolve ambiguity and reduce uncertainty to a level where the innovation team's mental model of the system to be designed contains all variables with their value ranges relevant to start application development. The CR project faced an initially ambiguous situation but variables and

their functional relationships relevant for proof of concept performance were accessible to the team, which led them to the choice of *'Concept Characterisation'* as complexity-handling mode. The right choices in a number of uncertainty areas were still to be made to get to an optimal kinetic profile and to make the technology ready for implementation in an application to be delivered to the market.

> "The challenge was to develop a CR profile that would rise very quickly, stay up the whole day, and fall down in the evening so patients can sleep, all this without peaks to prevent side-effects"…The overall proof of concept for this was known. However, what was not known was how to get this significant quick rise [the application domain of interest] at the beginning" (CR)

> "…We knew theoretically how it should work, techniques were described in the literature how to do it at lab scale. However nobody had ever done clinical studies with it…so we tried a number of solutions in parallel without being able to fall back to standard technology platforms" (CR)

For the Nanosuspension project the innovative technology needed to be made sterile and ready to accept company specific compounds.

> "During Phase A you select a concept, in the second step you optimize your concept" (Nanosuspension)

> "…finally we had a working formula…this was the end of Concept Selection…Then we started characterising the product…every parameter; particle size as a function of time, sedimentation as a function of time, etc. we checked for robustness" (Nanosuspension)

After delivery of the initial proof of concept that certain polymers in interaction with $SCCO_2$ can lower the required meltextrusion temperature hence extending its application range to thermolabile active ingredients, the SCF project team chose this complexity-handling mode to further develop the technology to a level were it can be used for application development.

> "…As soon as we have a bit of a view of the process [Phase A] we will use Design of Experiments to characterise and optimise the process" (SCF)

> "As soon as we find out if and how we can get the active substance into the SCF we will have to check *how good* the active can be solubilized, or *which volume* you can usefully bring into the extruder, what is the stability, bio-availability, of the mixture?"(SCF)

Finally, I propose that the choice to manage the innovation project following the '*Concept Application'* complexity-handling mode is made by the team whenever it is facing an uncertainty type sense-making opportunity with a very low level of residual ambiguity. Only uncertainty, no or few residual ambiguity is experienced since for a specific application domain a mental model containing all relevant variables with their respective value ranges has been formed within the team. Since the level of experienced ambiguity is negligible it is not relevant for the choice made by the innovation team. Therefore, I call this experienced complexity type uncertainty-based. This leads me to formulate the following proposition which is in line with the proposition formulated by Schrader *et al.* (1992) that problems will be framed involving little ambiguity if the problem-solver has successfully solved apparently isomorphic or related problems previously:

Proposition 1-1c: A 'Concept Application' complexity-handling mode will be chosen if the innovation team experiences uncertainty-based complexity.

The Alfa IR Tablet project started with a full mental model of the problem-solving situation available. Residual ambiguity was negligible since both projects could be developed using platform technologies like concept formulations and documented best practices. As soon as the steep kinetic profile CR concept was characterised, in its last phase it could be handled using standard development approaches.

> "The target in Phase A and B is evolving in the sense that you're discovering domain limits. While in Phase C you know the limits but now your efforts are directed to developing a product…The endpoint of Phase B is a working concept that still needs to be developed. In Phase C you know it works" (CR)

The innovation teams chose this mode while they had successfully solved apparently isomorphic problems before leading them to rule out ambiguity and to choose for a complexity-handling mode that is strong in delivering results at pre-planned stage-gates, using rigorous experimentation techniques.

2.6.2 Experienced complexity dynamics

The incapacity of project exogenous characteristics to explain complexity-handling mode transitions is evidenced by a number of my cases. The Microemulsions project will soon change from '*Concept Selection'* mode to '*Concept Characterisation'* mode, although it will still be a 'Super-High-Tech' technology while still based on non-existent technologies at project initiation. The same holds for the SCF project that has changed complexity-handling mode but is still a 'High-Tech' project. Also, it remains to have a 'project uncertainty profile' (Loch et al, 2000) characterised by ambiguity, variation and risk.

Instead, by taking an interpretative approach my case data lead to a better understanding why experimentation approaches vary dynamically over the course of the project. By taking the perspective of the mental model of the problem-solving situation, one can see that the team gradually builds up understanding of the innovative system to be designed. The decision to transit to a new complexity-handling mode is driven by the perceived completeness of the team's mental model. As long as ambiguity has not been sufficiently resolved or uncertainty sufficiently reduced, the team will stay in a certain complexity-handling mode. This leads me to formulate the following propositions;

Proposition 1-2: The decision to change complexity-handling mode is determined by the innovation team' perceived completeness of their mental model of the problem-solving situation at hand, operationalised in the proven delivery of minimum system requirements.

The decision to transit from '*Concept Selection'* to '*Concept Characterisation'* mode is made by the team as soon as solution critical requirements are met. High initial ambiguity must be resolved to the level that the emerged mental model contains all critical variables and their relationships necessary to deliver Proof of Concept.

Proposition 1-2a: The decision to transit from 'Concept Selection' to 'Concept Characterisation' complexity-handling mode is made by the innovation team as soon as a mental model has emerged that contains all critical problem-solving mechanisms, solution variables and their relationships necessary to deliver Proof of Concept.

Nor the DNA-Zyme, nor the micro-emulsions project have reached this transition since no delivery solution has been reached in the former nor has a definitive set of self-emulsifying polymers been found yet in the latter. The team does not know the definitive set of critical variables nor their relationships delivering proof of concept yet.

> "The transition to the next phase [from A to B] is really dependent upon the results you get. And if you don't get a result you have to stop unless the literature provides you with further clues" (Microemulsions)

Mid 2000 proof of concept was delivered in the Supercritical fluids project when polymers in interaction with $SCCO_2$ showed that they can decrease the meltextrusion temperature and that active substance can be made soluble in $SCCO_2$ under certain critical conditions. This led to the project team's decision to make the transition to the *'Concept Characterisation'* mode where the design of the melt-extruder using $SCCO_2$ is further characterised.

> "…First, we looked whether products of our pipeline could be brought into the extruder using supercritical fluids, then [in Phase B] we do a full physico-chemical characterisation of the whole extrusion system, later [Full Development] it will be scaled-up to full production level (SCF)

Second, the decision to transit from *'Concept Characterisation'* to *'Concept Application'* mode can be made by the team as soon as the application domain for the concept is characterised and solution critical requirements are met. Ambiguity must be resolved to the level that the emerged mental model contains now all variables, their relationships and value ranges necessary to deliver an application within the characterised concept domain.

Proposition 1-2b: The decision to transit from 'Concept Characterisation' to 'Concept Application' complexity-handling mode can be made by the innovation team as soon as a mental model has emerged that contains all problem-solving mechanisms, solution variables, their relationships and value ranges necessary to deliver an application.

In the Controlled Release project the transition decision to go for application development was made after the concept showed performance following a kinetic release profile that was hypothesized by the team. Then the team could switch to a project management approach tailored to manage 'incremental' projects (see process flow in Figure 2-6). The same situation held for the Nanosuspension project, which was characterised immediately on the commercial product and scaled up to full production (see process flow in Figure 2-5).

No other case projects made this transition yet since the mental model depicting the innovative delivery system has not crystallised to a level that it can be used to develop an application – as applicable to DNAZyme, Microemulsions, and Supercritical fluids cases.

> "I think the transition into Full Development is made as soon as there's a compound in need for this technology in its relevant application domain…Now the technology is already sufficiently characterized for one type of polymer to allow us to test its solubility in micro-emulsions for one specific compound falling within this application domain" (Microemulsions)

> "You develop an application on model compounds, not for a final R-number [specific drug project] We will only transfer into Full Development when we will be sure it will be bio-available, enhances solubility and is non-toxic. This needs to be shown first on model compounds (Microemulsions)

Finally, during Concept Characterisation and Concept Application modes a decision can be made by the team to go back to one of the previous modes if a fundamental problem arises or a new situation emerges preventing moving forward.

> *Proposition 1-2c: The decision to move back from 'Concept Characterisation' or 'Concept Application' complexity-handling modes can be made by the innovation team as soon as a situation arises where the present complexity-handling mode does not lead to further absorption and/or reduction of experienced complexity.*

An example of a problem where the innovation team decided to go back from Concept Application to Concept Characterisation is provided by the Gamma CR project where a drug candidate was in Development featuring a specific kinetic profile. However, the latter showed to be insufficient to meet the needs of the patient population. Therefore, the team was forced to revisit the problem and characterize a CR solution that was known to work at proof of concept level but that had not been applied before in the project-specific context. The SCF case is an example of a technology that had shown proof of concept and was characterized for the meltextrusion application domain. However, after this characterisation the team could have decided to go for a human bone-joining application, in which case it would have to go back to Concept Selection to find ways to get to proof of concept for this new application domain.

2.7 CONCLUSIONS

2.7.1 Implications for managerial practice

There seems to be a desire on the part of management to understand how to manage effectively the development of radically new products if, of course, it can be managed (O' Connor, 1998; Veryzer, 1998). Previous studies focussing on the difference of the radical versus incremental innovation process conclude that the development of disruptively innovative products does not seem to follow conventional stage-gate

processes and find a degree of informality with respect to how this development process is managed (Veryzer, 1998).

However, by focusing on the complexity experienced by the innovation team and the resulting mental modelling process my results indicate that there is not more informality but more ambiguity involved in managing for 'radical' as opposed to 'incremental' innovation. Experienced complexity, characterized into types by specifying levels of ambiguity and uncertainty facing the innovation team, is proposed to drive their choice for a specific experimentation and project management approach, I called complexity-handling mode.

'Concept Selection' is a complexity-handling mode leading to a mental model characterising the *innovative solution's core*, depicting the critical variables and their relationships affecting solution proof of concept level performance. During the process ambiguity is gradually resolved by selecting one or more candidate solutions that fit solution Proof-of-Concept-level critical requirements.

'Concept Characterisation' is a complexity-handling mode leading to a mental model characterising the *innovative solution's application domain* and its boundaries, depicting relevant variables and their relationships affecting solution performance. Using an adaptive learning process the wanted application domain gets characterised, resolving ambiguity beyond Proof of Concept level down to all variables and their functional relationships affecting solution performance. Also, critical value ranges are defined for all variables to indicate the feasible boundaries of the application domain.

'Concept Application' is a complexity-handling mode leading to a mental model characterising an *innovative solution's application* to the level of all relevant variables and their relationships, with a specification of the values delivering a working application system.

The conceptual framework presented above offers project managers in practice a diagnostic tool they can use to dynamically choose over the course of the project for a specific complexity-handling mode, contingent upon the type of complexity they're experiencing. Present diagnostic tools fall short against this framework while they are not dynamic, only considering the situation at the outset of the project, and take into account project exogenous characteristics as opposed to characteristics related to the complexity experienced by the innovation team (Shenhar and Dvir, 1996; De Meyer et al. 2001). Or, they don't consider the role of ambiguity and only focus on uncertainty types to guide the choice for a specific project management approach (De Meyer et al. 2001).

2.7.2 Toward a more fine-grained understanding of managing radical innovation

The purpose of this exploratory research project was to develop a more fine-grained understanding of radical innovation experimentation practice. A change in paradigm was needed to develop this more detailed understanding. This led to a number of new insights that can help me further develop a theory of experimentation strategy needed to manage radical innovation projects for performance.

My empirical results confirm the widely claimed distinction between managing incremental and radical innovation projects. However, by taking a complex systems perspective I find that, contingent upon the type of complexity experienced by the team, innovation projects can be managed using three different approaches. This result is highly relevant to practicing innovation teams who are offered a framework to dynamically decide upon the experimentation approach to be followed. Also, the framework deals with the oversimplifying binary distinction between radical and incremental innovation projects and develops a more fine-grained understanding of the process.

However, whether or not the conceptual framework identified in this research context is also applicable to other contexts is an empirical question. Further empirical studies in other contexts are needed to test internal and external validity of the framework set out above, which will further enrich my understanding of the radical innovation process. Second, to build a theory of experimentation practice for radical innovation I need to investigate and explain possible relationships between complexity-handling approaches chosen and process or outcome performance.

My future research will explore these questions in a pharmaceutical Discovery context. In the next project I will literally replicate the propositional model developed in this exploratory case study and try to document experimentation strategies used in the different modes. The latter will provide the basis for theory building around their performance in the innovation process.

3 Exploring Experimentation Strategies for Radical Innovation in Pharmaceutical Discovery: A Bayesian Perspective

3.1 INTRODUCTION

Whilst the competitively differentiating role of radical innovation[20] projects is recognised, its actual detailed operation is far less understood. Therefore, the purpose of my research is to build a theory of radical innovation experimentation practice. Previous literature suggests that radical innovation projects are managed using an experiential learning approach as opposed to incremental innovation projects being managed using a stage-gate driven planning approach. However, it fails to indicate if and how the experimentation approach evolves over the course of the radical innovation process. I argue that this shortcoming has its roots in the overemphasis of the extant literature on exogenous overall innovation project characteristics like innovation outcome, technological uncertainty or complexity level and their relationship to project and firm performance.

In contrast, in the exploratory research project described in the previous Chapter I proposed to shift paradigms by taking an interpretative stance modelling uncertainty and ambiguity as two dissimilar components of complexity experienced by the innovation project team. My empirical evidence suggested the fresh insight that experienced complexity type endogenously drives the experimentation approach or *'complexity-handling mode'* chosen by the innovation team to run innovation projects. A mental model of the product definition depicting variables, their interrelationships, and their value ranges was proposed to emerge during the radical innovation process, and used during a subsequent incremental innovation project.

However, no claims were made as to the performance implications of the innovation teams following a specific complexity-handling mode. Therefore, in the remainder of my research I explore this process-outcome question in two phases; In a confirmatory case study project described in this Chapter I will literally replicate (Yin, 1989) the complexity-handling mode findings of my exploratory case study and describe experimentation strategies used during 'Concept Selection' mode. My case evidence suggests that the findings pertaining to this latter mode of the proposed model and its transition to the subsequent 'Concept Characterisation' complexity-handling mode can be replicated to the context of pharmaceutical Discovery. However, the evidence provided by this case cannot be used to corroborate the 'Concept Characterisation' and 'Concept Application' parts of the model. In addition, this case study documented alternative experimentation strategies used in pharmaceutical

[20] In my research I define a radical innovation "as a product, process, or service with either unprecedented performance features or familiar features that offer potential for significant improvements in performance or cost, using non-existing or non-proven technologies that did not fully exist at the start of the project *with* high market uncertainty (Leifer et al. 2000; Lynn and Akgun, 2001).

Discovery, all specific applications of the 'Concept Selection' complexity-handling mode. Variance in experimentation strategies used is proposed to be explained by the extent 'front-loading' is used by Discovery management. Research propositions are formulated, linking alternative experimentation strategies to performance.

Furthermore, a Bayesian methodology is proposed to evaluate the qualitative and predictive performance of these front-loaded experimentation strategies during 'Concept Selection' mode. Finally, a rationale is developed to test the research propositions using simulation. In the next and final project of this research, this will provide the basis for a formal representation of these strategies as adaptive systems, which is needed to build a computer simulation model of the alternative pharmaceutical discovery experimental designs and to use the latter as an instrument for theory development.

The remainder of this paper is focused around the following research questions in this specific context: (1) 'Can the Pharmaceutical discovery process at PharmaCo be used to literally replicate the proposed model on complexity-handling? (2) 'Which experimental designs are used?' and, (3) 'How to describe performance of an experimental design?'

3.2 RESEARCH METHODS

I studied experimentation approaches used –my unit of analysis- in the European Discovery Unit of PharmaCo, the same global top-10 pharmaceutical company I was involved with in my previous research project. Only now, I chose pharmaceutical Discovery as a research context since its 'fuzzy front end' nature gives the best chance of finding experimentation approaches handling high levels of ambiguity as described in my 'Concept Selection' complexity-handling mode cited above. In contrast to the exploratory research conducted in Project 1 where I studied technology and product innovation projects, here I focus solely on radical *product* –called chemical compound or NME[21] - innovation projects.

3.2.1 Literal case study replication logic

Desk research was conducted of Discovery Research project specific archives, written departmental procedures, and the formally documented NME discovery process of the European part of PharmaCo[22]. I used visual mapping and temporal bracketing (Langley, 1999) to make sense of the process data. The resultant high-level process map (Figure 3-4) was used to conduct 15 interviews with Discovery Research functional and project

[21] New Molecular Entity

[22] A process flowchart of the European operation of PharmaCo's Discovery Research process is available with the author and referenced in Appendix C under number (10). It was not included into this document for reasons of confidentiality.

managers around the 'How' and 'Why' of the 'Portfolio' part[23] of the discovery process they had been using in their projects to solve their radical innovation problem. Finally, I used the theoretical framework developed in Project 1 as a vehicle for generalizing to the discovery process case studied in Project 2.

3.2.2 Data collection and coding

To probe into the 'How' and Why' of the experimentation approaches used in Discovery Research a total of 15 interviews were conducted with representatives from medicinal chemistry, ADME and Toxicity Research and a senior scientist of the bio- and cheminformatics group. Interviews lasted between one and two hours, were tape-recorded and transcribed. For each interviewee the first interview session focussed on the fit of the proposed Project 1 framework with the experimentation approach they had been using in their cited 'Portfolio' projects to solve their innovation problem. Thus probing into the How? and Why? of the experimentation approaches used for their innovation projects. Subsequent interviews with selected Discovery scientists focussed on deepening specific topics like the functioning of surrogate marker chains, front-loaded experimentation strategies, and details of the multi-factorial compound optimization process used in Discovery Research. Also, interviews were used to validate the functioning of the simulation model.

Data coding consisted of mapping empirical indicators of the phases that emerged from visual mapping and temporal bracketing described above, to the descriptors of the various approaches used as described in the proposed Project 1 model.

3.2.3 Reliability and validity considerations

Whenever confirmatory case study findings did not fit the proposed Project 1 framework, the latter was modified to fit the new empirical data (Yin, 1989) hence improving its external validity. Internal consistency was verified by using additional sources of data or through verification by the original informants.

To ensure reliability a case study database was developed to formally assemble qualitative and quantitative evidence material. Empirical indicators linked to this material providing evidence for replicating the proposed Project 1 model can be found in Appendix C.

[23] The 'Portfolio' part of the Discovery process is PharmaCo specific terminology used to indicate the part of the Discovery starting when a biological target has been validated and the compound search and optimization process can start.

3.3 CASE STUDY RESULTS

3.3.1 PharmaCo's Drug Discovery Research context

As in every global top-10 pharmaceutical R&D operation, PharmaCo's Drug Discovery research (DD) is the front-end part of the overall product innovation process for chemical compounds typically lasting between 9 and 12 years (Figure 3-1).

Starting from biological targets defined in Disease Area's like Gastrointestinal (GI) or Central Nervous System (CNS), Drug Discovery's mission is to generate 6 to 8 new molecular entities a year with the desired biological action on specified clinical targets. It focuses on (1) target identification and target validation, which comes down to finding the biological areas for which an active compound needs to be developed. (2) A compound screening, profiling and optimisation process leads to a candidate compound being promoted to NME status.

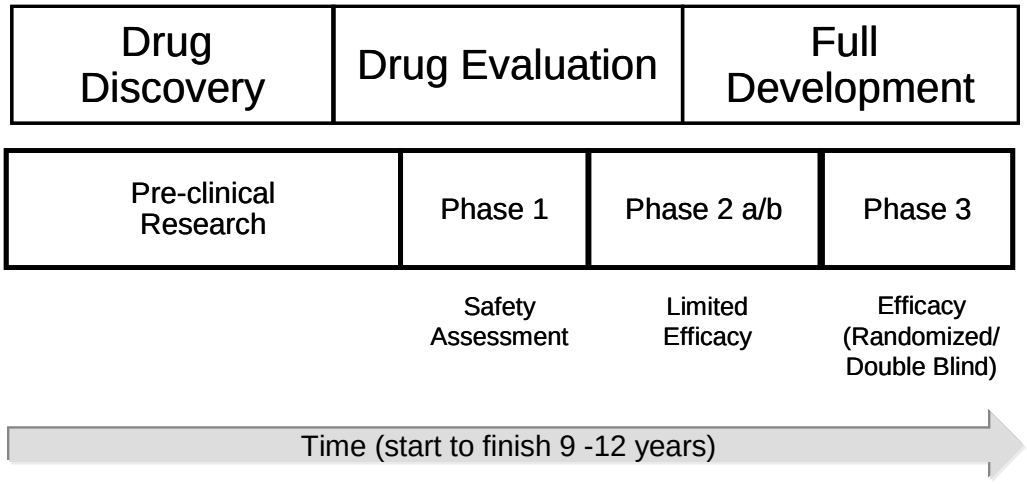


Figure 3-1: PharmaCo's R&D organisation and process

The aim of Drug Evaluation (DE) is to move candidate compounds as quickly as possible into human trials in Full Development. The DE process is designed to move quickly through pre-clinical evaluations of toxicology, human metabolism, and chemical-pharmaceutical formulation potential into early clinical trials in humans.

Full Development (FD) is responsible for further clinical and non-clinical development of the candidate compounds. Clinical trials are driven by a clinical development plan that tries to answer key questions on the compound's therapeutic and pharmacological effects in patients. In addition, safety margins and pharmacoeconomic parameters of the drug treatment are determined. The key clinical development process stages are increasingly complex and costly. Therefore, they are divided into three stages; Phase I studies are normally conducted in healthy volunteers and provide information on acute tolerability, safety and basic pharmacodynamic and pharmacokinetic properties (PK/PD). Phase II examines pharmacological and therapeutic effects in patients, defines dose response relationships, therapeutic ranges

and safety margins. Phase III studies conduct large-scale confirmatory trials to demonstrate efficacy, safety, and pharmacoeconomic parameters of the drug treatment.

In parallel to this process, the NME or active drug substance, and drug product, its drug delivery vehicle, are further developed from Phase IIb through Phase III and up-scaled to commercial manufacturing scales in the non-clinical 'Chem-Pharm' division of Full Development. All product and technology development projects case studies used in my exploratory research Project 1 originate from this latter part of the organisation.

3.3.2 Complexity-handling throughout the R&D process

The R&D of a drug candidate is a milestone-driven and highly regulated problem-solving process, increasingly engaging more resources to resolve the complexities of therapeutic and pharmacological effect. **Figure 3-2** provides an overview of the parameters involved in the design of a drug at PharmaCo. Key milestones include 'Compound Transfer' (CT), the handover between DD and DE. 'First in Human' (FIH) indicates the completeness of the drug candidate characterisation up to a level that an Investigational New Drug (IND) application can be submitted to the regulatory authorities. Handover of drug product and drug substance between DE and FD is done during an 'At Risk' (AR) meeting around the clinical Phase IIa milestone. Finally, the product's therapeutic effect and safety profile is fully characterised when a New Drug Application (NDA) is filed and approved by the regulatory authorities.

Parameters	DISCOVERY RESEARCH	PRE-CLINICAL RESEARCH	PHASE I	PHASE IIa	PHASE IIb	PHASE III
Biological Activity • Potency • Specificity • Dosing regimens	NME proven to show biological activity on a human receptor or enzyme assay	No change	initial estimates in healthy volunteers are made of parameter ranges like dose-plasma concentration profiles, maximum safe doses and concentrations	Assessment and confirmation of the therapeutic concept in patients	Further characterization of dosing regimes including strategies to optimize dosage for individual patient groups based on identification of subgroup specific PK/PD exposure-response relationships.	Further documenting clinical efficacy and variability estimates of dose-response due to PK/PD. This information allows getting even more specific on dosing regimens for targeted populations.
Pharmacological properties • PK/PD • ADME • Solubility/permeability • Toxicity/ Cardiovascular safety	NME characterised for basic ADME, solubility, and permeability in animal models (in-vivo).	NME further characterisation of PK and toxicity	initial estimates in healthy volunteers of routes of metabolism and excretion.	Affirmation of acute tolerability, maximum safe dose, and lack of acute safety issues in patients.		Further documenting clinical safety, adverse reaction profiles, dose-response due to PK/PD.
Drug substance properties	Basic NME chemical structure	Preliminary Drug Substance characterization. Scale-up first GMP batch	Synthesis route optimisation	Synthesis route optimized	Up-scaled manfacturing process defined	Up-scaled manfacturing process, cleaning validated, and transferred
Drug product properties		Phase I oral solution/ suspension formulation		Take home formulation for Phase IIa	Commercial formulation, methods and packaging, manufacturing process defined	Commercial formulation, methods and packaging, manufacturing process validated and transferred

Figure 3-2: Parameter overview involved in designing a drug

As can be verified in Figure 3-2 ambiguity gets resolved in the course of the process by gradually adding more parameters into the drug design. During discovery and pre-clinical research, the focus is on selecting those candidate compounds showing maximal biological activity with minimal pharmacokinetic problems in a reduced and highly controlled environment of human tissue assays and animal models. From clinical Phase I through III, uncertainty on biological activity and pharmacokinetic parameters is reduced by characterizing dosing ranges for specific target patient groups. In parallel, starting from the characterized NME in Discovery, ambiguity gets resolved by gradually adding more drug substance and drug product properties into the development equation, until an up-scaled optimal chemical synthesis route and a commercial drug delivery formulation is fully characterised by Phase IIb.

3.3.3 Complexity-handling throughout the discovery research process

PharmaCo's discovery and pre-clinical research process[24] is organised in two major phases; an Exploratory and a Portfolio-managed phase (Figure 3-3). During exploratory research the focus of the Disease Area teams is on finding biological targets, understanding mechanisms of action of candidate compounds, and finding first confirmed hits.

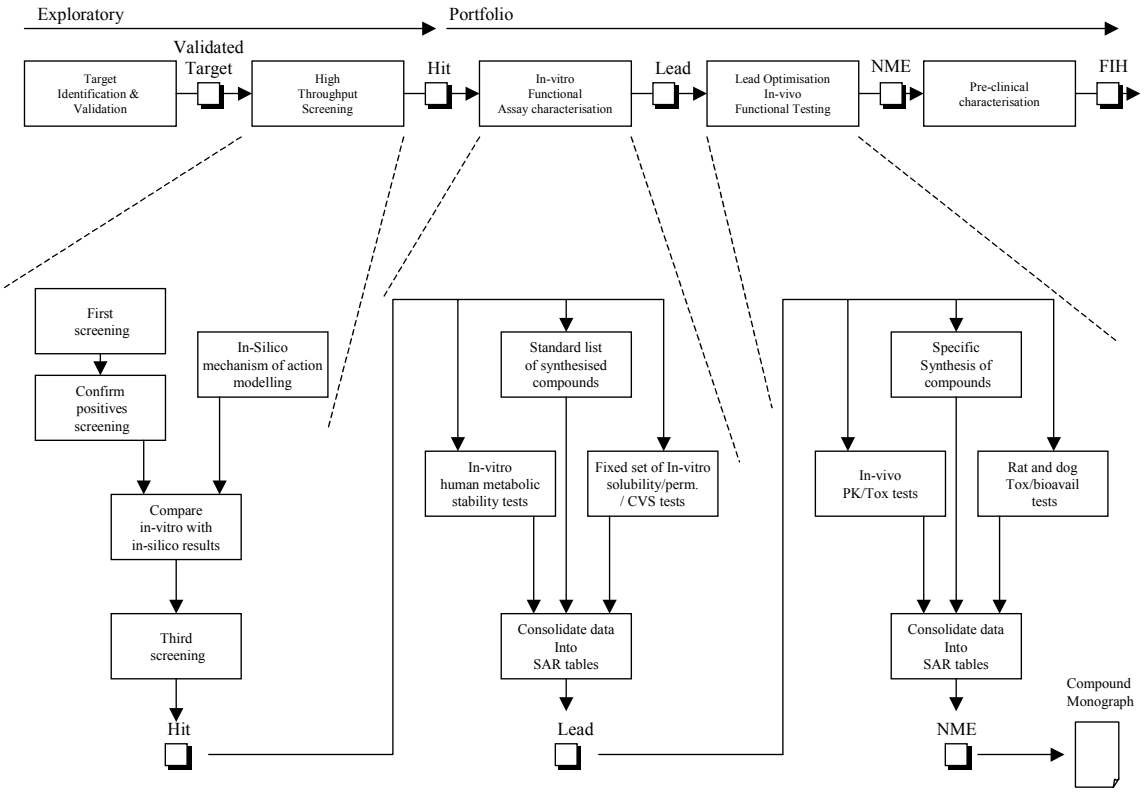


Figure 3-3: PharmaCo's discovery and pre-clinical research process

[24] See footnote 22 above

During the Portfolio phase a Project Team is created that optimizes the hits further and tries to turn them into a NME, to be transferred at *'Compound Transfer'* to DE for further pre-clinical characterization and testing.

Target Identification & Validation (TI/TV) can be based on known compounds from the Disease Areas for which a biological target or mechanism of action needs to be found, or it can involve looking at gene expression patterns in response to drugs in animal models of disease, or it can be de novo target finding. Of the 20 targets going to High Throughput Screening (HTS), 2 to 3 are genuinely novel. Once the target is validated, a receptor binding or enzyme assay is developed and transferred to HTS. PharmaCo strongly believes in having a validated target before starting the chemistry process. As the VP of Medicinal Chemistry testifies;

> "Very often you take the high- risk approach for a new target, and then you have only about 30% chance of getting some efficacy meaning your compound has no effect on the disease. A lot of companies who have gone through this genome-based drug discovery have experienced this and therefore we stress target validation so much".

HTS is run in three screens. The first could screen up to 150 to 200 thousand compounds. Positives are selected and confirmed in a second screen. Identified hits and non-actives are compared with the virtual hits of an in-silico model that also takes into account extra compounds in the library not screened in the first assay. The comparison of theoretical predictions with experimental in-vitro results in a list of false positives, or compounds that were positive in HTS but shouldn't be, and false negatives, or compounds that should have been positive in HTS. The third screen runs these lists and can already include a dose-response study. Finally, HTS results in a report identifying numbers of confirmed hits, active compound classes, and applied statistics[25]. Since PharmaCo believes in the benefits of a 'Frontloaded Discovery' experimental design (see later) both functional biological activity and drug-likeness have now been studied in-vitro and in-silico. Thousands of candidate compounds have been screened in a cell-based model for their binding affinity with clinically relevant receptors, and their selectivity relative to other systems. Mechanisms of Action and predictive drug-likeness indicators have been studied in-silico. About three to five compound classes of the twelve to twenty coming out of HTS -a compound class contains a series of chemical compounds featuring a similar chemical structure- are passed to Hit-to-Lead for further research. This is a key milestone for which Discovery management approval is required, since it marks the transition between 'Exploratory' and 'Portfolio' research. To qualify for transition, management checks amongst others the availability and quality of multiple promising compound classes, the drug likeliness of the hits, ease of chemistry, chemical and biological action plans, attractiveness of the target, patent situation for the hits, and the competitive situation[26].

Chemistry resources are allocated at Hit-to-Lead (H2L). The process is designed to search for problems to be expected with the 3 to 5 compound classes and, based on

[25] See footnote 22 above

[26] See internal policy document: 'Assessment for transition from HTS to H2L' is available with the author and referenced in Appendix C under number (11). It was not included into this document for reasons of confidentiality.

this knowledge, to reprioritize the compound classes for final selection and optimization in Lead Optimization (LO). Or, as described by the VP Medicinal Chemistry:

> "In H2L I'm looking at series as an entity and ask questions about these series. One of the questions is whether I can make molar modifications to this class of compounds to improve my primary activity? What I would not do is take one position in the molecule and make many changes there. Obviously, doing one or two already shows me how I can improve so scientists could have the tendency to do it and find the best one. However, this is not my purpose. I want to know whether there is a trend so I can improve from that model the parameters of interest. Chemistry-wise that means I will make some analogs on all of the positions of the molecule, which I can modify easily. Ease of chemistry is another one I would consider at that moment in time. It is really exploring the potential of the series. In LO you have explored the potential of a series, now I think I can find within that series the molecule I'm looking for".

Ambiguity gets further resolved by including an extra set of properties probing in-vitro for human metabolic stability (ADME), solubility, permeability, toxicity, and cardiovascular safety (CVS). Ideally, 2 to 3 lead classes with maximal effect and minimal ADME-Tox-CVS problems are selected for further optimisation in in-vivo animal models during LO. Essentially, following a flowchart of tests, new compound analogs are synthesized to get a feel for the optimization potential of the chemical classes. However, the unit of analysis at this stage is still at the level of the chemical class and not at compound level yet. The objective of H2L is to get a preliminary Structure-Activity relationship (SAR) and a feel for the type of problems one might encounter. As testified by the Head of ADME Research;

> "At the end of H2L you've identified what your compound's problem is because all compounds have problems. You have brought up your preliminary SAR whether in potency, permeability or whatever, and you've shown that you can get around that problem. Then you go into Lead Optimization and the more laborious in-vivo work coupled with the functional in-vivo work."

As such, the new compound class can be seen as one of the prototypes that get selected in H2L. The two to three best ones go through to LO where, hopefully, a compound comes out of the selected classes. Analog compounds are synthesized following classical or HTS methods. H2L typically takes 3-6 months.

Lead Optimization takes the H2L selected compound classes through in-vivo animal models. Pharmacokinetic, toxicity, bioavailability, and tissue distribution properties are characterized in rats and dogs using a battery of flowcharted tests. The synthesis plan for the prototype compounds is adapted based on the problems identified during H2L. Flowcharts of tests are used throughout the Portfolio process as quoted by the VP Medicinal Chemistry;

> "A flowchart splits testing into different components, at the beginning you have high-throughput and predictive methods, then you have in-vitro methods, and then you obviously want to get to the more mechanistic studies where the throughput is lower but the predictability is higher. Criteria get harder as you move along, but also the methods become more resourceful, more difficult to perform. In H2L it is more a generic type of flowchart while in LO it's an adjusted one related to the issues you have seen. The flowchart in LO is built upon the perception of the series in H2L and you get specific biology questions, which are disease related" [and further]... "In H2L we make about 50 analogs within each series. Ideally you don't do any chemistry in H2L because you have enough information in the

> compounds you already have. But this is unrealistic. So, I just want to do a minimum of effort to answer the question; Is this a good series? The goal is not do chemistry. In LO we go with 2 or 3 series and make about 2000 analogs of a specific molecule to find the best one within that series".

A multi-factorial optimisation cycle is conducted until the targeted SAR is reached. After 'First Glance', it is decided whether further lead optimisation is necessary, if not, at *'Compound Transfer',* the compound is transferred as an NME into DE.

3.3.4 Building a compound's mental model through surrogate marker chains

During the Discovery process outlined above a mental model is built of the innovative solution to the problem; the biological target that needs to be interacted with as shown in a positive SAR. In the following, I will focus on the interpretive dimensions of the mental modelling process in the specific context of the Discovery case, and on its implications for performance management.

Three critical areas can be distinguished in the mental model of the solution taking shape in the collective mind of the Discovery Team; biological activity (P), bio-availability (B), and toxicity (T) of the chemical compound to be designed. The first area is about characterizing pharmacophores, particular structures within the molecules, delivering potency and selectivity of the candidate compound against the target. Bio-availability is a set of complex in vivo disposition processes[27], generally called ADME or drug-likeness properties, needed for a drug to act systemically, which is solubility, permeability, and survival past the liver. Toxicity deals with characterizing the drug's genotoxicity, its cardiovascular safety, and drug-drug interactions.

The characterization process is all about getting to understand and moulding the biological effect the chemical scaffold has on the biological target. As testified by the VP Medicinal Chemistry;

> "I want to understand which parts of the molecular structure are relevant for specific parameters. Suppose you have 3 positions in your chemical scaffold you can modify. If I change this position I loose all of my activity, so I don't touch it in future optimization cycles. In a second position, you know, depending on what I put there, I can fine-tune my activity, which is another parameter. The third aspect of your molecule says whatever I put there it doesn't make a difference in activity but it means this is the position I can use to optimize all the other parameters".

The experimentation strategy used to characterize the three critical areas has evolved over the years. Before, the focus of Discovery Research was on characterizing potency and selectivity, leaving the two other areas for later during pre-clinical research. As mentioned by the Head of ADME Research;

[27] See PharmaCo internal document: 'Bioavailability=Complex in vivo disposition processes: Can they be reduced to discrete mechanisms that can be modelled?" referenced in Appendix C under number (12)

> "In the older days discovery scientists were looking for increased potency in animals, and in a way they were also looking for good ADME properties but somehow they didn't realize it. They were not separating these factors out though. Whereas now, we isolate them [drug-likeness indicators] out to be able to understand each individual process. Then we put them back together again at the end. And that's where you look for overall developability of candidate compounds".

Nowadays, discovery scientists try to separate out the factors of all three critical areas that have an effect on biological activity and drug-likeness as early as possible in the process. It's taking a macro-parameter and splitting it up into its constituent components as testified by the Head of ADME Research;

> "It has to do with our understanding of science. Before we looked at a macro parameter; what happens if you shoot it into an animal? Since we didn't know what to optimize we made something similar and we picked the best without really understanding the different components of that macro-parameter. Now, we want to understand these different components so activity in an animal means it needs to get into a cell, secondly it needs to get to the right place in the body and stay there long enough, this means drug permeability. We now get to a level where we play with single parameters to use in our design. It's taking a macro-parameter and splitting it up into its constituent components. This is the difference between old and new paradigms".

However, this leads to schools of thought on how to build the mental model of the solution. Unravelling macro-parameters by looking at their constituent components at very early stages of the discovery process implies that measurement errors are made and questions can be asked about the prediction capacity of these early tests carried out on surrogate markers emulating the true test. As an example, how does Lipinski's Rule of Five (Ro5), a classification method used in silico to assess the drug's absorption potential, correlate with the real test in a human where the Fraction of Dose absorbed in the Portal vein (FDP) or between the stomach and the liver, is measured? Summarizing the schools of thought, the VP Medicinal Chemistry states;

> "Some people say that you should forget all this ADME stuff [early in the discovery process] since the animal model tells you more than enough that the compound is active in this animal, and then the next step is to get it into human, because all the ADME testing you do is not predictive anyway. This is one way of looking at it. The other way is saying we want to understand this compound better so that we can extrapolate the individual data on this enzyme. We say you should be able to modify all these parameters and see what comes out then".

Front-loading, then, is an experimentation strategy opting for a multi-factorial optimization process where the three critical areas (P, B, T) get characterized from the beginning at HTS. By looking at surrogate markers emulating the true test in humans, a picture is painted of the effect the chemical scaffold has on performance variables in the three areas. Whether a picture is predictive for the picture later on in the process depends on the *surrogate marker chain tightness*, essentially the correlation between the different markers of the chain. To make predictions, scientists need to find out what the chains are, which indicators are surrogates of which ones? The higher the correlation between surrogate markers used at subsequent phases the tighter the chain.

As an example, the surrogate marker chain for bio-availability starts in HTS where Lipinski's Rule of 5 is used to get a first idea about this critical compound's

property. Then, in the next phase, H2L, another indicator called PAMPA is used, and further down the process, during LO, FDP in animals is used. Since all tests have measurement errors, a correlation between the test results can be calculated. Also, measurement methods become more accurate as one proceeds in the discovery process. More aligned test results lead to higher, tighter correlations between results captured in HTS, H2L, LO, and ultimately in humans. The higher the correlation between the results in the various phases leading to human testing the better will be the predictions in humans.

The tightest chain realizable for the moment is the one describing bio-availability B. The problem with the P and T chains, the two other ones, is that they are still in their infancy, and PharmaCo's scientists are still trying to understand the science behind them.

Summarizing, pharmaceutical discovery is a noisy search process where proxies or surrogates of the 'real' parameter of interest are used at various moments to make predictions about the potential of the candidate compound in the real world of human testing. A final fundamental question remaining for this research then is; what is the minimum tightness required for front-loading to be meaningful as an experimentation strategy?

3.3.5 Discussion

To discuss the results of this confirmatory case study I will follow a literal replication logic (Yin, 1989) predicting the same results as in the previous exploratory study. PharmaCo's discovery process is documented as a process flow chart[28] that the Discovery Teams follow. Visual mapping (Langley, 1999) of this discovery and pre-clinical research process discussed led me to identify two distinct phases A and B of complexity-handling behaviour (Figure 3-4).

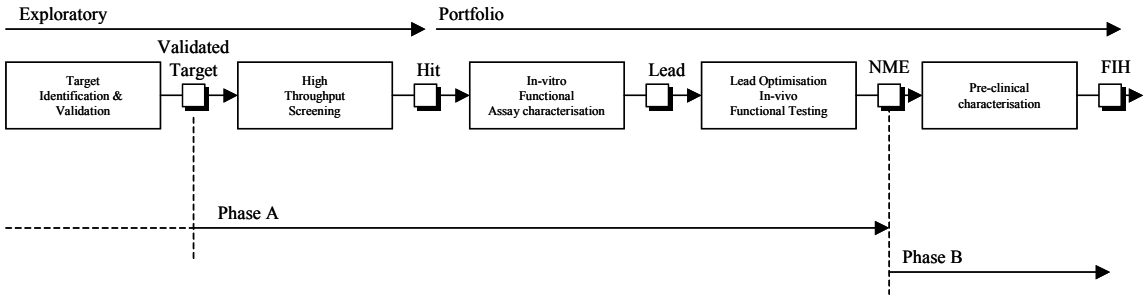


Figure 3-4: Visual mapping of the discovery and pre-clinical research process

Before Phase A the clinical target is searched for and validated. The search for the candidate chemical structure solution has not started yet. Confirmatory case data about the process followed by the innovation teams in this early 'Exploratory' phase of

[28] See footnote 22 above

research are insufficient to formulate conclusions about the experimental design used, and whether they fit my proposed prescriptive framework.

Mapping the descriptors of the three complexity-handling modes of my proposed model to the two identified phases A and B (see Appendix C), indicates the best match of empirical indicators for Phase A with the 'Concept Selection' complexity-handling mode descriptors. Also, the data fit with the proposed reason why this mode is followed by the team. During phase A, innovation teams experience high ambiguity-based complexity as previously defined in Project 1[29]. Solution properties are unknown to deliver proof of concept, and several options need to be characterized, selected, and optimized until target criteria are met. Once target criteria are met, a candidate compound is promoted to the status of NME, and transferred to Phase B for further optimization. Essentially, a concept screening approach is followed until the winning concept is selected against an objective, defined up-front. It is a structured solution search process in a multi-factorial fitness landscape.

During Phase B, in pre-clinical research, innovation teams face ambiguity-based complexity as previously defined[30]. Mapping the Discovery case data to the 'Concept Characterisation' Coordination Approach mode descriptors (see Appendix C) acknowledges that the selected NME's that have shown proof of concept in 'Concept Selection' are now characterised to a level where it is possible to focus their further characterisation and development. Based on 'First Glance' and 'Compound Transfer' results DE takes the decision to accept the candidate NME for further development. However, these case data provide no further evidence to corroborate the Target Setting, Experimentation Approach, or Learning Approach descriptors for this mode since no interviews were conducted within this part of PharmaCo's research organisation.

3.3.6 Conclusions

From the previous, I conclude first that the experimentation approach used by the innovation team in Phase A has the best fit with the 'Concept Selection' complexity-handling mode defined in my exploratory research (See Figure 3-5 below). Phase B could fit the 'Concept Characterisation' mode. However, since no extensive interviews were conducted in this part of the organisation known as 'DE', I cannot validate the set of descriptors mentioned for this complexity-handling mode. Nor can I do this for the 'Concept Application' complexity-handling mode in my proposed Project 1 model.

In addition, this case study confirms my previous propositions that the 'Concept Selection' mode will be chosen by the innovation team when experiencing high ambiguity-based complexity, and that 'Concept Characterisation' will be chosen when facing ambiguity-based complexity. However, case study data do not allow confirming propositions made about the 'Concept Application' complexity-handling mode, since no empirical evidence was found for this mode in the Portfolio part of the Discovery process.

[29] See Table 2-3: Experienced complexity typology
[30] See footnote 22 above

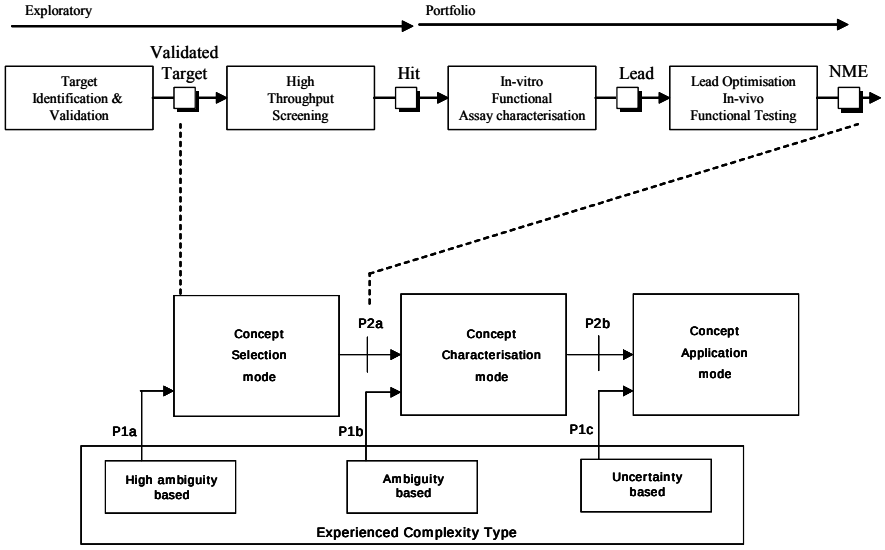


Figure 3-5 Pharmaceutical Discovery process mapped to proposed complexity-handling model

As described in Table 3-1 below, based on these confirmatory case data, a number of additions (indicated in italics) could be made to the 'Concept Selection' descriptor sets documenting the different approaches.

Also, the Discovery case data support my exploratory case study results that experimentation approaches vary dynamically over the course of the project, even if judged on outcome it could be classified uniquely as a radical innovation project. By taking the perspective of the mental model of the problem-solving situation, the team gradually builds up understanding of the innovative system, the prototype compound, to be designed. The decision to transit to a new complexity-handling mode is driven by the perceived completeness of the team's mental model. As long as ambiguity has not been sufficiently resolved, the team stays in 'Concept Selection' mode. The decision to transit from 'Concept Selection' to 'Concept Characterisation' mode is made by the team as soon as solution critical requirements are met; in this case the acceptance of the NME by DE at 'Compound Transfer'. The latter milestone indicates that high ambiguity is resolved to the level of an emerged mental model of the innovative solution containing all critical variables and their relationships necessary to deliver Proof of Concept.

Therefore, I conclude that these case data support my previous proposition that the decision to transit from 'Concept Selection' to 'Concept Characterisation' complexity-handling mode is made by the innovation team as soon as a mental model has emerged that contains all critical variables and their relationships necessary to deliver Proof of Concept. However, these confirmatory case data do not allow being conclusive about the transition from 'Concept Characterisation' to 'Concept Application' mode.

	Concept Selection mode descriptors
Target Setting Approach	• Target defined up-front as minimal system critical requirements to pass
Experimentation Approach	• ***Define uncertainty areas to characterise critical variables and to anticipate problems to be solved (added)*** • Run parallel experiments to characterise critical variables affecting response for different candidate system solutions • System solution selection • Show Proof of Concept by spelling out assumptions about the set of relevant variables and their functional relationships
Learning Approach	• External explicit: Mainly at project start learning from published science. Later ad-hoc for problem solving. • External tacit: On-going knowledge transfer by interaction between teams and external technology suppliers. • ***Internal tacit:*** ***Use of pockets of previous knowledge (added)***
Coordination Approach	• Define milestone targets that are reached if results can be shown • Estimate work package effort/ timeline based on first experiments or expert knowledge • Focus on experiments capable of selecting as quickly as possible solutions that meet all system critical requirements • Through close monitoring of progress: Eliminate as quickly as possible candidate system solutions not meeting one of the system critical requirements

Table 3-1: Revised 'Concept Selection' descriptors

Finally and most importantly, this case study allows for exploring further the dynamics involved in the build-up of the mental model. More specifically, internal PharmaCo presentations[31] reveal a debate around the set of critical variables and their relationships needed to deliver Proof of Concept of a NME. Thus, an internal PharmaCo study[32] revealed that in the 1995-2000 period about 50 compounds failed in pre-clinical and clinical development programs due to poor 'drug-likeness' of NME's, meaning they showed too low performance on PK/PD, toxicology, or could not be suitably packed in a drug delivery vehicle. The same study reveals that, in retrospect, the present discovery process, taking into consideration both biological activity and pharmacokinetic property classes, would have 'caught' about 20 compounds being unrightfully promoted to NME status. Acknowledging that in discovery both biological activity and pharmacokinetic property classes should be used in a multi-factorial optimization, the experimental design question remains when each property class should be taken into consideration to be as effective and efficient as possible.

[31] Hoflack, J. (2002)' The drug discovery challenge: Better targets, better compounds, better processes', internal PharmaCo publication as referenced in Appendix C under number (13).
[32] Mackie, C. (2002) 'Pre-clinical expertise in drug discovery: Changing the paradigm', internal PharmaCo publication as referenced in Appendix C under number (14).

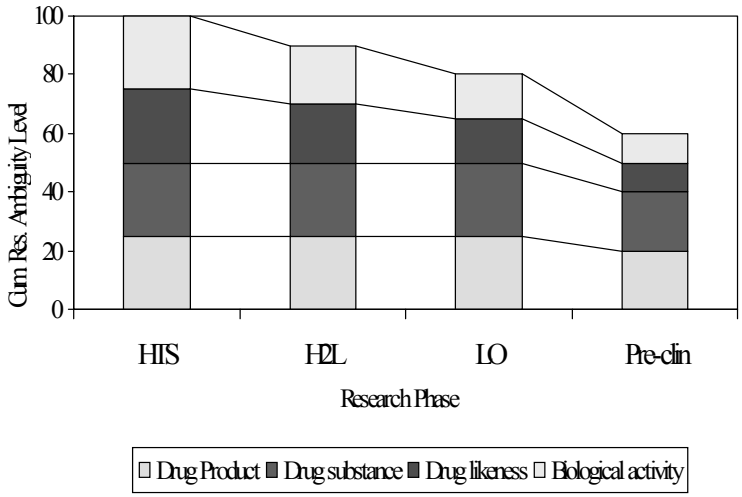


Figure 3-6: Cumulative Residual Ambiguity Level as a function of Research Phase

In terms of my conceptual framework, this debate turns (1) around the question of the desired level of residual ambiguity required to transit from 'Concept Selection' to 'Concept Characterisation', and (2) around the speed with which ambiguity gets resolved. The experimental design choices made by PharmaCo are conceptually depicted in Figure 3-6.

If we define residual ambiguity as the number of solution properties that remain to be characterized in the R&D process, then, at the start of HTS 100% still needs to be done while only the biological target is known. As specified above, of the property classes relevant for the design of a drug (see **Figure 3-2**), PharmaCo considers both biological activity and bio-availability –also called drug-likeness- to promote a candidate compound to NME status. Residual ambiguity at the end of Discovery research is composed of all properties pertaining to drug product and drug substance design, and all biological activity and drug-likeness properties which cannot be tested in animal models but, instead, need human models to be characterized. As to the residual ambiguity reduction speed, it can be verified on Figure 3-6 that PharmaCo chooses to focus already in HTS on biological activity and drug-likeness, which positions it as a discovery organisation that believes in a 'Front-loaded' experimentation policy. From HTS on pharmacokinetic or drug-likeness properties are used in a multi-factorial optimization cycle of the prototype compound. H2L further characterizes potency and drug-likeness of the compound classes and identifies potential problems that will need to be dealt with during LO, where the final NME gets selected and optimized for passing minimal proof of concept requirements on potency, selectivity, drug-likeness, and toxicity. Drug product and drug substance property classes are only characterised from pre-clinical on. The latter is common industry practice, but the question whether pharmacokinetic or drug-likeness properties should already be taken into account during the in-silico part of HTS –as PharmaCo does- remains open for discussion.

3.4 EXPLORING 'CONCEPT SELECTION' EXPERIMENTAL DESIGN

Both internal policy documents[33] and practitioner literature (Oprea, 2002; DeWitte, 2002) develop a scientific viewpoint on a number of experimental designs available to run pharmaceutical Discovery research. They can be summarized (Figure 3-7) to the extent *parallelism* and the level of *front-loading* is used in the Discovery process, the latter indicating the shape of the Discovery funnel.

The level of parallelism will be measured by the number of chemical classes taken into account for each drug target during the various stages of the Discovery process, indicating the shape of the Discovery funnel. Typically in PharmaCo one to five classes showing activity against a target are selected in HTS screening, and one to three classes are further analysed in H2L and carried through to LO. A narrow funnel, then would be characterized by the couple (HTS, H2L/LO) = (1, 1). Conversely, a broad funnel would be characterized by the couple (HTS, H2L/LO) = (5, 3). The main cost driver is the number of classes carried through to LO[34].

I will measure the level of front-loading by the number of variables –P and/or B- considered at differing stages during Discovery; the most front-loaded strategy uses P and B from HTS on, the least front-loaded strategy only uses P during the entire process, disregarding B.

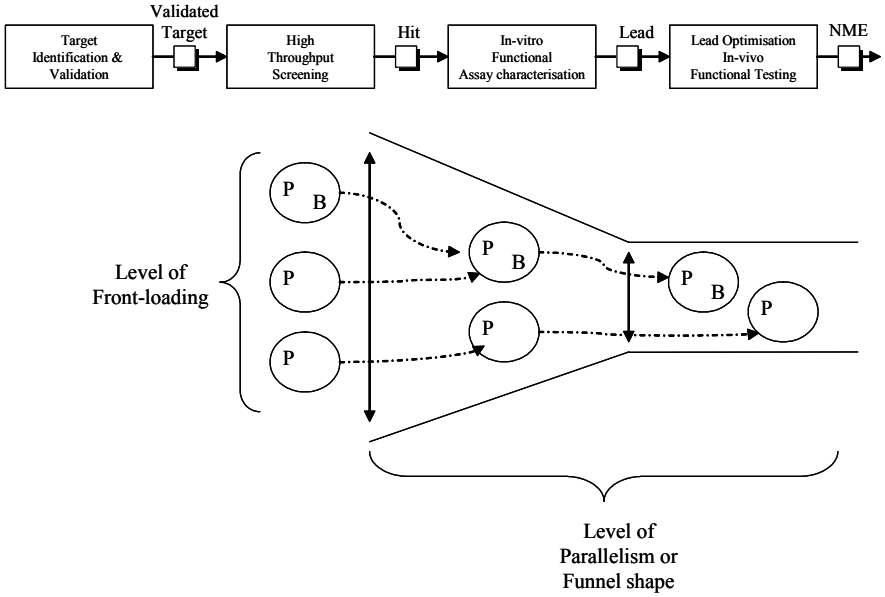


Figure 3-7 Alternative Discovery experimentation strategies considered

The three front-loaded experimental strategies for discovery research considered for comparative analysis are summarized in Figure 3-8 below. The arrows indicate

[33] See footnotes 31 and 32.
[34] See reference (15) in Appendix C.

which property classes are used in the prototype compound optimisation cycles to converge to a NME. A vertical arrow indicates that only biological activity (P) or potency and selectivity is optimized, an arrow along the diagonal indicates that both biological activity and drug-likeness or bioavailability properties (B) are used in multi-factorial optimization cycles.

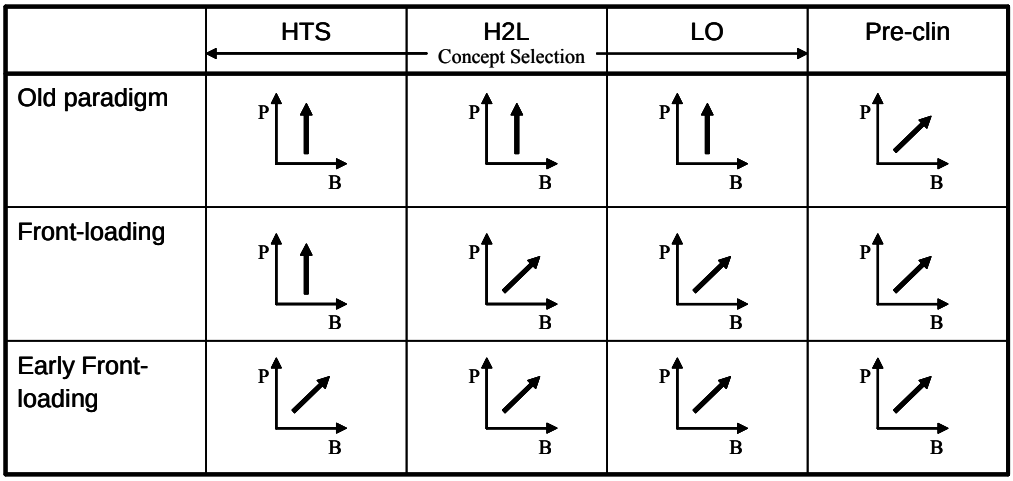


Figure 3-8: Alternative Front-loaded strategies for 'Concept Selection' in a pharmaceutical context

In the *Old paradigm*, applied in the industry some time ago, discovery research was only concerned about biological activity or potency and selectivity of a candidate compound. Drug-likeness properties were only taken into account from pre-clinical research on. As testified by the Head of ADME Research:

> "It was just unfortunate and the way of thinking at that time; everything was first done on pharmacology, then we used to do the check-box approach looking for solubility, permeability, absorption, toxicity, can we formulate this, what's the half-life? And then we would hand it on to the next phase and just pray. Now we look at these other characteristics at the same time as the pharmacology so that we can get a better drug. It's not just the pharmacology driving it".

Figure 3-9(a) graphically depicts this optimization process. The jagged line represents the sequence and number of adaptations or re-synthesis steps to the prototype compound to, first, reach good biological activity and, second, obtain satisfactory drug-likeness properties. This two-step optimisation process is a rather complex undertaking that is likely to require changes in those molecular determinants that are responsible for binding affinity and specificity and can lead to significant drops in potency, which is a time and effort consuming process (Oprea, 2002: 54). This strategy was not entirely successful since within the industry only one in ten candidate drugs succeeded through clinical trials to reach the market (Drews, 1998). Thus, in terms of my conceptual framework, the main reason for this can be found in the too high level of residual ambiguity at the end of discovery research, leaving pharmacokinetic properties uncharacterized. Also, while this experimental design only takes into consideration those compounds with the highest potency for further pharmacokinetic characterisation, compounds with potentially good overall performance profiles are likely to be discarded

due to modest performance on potency or selectivity measures (Oprea, 2002; DeWitte, 2002). This represents a significant opportunity cost.

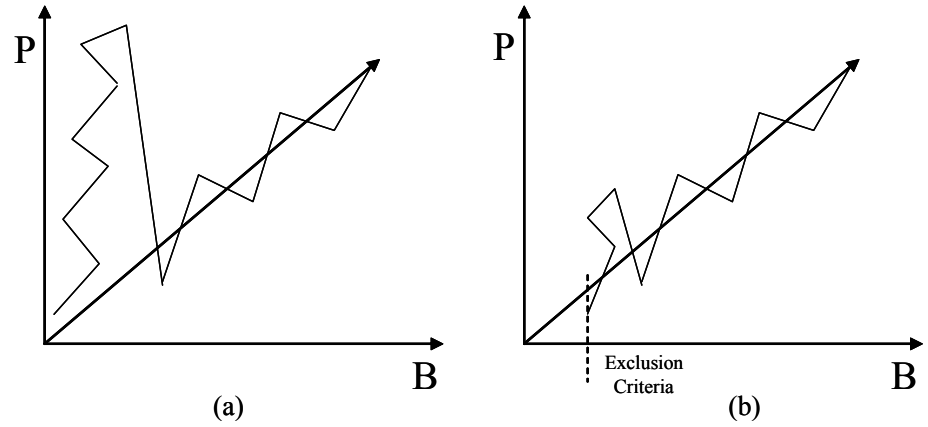


Figure 3-9: Old (a) versus New (b) paradigm experimental design

In the *Front-loaded paradigm*, Figure 3-9(b) shows that both biological activity and pharmacokinetic properties are optimised simultaneously from the Hit-to-Lead (H2L) stage on. Here, confirmed hits series from HTS are prioritized and analyzed for problems. Progressive increments in binding affinity and pharmacokinetics are obtained by addressing the appropriate molecular determinants that define the desired compound characteristics. A second component of this experimental strategy deals with exclusion and promotion criteria used along the discovery research process. It is argued (DeWitte, 2002) that prior to entry into H2L, stringent pharmacokinetic exclusion criteria should be applied to all candidate compounds, with no exception. This would prevent insoluble compounds being synthesized and eliminated after months of useless work. Second, one should be more generous in promoting selected candidate compounds through the H2L and LO process. This argues against introducing too heavy attrition screens along the way to a NME and means more candidates will reach LO status. Although this is an expensive recommendation, it does prevent candidate eliminations due to tests that mimic reality with statistical inaccuracy, the latter compounded to the number of tests conducted. The idea, then, is to build up knowledge through to in-vivo of *a series* of candidates instead of going for a fast attrition based on in-silico or in-vitro tests with low prediction power, and running LO with a small subset of 'winning' candidate series of compounds. Overall, it is argued that this experimental design is a first step into front-loading that selects the best *overall* candidate and that it could be a less time-consuming and resource-demanding strategy than the old paradigm.

Finally, *Early Front-loading* takes the Front-loaded paradigm even further including pharmacokinetic properties during High Throughput Screening (HTS). This step now combines in-vitro with in-silico data on drug likeness of compound classes and selects the most promising classes based on the fullest multi-factorial picture available. Like in the *Front-loaded paradigm*, H2L identifies problems upfront with the series that need to be taken care of. The real multi-factorial compound optimization cycle only starts from LO.

3.5 MEASURING THE QUALITY OF 'CONCEPT SELECTION' EXPERIMENTATION STRATEGIES

The three experimentation strategies considered could be compared on efficiency or effectiveness. Discovery management is most interested in improving the predictive quality of the decision-making process enabled by an experimentation strategy, as testified by PharmaCo's VP Medicinal Chemistry;

> "Actually, if you would take this question higher up, people will tell you cycle time is important. Where we know seriously it is not, it is the quality of the compound and the real thing is to get compounds which make it through development. If a scientist could give me high quality compound but it would take him two more years, I would be extremely happy."

Also, in terms of my model I'm interested in determining the level of residual ambiguity at the end of 'Concept Selection' that maximizes the quality of subsequent decision-making. Even when considering the speed with which ambiguity gets reduced, I am not interested in the efficiency improvement potential of the experimentation process. Instead, I want to focus on the contribution ambiguity reduction speed makes to the quality of the resulting mental model of the solution at the end of 'Concept Selection'. In the remainder of this and the following paragraph I apply recent Bayesian thinking for modelling clinical diagnosis and decision-making (Parmigiani, 2002) to develop a framework to quantitatively evaluate the predictive performance of experimentation strategies used in 'Concept Selection'.

Effectiveness is intuitively defined as the quality of the promotion process of compounds to NME status. In other words; if a discovery experimentation strategy promotes candidates to this status and this gets confirmed by subsequent pre-clinical and clinical development, the strategy is said to be of high quality. In terms of my proposed model I probe into the quality of the decision to change from 'Concept Selection' to 'Concept Characterisation'.

	C^+	C^-	Total
H^+	$p(H^+,C^+)$	$p(H^+,C^-)$	$p(H^+)$
H^-	$p(H^-,C^+)$	$p(H^-,C^-)$	$p(H^-)$
Total	$p(C^+)$	$p(C^-)$	1

Table 3-2: Two-way table of probabilities of experimentation strategy results according to the outcome of 'Concept Selection' and actual subsequent testing results

Using more formal notation, I distinguish between H^+ and H^- being numbers of compounds declared respectively as active or inactive by the experimentation strategy during 'Concept Selection'. The experimentation strategy delivers an output, being a set

of candidate chemical structures H_i^+ that tested positive throughout the optimization process. During this process, a –virtually infinite- number of candidate structures H_j^- were eliminated from further optimization. Then, active compounds H_i^+ declared NME's by the discovery experimentation strategy, will subsequently be confirmed to be positive or negative in pre-clinical and clinical testing. I distinguish between C^+ and C^- as the number of compounds or NME's respectively passing these latter tests or not. Combining these two into one view (Table 3-2), leads to cross-classification of candidate compounds as they are classified by the experimentation strategy and by subsequent testing. Cells indicate probabilities of occurrence.

To test for experimentation strategy *effectiveness*, I analyse the quality or the accuracy of the NME decision made by the experimentation strategy. I distinguish between the sensitivity and specificity of an experimentation approach and its subsequent NME decision-making. Quantitatively, sensitivity is defined as the *true positive rate* β, or the probability of finding an effective compound as being active against a biological target. Defined in the context of Table 3-2 above it is the joint probability $p(H^+, C^+)$ divided by the marginal probability $p(C^+)$, the latter being the sum of true positives and false positives in 'Concept Characterisation' and beyond. Specificity is the *true negative rate* α, or the probability of rightfully classifying a non-active compound as non-active. Alternatively defined it is the joint probability $p(H^-, C^-)$ divided by the marginal probability $p(C^-)$, the latter being the sum of true negatives and false positives in testing beyond 'Concept Selection'. Incidentally, the expressions above are examples of a general relationship between conditional, joint, and marginal probabilities. Therefore, I propose to measure the *quality* of an experimentation strategy using two criteria χ;

$$\chi_1 : \beta = p(H^+ \mid C^+) = \frac{p(H^+, C^+)}{p(C^+)} \qquad \textbf{(3-1)}$$

$$\chi_2 : \alpha = p(H^- \mid C^-) = \frac{p(H^-, C^-)}{p(C^-)} \qquad \textbf{(3-2)}$$

A high-quality experimentation strategy is highly sensitive and selective, hence, has a high value for both β and α.

3.5.1 Measuring predictive performance of 'Concept Selection' experimentation strategies

Both α and β describe accuracy; the two types of correct classifications made by the experimentation strategy. However, to evaluate the predictive performance of an experimentation strategy, a related though different question must be answered: 'What

is the probability that a structure H_i^+ is *really* active, denoted as C_j^+? To answer this question, we must start from the universe of potential compounds H and examine how the experimentation strategy improves the odds of finding an active compound. If we indicate by π the fraction of really active compounds $p(C^+)$ in the universe of potential compounds H, called the prevalence, we can rewrite the cells of Table 3-2 to become Table 3-3 below.

	C^+	C^-	Total
H^+	$\pi\beta$	$(1-\pi)(1-\alpha)$	$\pi\beta+(1-\pi)(1-\alpha)$
H^-	$\pi(1-\beta)$	$(1-\pi)\alpha$	$\pi(1-\beta)+(1-\pi)\alpha$
Total	π	$1-\pi$	1

Table 3-3: Two-way table of probabilities of experimentation strategy results according to the outcome of 'Concept Selection' and actual subsequent testing results using compound prevalence

Now, using Table 3-3 we can define positive and negative predictive values of experimentation strategies in terms of their sensitivity, specificity and prevalence of active compounds in the universe H. This representation of uncertainty about parameters using probabilities is called Bayesian inference. It models the experimentation strategy as a learning process that modifies one's initial probability statement about the prevalence prior to observing the data during experimentation to updated or posterior knowledge incorporating both prior knowledge and the data at hand (Congdon, 2001: 3). The *positive predictive value*, then, denoted $p(C^+ \mid H^+)$ or π^+ is read as the probability that a compound will pass pre-clinical and clinical tests $p(C^+)$, given it has been declared active $p(H^+)$ by the experimentation strategy. Similarly, the probability that a compound will not pass pre-clinical and clinical testing $p(C^-)$, given it has been declared inactive $p(H^-)$ by the experimentation strategy, is called the *negative predictive value*, and denoted $p(C^- \mid H^-)$. Knowing that the fraction of really active compounds, given they were declared inactive by the experimentation strategy is called π^- or $p(C^+ \mid H^-)$, the negative predictive value is denoted as $1-\pi^-$ (Parmigiani, 2002). Hence, I propose to measure *the predictive performance* of an experimentation strategy using two criteria χ;

$$\chi_3 : \pi^+ = p(C^+ \mid H^+) = \frac{p(C^+, H^+)}{p(H^+)} = \frac{\pi\beta}{\pi\beta + (1-\pi)(1-\alpha)} \quad \textbf{(3-3)}$$

$$\chi_4 : \pi^- = p(C^+ \mid H^-) = \frac{p(C^+, H^-)}{p(H^-)} = \frac{\pi(1-\beta)}{\pi(1-\beta) + (1-\pi)\alpha} \quad \textbf{(3-4)}$$

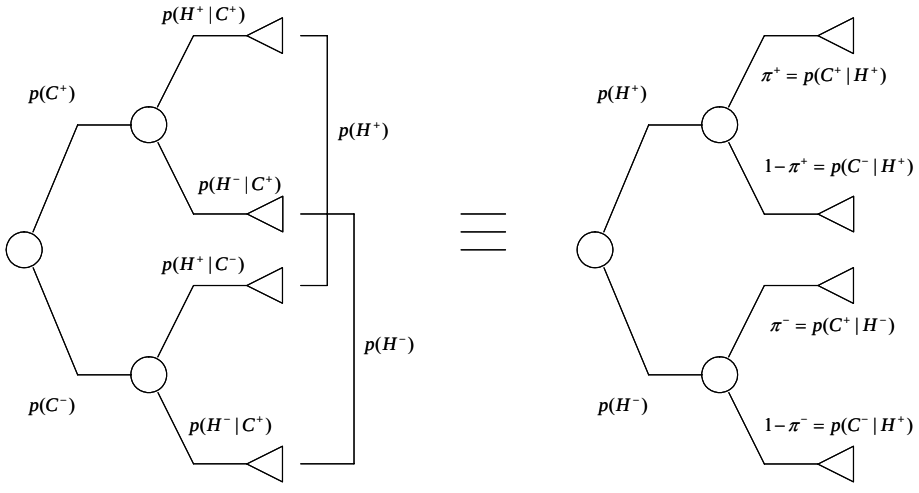
An experimentation strategy featuring high positive and negative predictive performance, then, has a high value for π^+ and a low π^-. The transition from π to π^+ and π^- models the learning about the true status of the universe of potential compounds H. Using the Bayesian logic set out above, it quantifies how inferences about the universe of potential compounds H are updated in the light of new evidence, provided by the experimentation strategy.

3.5.2 Measuring business performance of 'Concept Selection' experimentation strategies

To formally state the problem of an R&D manager acting as a normative system[35] having to select between experimentation strategies, the optimal business decision for a strategy d^* can be modelled as follows (see a.o. Müller, 1999; Parmigiani, 2002)

$$d^* = \arg\max_{d \in D} U(d) \quad \text{where} \quad U(d) = \int u(d,\theta,y)\, p(\theta)\, p_d(y \mid \theta)\, d\theta\, dy \qquad \textbf{(3-5)}$$

U(d) is the expected utility of an experimentation strategy d, an element of the universe of possible experimentation strategies D. The utility function $u(d,\theta,y)$ is in our case specified by solving a decision tree of the outcomes of the various (H, C) combinations. Ordering these combinations in a decision tree, and based on cost assumptions for each experimentation strategy a financial outcome can be calculated and used for comparison to make an optimal decision d^*.



(a) **(b)**

Figure 3-10 Experimentation strategy outcomes Bayesian tree reversal

As a starting point for this calculation, Figure 3-10 (a) gives the decision tree representation of Table 3-2 using conditional probabilities. In this representation, circles

[35] See Chapter 1 for a definition of a normative system

indicate uncertainty or chance, triangles represent an outcome. Summarizing the outcomes for both states of H using tree reversal following Bayes' theorem[36], the resulting conditional probabilities are now expressed in Figure 3-10 (b) as forms of the predictive performance indicators π^+ and π^-. The latter decision tree representation is the basis for calculating the utility or business value of each of the experimentation strategies considered in this study; Old paradigm, Front-loading, and Early Front-loading.

Figure 3-11 gives a decision tree representation of the optimal choice d^* (depicted as a rectangle) to be made between the three experimentation strategies. A number of assumptions have been made to calculate the financial value of each branch in the decision tree. Overall, the depicted tree, identical for each experimentation strategy, represents the two-step decision to take a compound to market. The first – Discovery- step leads to a hypothesized activity H of the candidate compound against the biological target, to be confirmed in subsequent clinical testing as being *really* active, in which case it will be taken to market, or not. To avoid complexity in calculations not being the subject of this thesis, it was assumed that compounds only get promoted to the following status at the end of a phase meaning all costs related to that phase are always incurred. In other words, no compounds get eliminated from the process during a phase. Then, the branch $\langle H^+, C^+ \rangle$ leads to outcome A, which is revenue R after deducting the costs of taking the compound through Discovery research (C_D) and Clinical development (C_{Clin}). In the utility calculation, the latter cost and the revenue R will be held constant while not the object of this study. To respect confidentiality, actual numbers for R and C_{Clin} will be fictitious numbers provided by PharmaCo[37] with the objective to represent reality as good as possible. The Discovery research cost is contingent upon the experimentation strategy followed; PharmaCo case study numbers will be used to calculate $C_{D\,Old,FL,EFL}$. The branch $\langle H^+, C^- \rangle$ does not lead to revenue R although the candidate compound has been promoted to development status after Discovery, so outcome B represents only costs incurred by Discovery and Development. The branch $\langle H^-, C^+ \rangle$ is not observable in practice since, in reality and in principle, a candidate compound declined in Discovery will not be pursued further in Clinical Development. However, to calculate experimentation strategies' utility based upon both positive and negative components of their predictive performances, the outcome C of this branch was modelled as follows; Discovery research costs are incurred since the candidate will only be declared to be inactive against the biological target at the end of Discovery research. Then, Clinical development cost is actually saved, represented by $+C_{Clin}$, since the compound would have been eliminated from the development process. However, this would have resulted in a missed opportunity valued -R. Likewise, the branch $\langle H^-, C^- \rangle$ would not be observable in practice for the same reason as above. Only now no revenue will be the result and only Discovery research costs would have been incurred.

---

[36] $p(C|H)\,p(H) = p(C)\,p(H|C)$

[37] See reference (15) in Appendix C.

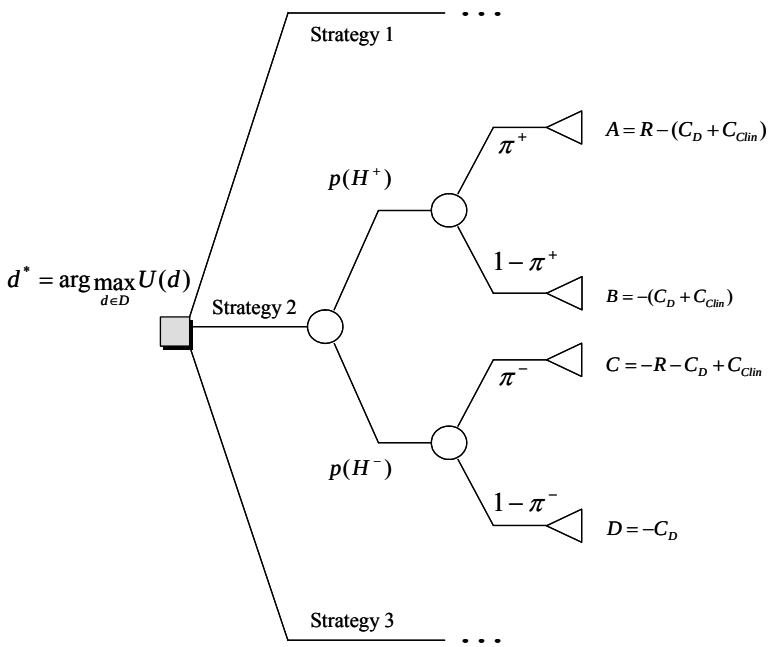


Figure 3-11 Utility calculation and optimal choice decision tree model

To solve this decision tree for each experimentation strategy the payoff values at the outcome nodes need to be propagated backwards from the end-states. Two basic rules are followed to go from the end nodes to the decision point at the beginning; (1) at a chance node the expected value (EV) is obtained by multiplying each value at the end of a branch leading from that node by the probability on that branch and then summing all these products; (2) Finally, at the decision node the EV is just the maximum value of the payoffs found at the ends of the arcs leading from the node concerned. The optimal choice at the decision node is the experimentation strategy option corresponding to the branch that leads to the maximum value (see a.o. Jensen, 2001). Applying this logic to Figure 3-11 leads to the following utility or business value U(d) for an experimentation strategy d:

$$\begin{aligned} U(d) = {} & p(H^+)\left[\pi^+(R-(C_D+C_{Clin}))-(1-\pi^+)(C_D+C_{Clin})\right] \\ & + p(H^-)\left[\pi^-(-R-C_D+C_{Clin})-(1-\pi^-)C_D\right] \end{aligned} \quad \textbf{(3-6)}$$

or, rearranging for a revenue-enhancing and a cost component, the last performance criterion for an experimentation strategy is given by

$$\begin{aligned} \chi_5 : U(d) = {} & \left[\pi^+(R-(C_D+C_{Clin})\,p(H^+)-\pi^-(R+C_D-C_{Clin})\,p(H^-)\right] \\ & -\left[(1-\pi^+)(C_D+C_{Clin})\,p(H^+)+(1-\pi^-)C_D\,p(H^-)\right] \end{aligned} \quad \textbf{(3-7)}$$

To calculate U(d) a statistical experiment needs to be set up to calculate a model $p_d(y|\theta)$, a distribution of observables y conditional on prior distribution $p(\theta)$. In our case, this model is not analytically solvable implying the need for numerical solution strategies (Müller, 1999). This is one of the reasons why in Chapter 4 Monte Carlo simulation (see a.o. Critchfield and Willard, 1986; Müller, 1999) will be used to explore the research questions and build theory on predictive performance of experimentation strategies.

3.6 RESEARCH CONJECTURES

3.6.1 Performance of front-loaded experimentation strategies

In my previous exploratory research project I proposed that during 'Concept Selection' the high ambiguity facing the innovation team gets gradually resolved to a level where the team's mental model of the system solution to be designed contains all variables and their functional relationships relevant to show Proof of Concept. The decision to transit to 'Concept Characterisation' is driven by the perceived completeness of the team's mental model or as soon as solution critical requirements are met. High ambiguity must be resolved to the level that the emerged mental model contains all critical variables and their relationships necessary to deliver Proof of Concept. In this project I called this the *residual ambiguity* level and I want to explore the performance impact of the dynamics of this concept on subsequent experimentation and decision-making during 'Concept Characterisation' and beyond.

Level of front-loading. The following research conjectures related to the transition question between 'Concept Selection' and subsequent 'Concept Characterisation' and 'Concept Application' modes will be explored:

> *Conjecture 2-1: There is an inverse relationship between the level of residual ambiguity at the end of 'Concept Selection' and the positive predictive value of the experimentation strategy followed within this complexity-handling mode.*

In other words, the more you know about a concept before stepping into exhaustive characterisation and application development, the higher the chance for the concept to survive 'Concept Characterisation' and 'Concept Application'. This is in line with my case study findings and with the recent literature on problem-solving conducted in new product development projects in other sectors. An exploratory study conducted by Verganti (1999) in Swedish and Italian companies, operating in the vehicle, helicopter, and white-goods sectors, provides an example. He identified four possible approaches to manage the early phases where problem anticipation and reaction have different balances and are not mutually exclusive nor contradictory, but strongly interact with each other through a mechanism called 'planned flexibility'. This is the capability to build flexibility into the new product development process due to decisions taken early in the project. Planned flexibility implies early identification of specific critical areas of a given project and early planning for reaction measures, enabling more efficient dealing with uncertainty during subsequent development. Thomke and Fujimoto (2000) define front-loaded

problem-solving as 'a strategy that seeks to improve development performance by shifting the identification and solving of design problems to earlier phases of a product development process', and report on findings in the automotive sector, where the methodology currently is being applied to the reengineering and shortening of product development.

Secondly, it is proposed that the earlier ambiguity gets resolved during 'Concept Selection' the higher will be the qualitative and predictive performance of this complexity-handling mode:

> *Conjecture 2-2: The earlier ambiguity gets resolved during 'Concept Selection' the higher the positive and negative predictive value of the experimentation strategy followed within this complexity-handling mode.*

Although this is in line with the thinking developed in Discovery Research at PharmaCo and corroborated by the practitioner literature (Pickering, 2001; Coty, 2002; DeWitte, 2002), no empirical evidence exists in this or other industries. A theoretical rationale could be provided by Information Theory claiming that 'the best searches [in problem-solving] are sensitive enough to return all or most of the desired data', …[and] 'the AND operator provides the greatest selectivity'(Bergeron, 2003: 164). Since the multi-factorial optimization used in a front-loaded experimentation strategy takes into consideration lead compound drug potency, and drug likeness, and toxicity, it should improve its selectivity α and consequently improve the odds of succeeding pre-clinical testing given a positive outcome in Discovery. Since the posterior odds are equal to the prior odds multiplied by the *weight of evidence* or *Bayes factor* (Parmigiani, 2002: 11), the latter will increase if α increases as can be shown if we give Bayes' rule in its multiplicative form:

$$\frac{p(C^+ \mid H^+)}{p(C^- \mid H^+)} = \frac{\pi^+}{1-\pi^+} = \frac{\pi}{1-\pi}\frac{\beta}{1-\alpha} \qquad \textbf{(3-8)}$$

Also, using a similar argument it will reduce the odds of finding positive compounds in pre-clinical and clinical testing given a negative outcome in Discovery since the related Bayes factor will decrease with an increasing α;

$$\frac{p(C^+ \mid H^-)}{p(C^- \mid H^-)} = \frac{\pi^-}{1-\pi^-} = \frac{\pi}{1-\pi}\frac{1-\beta}{\alpha} \qquad \textbf{(3-9)}$$

Therefore, I conjecture a front-loaded experimentation strategy to be more selective than an old paradigm strategy, which only took into consideration drug potency as a factor for optimization during Discovery. Applying multi-factorial optimization earlier possibly will compound the weight of evidence and, in principle, increase the selectivity. Thus, I tentatively conclude that an Early Front-loaded experimentation strategy is even more selective than a strategy following the Front-loaded paradigm provided the theoretical evidence for both conjectures stated above.

Level of parallelism. Previous studies in technology-intensive industries suggest the benefits of broadening the concept testing funnel (Sobek II et al. 1999) or at least propose to optimize the shape of the concept funnel (Dahan and Mendelson, 2001). However, the impact of concept funnel shaping strategies on predictive performance has not been studied before. Also, considering the recent implementation of the front-loading concept at PharmaCo, comparative historic success data are lacking so my case data do not allow me to formulate conjectures potentially guiding the simulation-based research described in the next Chapter.

3.6.2 Performance robustness of front-loaded experimentation strategies

A front-loaded experimentation strategy only makes sense if, at differing moments during the solution discovery process, a multi-factorial picture can be painted of the solution and its effect on performance parameters. The surrogate markers emulating the real tests in humans after the Discovery phase, used to do this need to be predictive in nature. A non-predictive chain would lead to an erroneous picture of the solution that does not get improved along the discovery process. This leads me to formulate the following proposition;

> *Conjecture 2-3: Front-loaded experimentation strategies used during 'Concept Selection' will feature higher positive and negative predictive value than old paradigm strategies provided a minimum tightness level of the predictive surrogate marker chain is realized.*

Summarizing, front-loaded experimentation strategies are conjectured to outperform classical 'Old Paradigm' and 'Front-loaded paradigm' strategies but their improved performance will only be robust for minimal levels of the tightness of the used surrogate marker chain.

3.7 SIMULATION MODELING RATIONALE AND FRAMEWORK

3.7.1 Simulation modelling rationale

I argue that the conjectures stated in the previous section can only be transformed into propositions using computer simulation. The reason for this is given by the fact that a number of variables like H_i^- or α or their derived probabilities like $p(H^- \mid C^+)$ or $p(C^+ \mid H^-)$ amongst others cannot be observed in practice; hence it is impossible to empirically test for them. Also, testing for performance robustness of front-loaded experimentation strategies requires a methodology allowing for sensitivity analysis of the different explaining variables. Finally, a simulation-based research methodology allows me to address in an integrated way the following tasks (Parmigiani, 2002): (1) making inferences about the parameters of complex models, (2) generating artificial cohorts of candidate compounds to run through experimentation approaches, computing summaries of interest, such as expected utilities or cost-effectiveness ratios, (3)

performing sensitivity analysis of results with respect to input parameters and assumptions, (4) facilitating the search for optimal experimentation approaches in high-dimensional spaces.

As before, in the remainder of this paragraph I will use a Bayesian perspective as developed by Parmigiani (2002) in the context of clinical diagnosis and apply it to build a methodological framework for a Monte Carlo simulation-based computational study.

3.7.2 A Bayesian framework for a simulation study

To test the research conjectures, I need to represent prior knowledge or uncertainty about the parameters describing the universe of potential compound and model how this gets improved by the alternative front-loaded experimentation strategies. From a probabilistic inference point of view, given π, the latter translate uncertainty about α and β into uncertainty about predictions π^+ and π^- of pre-clinical and clinical testing outcomes.

Defining a population parameter, take for example β, requires thinking of all possible really active compounds in the universe, and asking what fraction would be found to be active by an experimentation strategy. However, the universe of really active compounds is extremely large and contains compounds that are not even synthesized yet, which means that β is not observable in practice. A sample of compounds taken from the universe could help to determine possible values of β. The connection between a sample and the population can then be determined through probabilistic inference.

Thus, if we draw at random from the universe a really active compound conditioning on a hypothetical value of β for the population, then the probability that compound x_i will be found active ($x_i = 1$) or inactive ($x_i = 0$) is given by;

$$p(x_i \mid \beta) = \beta^{x_i}(1-\beta)^{1-x_i} \qquad x_i = 0,1 \qquad \textbf{(3-10)}$$

In other words, the probability that a randomly drawn compound will be found active by the experimentation strategy equals β. Assuming conditional independence[38] between compounds we can build a probability distribution for the whole sample, starting from the individual observations;

$$p(x_1 \ldots x_n \mid \beta) = p(x_1 \mid \beta) \ldots p(x_n \mid \beta)$$

[38] Mathematically, statistical independence between two variables requires that their joint probability distribution is the product of their marginals. Conditional independence is the same requirement, but applied to distributions that are conditional on some third variable, in our case β (Parmigiani, 2002: 32)

$$= \beta^{x_1+x_2+\ldots+x_n}(1-\beta)^{n-(x_{1+}\ldots+x_n)} \quad \textbf{(3-11)}$$

This summarizes the information provided by the sample about β. Relating this back to Table 3-2 or Table 3-3 the powers in the second formula would equal the number of compound entries used for probability calculations in the cells (C^+, H^-) and (C^+, H^+) respectively, allowing us to calculate the value of this likelihood function indicating probabilities of empirical evidence given the unknown of interest β.

However, having observed the sample, all values of β are still plausible, but some will be more plausible than others. To represent our uncertainty about α and β a probability distribution is needed. To speak of a probability distribution we need to imagine a metapopulation, or a universe of possible populations, each with a different α or β. The next step then is to specify the posterior based on the accumulated knowledge and derive –for the case of β- $p(\beta \mid x_1, x_2, ..., x_i)$ to represent what is known about which values of β are plausible and which are not in the light of the sample. Now, there is still only one universe of compounds and the probability distributions describe variance in the imaginary metapopulation, which represents an important philosophical debate while it is not as conceptually straightforward as it was in a real population to define parameters of interest. One problem is to define the prior probability $p(\beta)$, or the probability distribution of the possible values of β irrespective of sampling results. Since I know nothing about the prior probability distribution of potentially active compounds in the universe of compounds, and I'm interested in making conclusions that depend exclusively on the outcome of experimentation strategies, I treat all imaginary populations on equal footing and assume a uniform prior. Using Bayes' inference rule Parmigiani (2002: 36) now derives the posterior probability density of the parameter β, given the data provided by the samples;

$$p(\beta \mid x_1, x_2, ..., x_i) = \frac{p(\beta).p(x_1, ..., x_n \mid \beta)}{\int_0^1 p(\beta).p(x_1, x_2, ...x_n \mid \beta)d\beta} \quad \textbf{(3-12)}$$

This probability density function updates our knowledge of the experimentation strategy parameter of interest β. The same can be derived for α. I will use these functions as the basis for the simulation. Together with the prior knowledge of the prevalence in the universe of compounds they can be used to determine the comparative predictive value of experimentation strategies.

3.7.3 Monte Carlo simulation study methodology

To run a simulation study a virtual cohort of compounds needs to be generated and used as input to the probability density functions cited above. This is because the analytic calculation of the latter functions will prove to be very difficult considering the composite non-linear character of the virtual compounds I need to generate. More specifically, I chose to model a virtual compound having three fundamental virtual

properties P, B, and T; Potency, Bioavailability, and Toxicity. Scientist interviews confirm that all three are non-linear functions in reality, and for a compound to be declared active, a multi-factorial evaluation function would need to be designed. Therefore, I will replace these calculations of posterior quantities of interest with summaries of simulated values.

The simulation study will generate M virtual cohorts of compounds, each representing one sample $\{X_m\}$ of the universe of compounds. As indicated in Figure 3-12 below, the selector 'Pre-clinical and clinical testing' serves as an emulator for all tests conducted after 'Concept Selection'. It will serve as the '*Golden Standard'* for data comparison.

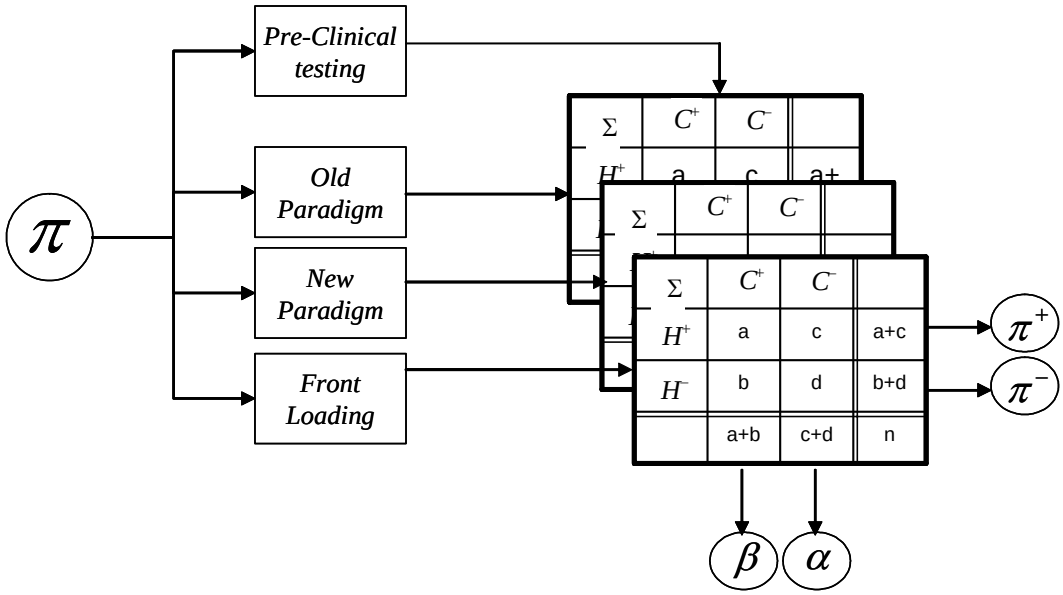


Figure 3-12: Simulation study concept

This way, a fraction π of the n virtual compounds of sample m, the prevalence, is declared to be active by the selector function and counted as an entry in the subset $\{C^+\}$ which consists of (a+b) elements. The (1-π) fraction of the sample counts as entries in the subset $\{C^-\}$ consisting of n-(a+b)=(c+d) elements.

Meanwhile, the same sample serves as input to the three experimentation strategies. Each strategy will make a distinct set of decisions during each of the Discovery phases described above based on specific information about the virtual compounds and will eventually promote compounds to NME status, which classifies them in the subset $\{H^+\}$, or it will eliminate them from the selection and optimization Discovery process, which classifies them in the subset $\{H^-\}$. It will be clear from Figure 3-12 that $\#\{H^+\}$= (a+c) and $\#\{H^-\}$= (b+d), both adding up to n, the sample size.

For each of the M samples the performance criteria $\chi_{1,...,4}$ cited above must be calculated using the summary cross-classified data above. Each distribution of sample values $\pi_m, \alpha_m, \beta_m, \pi_m^+, \pi_m^-$ with $m = 1,....,M$. If θ_m represents a general form of the before mentioned distributions of sample values, they can be summarized using the Monte Carlo average, which is an approximation of the real posterior distributions $p(\theta|x_1,...,x_{nm})$ we're interested in;

$$p(\theta|x_1,...,x_{nm}) \propto \frac{1}{M}\sum_{m=1}^{M} h(\theta_m) \qquad \textbf{(3-13)}$$

Finally, I will use the set of functions $h(\theta)$ to study the propositions set out in the previous paragraph. Although the Monte Carlo-based simulation method allows me to test for the effectiveness and predictive performance of the experimentation strategies, an important limitation does exist; while the simulation concept compares the outcome of each strategy with a gold standard, this does imply that the search space that can be used by the experimentation strategies, although made large in the simulation to emulate reality, is confined to a finite number n. However, as known by every medicinal chemist in reality $n \to \infty$ since it is only limited by the creativity of the scientist synthesizing compounds. Therefore, n will have to be sufficiently large to reasonably emulate reality.

3.8 CONCLUSION

This confirmatory case study successfully replicated part of the findings pertaining to my proposed Project 1 model relating experienced complexity to choices for a specific complexity-handling mode. More specifically, the 'Concept Selection' mode and its transition to the subsequent 'Concept Characterisation' complexity-handling mode could be replicated to the context of Pharmaceutical Discovery. However, the evidence provided by this case provides no evidence to corroborate the 'Concept Characterisation' and 'Concept Application' parts of the model.

In addition, this case study documented alternative experimentation strategies used in Pharmaceutical Discovery, all specific applications of the 'Concept Selection' complexity-handling mode. Various forms of 'front-loaded' experimentation strategies were proposed to be used by Discovery management. All manage the build-up of a mental model of the solution to a level of residual ambiguity the innovation team feels comfortable with to start 'Concept Characterization'. Also, a debate was highlighted pertaining to the tightness of the surrogate marker chain used to conduct experimentation. A minimum tightness level was proposed for front-loaded experimentation strategies to robustly outperform 'Old Paradigm' strategies.

Research conjectures were formulated, linking alternative experimentation strategies to performance. A Bayesian methodology was proposed to evaluate predictive performance of these front-loaded experimentation strategies and their robustness for varying degrees of tightness of the surrogate marker chain.

Finally, a rationale and methodology was developed to test the research propositions using simulation. In Project 3, this will provide the basis for a formal representation of these strategies as adaptive systems, which is needed to build a computer simulation model of various front-loaded pharmaceutical discovery experimentation strategies and to use the latter as an instrument for theory development.

4 Predictive Performance of Front-Loaded Experimentation Strategies in Pharmaceutical Discovery: Theory Building Using a Monte Carlo Simulation Model

4.1 INTRODUCTION

In the previous confirmatory case study analysis the pharmaceutical Discovery process at PharmaCo was mapped to the complexity-handling model proposed in the first exploratory case study analysis. As discussed then, a match could be found between the descriptors of the 'Concept Selection' complexity-handling mode and a number of "Exploratory" and "Portfolio" steps of the discovery research process.

Also, the confirmatory case study allowed me to document various experimentation strategies used during 'Concept Selection' in a pharmaceutical Discovery context. It was argued that they all lead to levels of *residual ambiguity*, aimed for by the innovation team, before the transition to 'Concept Characterisation' is made. Case evidence indicated that variance in experimentation strategies can be explained by the extent front-loading is used. Then, research propositions were formulated, linking residual ambiguity dynamics to predictive performance.

Furthermore, a Bayesian methodology was proposed to evaluate predictive performance of these front-loaded experimentation strategies. And finally, a rationale was developed to test the research propositions using computer simulation.

In this Chapter I will argue that an adaptive system paradigm (Holland, 1992) is amongst others the best choice to emulate the 'Concept Selection' experimentation process. However, to make the paradigm fit (1) my research objective of theory development on predictive performance of experimentation strategies, and (2) the specifics of the pharmaceutical Discovery process context I will use for the simulation, a number of annotations to the paradigm will be proposed.

The annotated adaptive system paradigm, then, will be the basis for designing the simulation model in Visual Basic Excel (VBA), which will be used for "top-down" theory development (Gilbert & Doran, 1994: 39), exploring predictive performance of front-loaded experimentation strategies without and with resource and scientific constraints on the tightness of the predictive marker chain[39]. Therefore, research questions driving this simulation-based theory development study include: *'Does a*

[39] As explained in my Project 2 paper front-loading is an experimentation strategy opting for a multi-factorial optimization process where three critical areas (Potency, Bio-availability, Toxicity) get characterized from the beginning at HTS. By looking at surrogate markers emulating the true test in humans, a picture is painted of the effect the chemical scaffold has on performance variables in the three areas. Whether a picture is predictive for the picture later on in the process depends on the *surrogate marker chain tightness*, essentially the correlation between the different markers of the chain.

front-loaded experimentation strategy increase the odds of getting a positive result at 'Concept Characterization', given a positive 'Concept Selection' outcome?', and, *'How robust is this result for scientific and resource constraints used during 'Concept Selection'?*

My simulation results indicate that Front-loaded strategies in a pharmaceutical Discovery context outperform other strategies on positive predictive performance, irrespective of the tightness of the surrogate marker chain. The number of classes used influences significantly the negative predictive performance of experimentation strategies. This practically means that conducting parallel explorations of concepts during 'Concept Selection' in a pharmaceutical discovery context significantly reduces the probability of missed opportunities in 'Concept Characterisation'. These results will be shown to be robust for varying levels of tightness of the surrogate marker chain.

4.2 THEORETICAL BACKGROUND

4.2.1 Simulating complex innovation processes

In the previous project I gave a rationale for choosing a simulation methodology to answer the performance question of alternative experimentation strategies used during 'Concept Selection'. It was argued that simulation is the preferred methodology to test the research propositions since (1) the dynamics of problem-solving behaviour are not analytically tractable while they have to be represented using discontinuous nonlinear systems, which are generally hard to describe in closed form (Devaney, 1989; Mihm et al. 2003), (2) some of the variables are unobservable in real life hence resisting real-life experiments and empirical research (Masuch and Lapotin, 1989), and finally (3) a simulation-based research methodology allows me to vary underlying assumptions and to virtually search for optimal experimentation approaches in high-dimensional spaces (Parmigiani, 2002).

Using a simulation methodology, a formal model of the discovery research process needs to be designed and a paradigm has to be chosen. To select a simulation paradigm I use a number of criteria. First, case study evidence from my previous project conducted in the pharmaceutical Discovery context suggests that 'Concept Selection' experimentation strategies can be described as a *selection & optimization* process. The chosen paradigm must cater for this. Second, the scientific objective of my computer simulation-based study is to use the model for "top-down" theory development, i.e. 'for elaboration of theoretical hypotheses which may then be [partially] tested by empirical observation'(Gilbert and Doran, 1994: 39). Finally, since 'managing size and complexity of an NPD project has not received widespread attention in the empirical literature' (Mihm et al. 2003) and this is one of the first attempts to model complexity-handling behaviour used in innovation projects, I take a positivist epistemological stance with its core ontological assumptions viewing reality as a fully observable, measurable and objective phenomenon. By limiting the representation of the socio-cognitive problem solving behaviour of the innovation team to an organizational process of selection and optimization, excluding the intricacies of the social

constructivist element of finding a solution to the innovation problem, I hope to limit the complexity of my simulation model.

4.2.2 The garbage can model of organizational decision-making

The garbage can model (Cohen et al. 1972) is a top-down developed theoretical model of organizational decision-making, arguing that organizational choice is ambiguous. 'To understand processes within organizations, one can view a choice opportunity as a garbage can into which various kinds of problems and solutions are dumped by participants as they are generated. The mix of garbage in a single can depends on the mix of cans available, on the labels attached to the alternative cans, on what garbage is currently being produced, and on the speed with which garbage is collected and removed from the scene' (Cohen et al., 1972: 2). Hence, the garbage can model could represent the 'selection' part of my simulation model since its ontological view of the problem-solving process is one of problems meeting solutions in choice opportunities. If a problem meets its solution at the right choice opportunity, a rational outcome is made. Their simulation showed that only specialized access of problems to choice opportunities, combined with unsegmented access of choice opportunities to solutions yields high rates of resolution (Cohen et al. 1972; Masuch and Lapotin, 1989).

Applied to the innovation process context the garbage can model views the innovation team as a machine moving through the research process in an exogenous flow, subject to random simultaneity with the decision to be made at the end of the phases (HTS, H2L, LO) and the solution; whether to pursue or not with the candidate compound. Optimal decisions will be made if specialized scientists make decisions, and if they are allowed to be flexible in their decision-making. The team *responds* at the choice opportunity by demonstrating rational problem-solving behaviour, being unaware of organizational goals or guided by individual preferences, thus ruling out the interpretive view of the team as being a socio-consciousness, a constructor of social reality and manipulator.

Although the garbage can model of decision-making –of *selection*- would be ideally suited for "top-down" positivist theory development, for building a Monte Carlo simulation model and for elaborating theoretical propositions (Gilbert and Doran, 1994: 39), it does miss a critical element of the representation of the complexity experienced by the innovation team; the *optimization* part of their problem-solving behaviour.

4.2.3 Physical symbol systems versus connectionist models

Optimization implies the search for a solution to the innovation problem. There are two approaches to modelling the complexity involved (Cilliers, 1998). *Physical symbol systems* constitute the classical approach to the modelling of complexity. On the other hand, neural networks, parallel distributed processing systems or *connectionist models* are inspired by the working of the human brain, capable of performing complex tasks like pattern recognition followed by intelligent action. Both approaches to modelling complexity receive strong support. In the remainder of this section I will investigate the

applicability of both paradigms to modelling the complexity experienced by the innovation teams involved in a selection & optimization process.

Physical symbol systems model complexity on an abstract semantic level. 'A physical symbol system consists of a set of entities, called symbols, which are physical patterns that can occur as components of another type of entity called an expression (or symbol structure). Thus, a symbol structure is composed of a number of instances (or tokens) of symbols related in some physical way (such as one token being next to another). At any instant of time the system will contain a collection of these symbol structures. Besides these structures, the system also contains a collection of processes that operate on expressions to produce other expressions: processes of creation, modification, reproduction and destruction. A physical symbol system is a machine that produces through time an evolving collection of symbol structures. Such a system exists in a world of objects wider than just these symbolic expressions themselves' (Newell and Simon, 1976: 116). Applied to problem solving behaviour; 'to state a problem is to designate (1) a test for a class of symbol structures (solutions for the problem), and (2) a generator of symbol structures (potential solutions). To solve a problem is to generate a structure, using (2), that satisfies the test of (1)' (Newell and Simon, 1976: 121). Thus, physical symbol systems express intelligence by being involved in a heuristic search by generating and modifying structures until a solution structure is identified. Physical symbol systems simulation is a "bottom-up" methodology requiring empirical data to be captured from experts to feed the rule-based system underlying the selection & optimization processes. Since rule-based representations of reality may encompass aspects of both interpretive –social constructivist- and positivist ontology, 'they have the necessary and sufficient means for general intelligent action' (Newell and Simon, 1976: 116), and can be used for both '[teleological and causal] explanation of relations between social values, beliefs and actions, and the prediction of the effects of values and beliefs upon social actions' (Gilbert and Doran, 1994: 38).

Masuch and Lapotin (1989) successfully used the physical symbol system approach to challenge and extend the initial garbage can findings of Cohen et al. (1972) to make it fit better with empirical reality. Also, it could be used to represent the selection part of the pharmaceutical discovery process. However, once again the challenge to use this model as an emulator of reality would be in the optimization part for which it would be hard to formulate rule sets.

In general, many scientists believe that it would be very difficult to elicit all the rules underlying this intelligent behaviour. Instead, they believe that the best route to artificial intelligence is through connectionism, a 'paradigm in which humans write only simple rules, and complex behaviour such as intelligence emerge from the massively parallel application and interaction of these simple rules' (Mitchell, 2001: 4). Neural networks are one type of connectionist models, mathematically represented as interconnected networks of neurons, capable of pattern recognition, information processing, regulation, prediction and replication by being given simple neural thresholds and maps of strengthening or weakening of connections and capable of learning from indications of success. More specifically, (1) regularity discovery, in which the team learns to respond to interesting input patterns, (2) pattern association, in which the goal is to find a set of connections so that whenever a particular pattern

reappears at the input an associated pattern will appear on the output, and (3) auto-association used for pattern completion (Rumelhart et al. 1986), could be used to model distinctive elements of the innovation team's searching behaviour for an optimal solution. Connectionist models, like physical symbol systems are capable of intelligent behaviour. During training, patterns can be simultaneously presented at the input and the output, weights in the input-output connectivity matrix are modified, which emulates the learning of the model. Then, patterns are presented at the input and the output pattern is measured. This means that knowledge is in the connections, which is the most profound difference with the physical symbol system, where knowledge is stored in the rule-base underlying the model. In connectionist models knowledge is not directly accessible to interpretation by some separate processor, but it is built into the processor itself and directly determines the course of processing. It is acquired through tuning of connections as these are used in processing, rather than formulated and stored as declarative facts in the rule-base of a physical symbol system (Rumelhart et al. 1986).

Although connectionist models are capable of sophisticated pattern recognition and learning, again they do fall short representing a decisive element of the experienced complexity of the innovation team's problem-solving behaviour; the teleological aspect of moulding the candidate solution to fit the design objective.

4.2.4 Evolutionary computation

Evolutionary computation is another example of a massively parallel paradigm with minimal rules, capable of intelligent behaviour. These rules are 'typically "natural selection" with variation due to crossover and /or mutation; the hoped-for emergent behaviour is the design of high-quality solutions in the face of a changing environment' (Mitchell, 2001). Inspired by biological evolution, a massively parallel adaptive search for genetic sequences delivering a highly fit organism can also be used as a method for designing innovative solutions to complex problems. A typical example of an adaptive system, emulating biological evolution while lending itself to computer implementation, is a Genetic Algorithm (GA). GA's 'are search algorithms based on the mechanics of natural selection and natural genetics. They combine survival of the fittest among string structures with a structured yet randomized information exchange to form a search algorithm with some of the innovative flair of human search. In every generation, a new set of artificial creatures (strings) is created using bits and pieces of the fittest of the old; an occasional new part is tried for good measure. 'While randomized, genetic algorithms are no simple random walk, they efficiently exploit historical information to speculate on new search points with expected improved performance' (Goldberg, 1989: 1). I argue that an adaptive system is the best artificial reconstruction of the mental modelling process during 'Concept Selection' research since it mimics best the blind "selection & optimization" search of discovery scientists for a NME in a vast, discontinuous multi-factorial solution space.

Blind search as a problem-solving strategy in radical innovation contexts is not a pejorative term but should be seen in contrast to enumerative search. The latter consists in enumerating different solutions to a problem, evaluating the benefit of each solution, and selecting the best performing one for execution. Of course, considering the intricate

hard to unravel nature of chemical structures and their supposed effect on a biological target, this method becomes worthless because it would take too long to find a solution. Even with computer help, this would be a formidable task. Mitchell provides an example: 'Suppose you want to search for a protein –a sequence of amino acids- that folds up to a particular three-dimensional shape so it can be used to fight a specific virus. The search space is the collection of all possible protein sequences –an infinite set of possibilities. To constrain it, let us restrict the search to all possible sequences of length 100 or less –still a huge search space, since there are 20 possible amino acids at each position in the sequence. Instead, the more efficient blind searching in these kinds of large solution spaces where information on solutions is not readily stored for evaluation, involves candidate solutions being created as the search process proceeds. This adaptive search method '…(1) initially generates a set of candidate solutions, (2) evaluates the candidate solutions according to some fitness criteria, (3) decides on the basis of this evaluation which candidates will be kept and which will be discarded, and (4) produces further variants by using some kind of operators on the surviving candidates (Mitchell, 2001: 7).

This is why I will use adaptive systems as a formal model for the artificial reconstruction of the experimentation process during the 'Concept Selection' complexity-handling mode, and operationalise it later using a simulation model, tailored to the specific environment of pharmaceutical Discovery research. However, to be compliant with my research objective investigating predictive performance of experimentation strategies used during this complexity-handling mode, a number of modifications will have to be made to the generic adaptive systems model. These will be dealt with in the next paragraph building a formal framework for the computer simulation-based study.

4.3 BUILDING AN ADAPTIVE SYSTEM SIMULATION FRAMEWORK

To build a formal framework for the discussion above I follow Holland's (1992) notation specifying adaptive systems by the set of objects (η, Ω, I, τ). I distinguish for my research case between:

E; the environment, or biological target, against which, a drug must be developed. $E \in \varepsilon$, the total set of biological targets,

h_{ij}^{+}; a hypothesized solution to the biological target; which is a candidate compound structure with chemical properties. $h_{ij}^{+} \in H_i$, the latter being defined as a chemical class, and $H_i \subset \eta(t)$, the total set of chemical structures attainable at moment t in time. The initial set $\eta(0) = \eta$, represents the universe of attainable compounds,

τ; the adaptive plan determining successive structural modifications to the hypothesized solution h_{ij}^{+} in response to the target, $\tau \in T$, the set of feasible plans,

I; the total range of signals receivable by the adaptive system on the h_{ij}^{+} properties, measured across the discovery research process. The information I(t) received by the system at time t will be constrained to the subset $I_{\eta(t)} \subset I$, where $I_{\eta(t)} = \{\delta_1(t), \delta_2(t), ...\delta_i(t)\}$ is the set of values δ received at t,

Ω; the set of operators used by the adaptive system to modify the candidate compound structure through the discovery research process. $\Omega = \{\omega_s, \omega_M\}$; Selection and maximization are the operators used,

μ; a measure of performance of the different hypothesized solutions in the environment.

However, to make the formal model fit my research context and objective I need to specify four annotations. The theoretical model underlying the simulation software system will then be the *annotated adaptive system* model.

4.3.1 The domain of action

As a starting point for the formalism, I consider the domain of action for the adaptive plan τ, the universe of potential candidate compound structures η. Each structure h_{ij} exhibits a set of properties, given by nature; its biological activity or potency and selectivity (P), its bio-availability (B), and its toxicity (T). Each property will be represented by a number between 1 and 9. For (P) and (B), a high number indicates high performance. For (T) a low number indicates high performance. Clearly, the latter is a virtual representation of compound properties. The connection to reality is in the exponential distribution of the P, B, T values in the domain of action η.

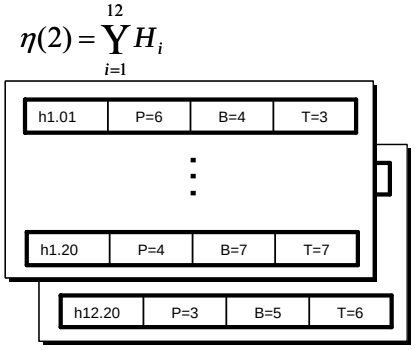


Figure 4-1: Example of a domain of action

Figure 4-1 above gives an example of this representation where the domain of action of the adaptive plan at t=2 is represented by twelve chemical classes H_i, each consisting of

20 candidate compounds h_{ij}, each featuring three properties P, B, and T. Properties can be summarized at chemical class level.

This leads me to the *first annotation* to the formal adaptive system specification I need to formulate to make it fit with my particular application. Holland (1992: 21) notes that 'the set η will usually be potential rather than actual. That is, elements [in my case h_{ij}] become available to the plan only by successive modification (e.g., by rearrangements of components or construction of primitive elements), rather than by selection from an extant set'. My scientific objective is to develop theory on predictive performance of experimentation strategies –represented as adaptive systems- using Bayesian inference. Therefore, as discussed in my previous project paper, I need a reference or Golden Standard, to compare the outcome of the experimentation strategies with "correct classifications" made by nature, or subsequently during 'Concept Characterisation'. The Golden Standard is provided by starting from an extant set and by classifying all candidate compounds in it as "really active" c^+ or "really inactive" c^-. By cross-classifying with the candidate compounds declared positive h^+ or negative h^- by the experimentation strategy, a table can be constructed as shown in the previous project and Bayesian conditional probabilities can be calculated. If I would allow the adaptive strategy to generate new candidate compounds –obviously not known at the outset- during the process, it would not be possible to cross-classify them, which would make further calculations impossible.

4.3.2 The adaptive plan

The adaptive plan τ produces a trajectory of candidate structures h_{ij} through η by making successive selections and optimizations using a set of operators Ω. The adaptive system is limited by what properties $I(t) \subset I$ get measured at a specific moment in time. As an example discussed in the previous project, in the *Old paradigm* bio-availability properties are not used at all for optimization during 'Concept Selection' in Pharmaceutical Discovery. Hence, given I(t) and η(t) the adaptive plan transforms the structures η(t) into η(t+1) by using operators from the set Ω, meaning that the detailed operation of the adaptive plan is described by

$$\tau : I \times \eta \rightarrow \Omega \qquad \textbf{(4-1)}$$

and

$$\tau(I(t), \eta(t)) = \omega_{(t)} \in \Omega \qquad \textbf{(4-2)}$$

I will limit the set Ω to two operators ω_s and ω_M (the latter to be discussed later) applied to candidate structures and chemical classes over the course of the discovery process. I will not use the mutation operator ω_m or crossover operator ω_c of an adaptive system paradigm as operationalised in a Genetic Algorithm.

The latter represents the *second annotation* I want to make to the adaptive system paradigm. To make the paradigm fit with my scientific objective of measuring effectiveness of experimentation strategies I need to exclude the use of mutation and crossover operators. In optimizing the set of candidate structures η(t) into a new set η(t+1), a Genetic Algorithm would reproduce, cross-over and mutate the most performing building blocks of candidate structures, forming a new set of candidate structures showing overall better performance against the objective function (Goldberg, 1989). However, as mentioned above I need to limit the search space to an extant set, randomly generated at the beginning of the search process to test for effectiveness of the different experimentation strategies. Hence, I cannot allow for the random generation of new *upfront unknown* compounds during the optimization process. By limiting the adaptive search engine to selection operators applied to candidate chemical classes and chemical structures, I also limit the possibilities of the paradigm I'm using. More specifically, it will not allow me to test for the *efficiency* of a search strategy to get to the design objective since the number of steps to get to an optimised fit will be fixed by the simulation process and will not be determined by the number of creative leaps made by the crossover operators. However, since the objective of my study is to probe into *effectiveness* of experimentation strategies, this is not considered to be a problem but a limitation of the simulation method.

Second, to fit the adaptive system paradigm to the specific context of the Pharmaceutical Discovery process, a remark should be made about the level of aggregation of information I(t) about η(t) one considers at different steps in the optimization process. From the Discovery case study analysis conducted in Project 2 it became clear that the scientists' unit of analysis during HTS and H2L is the chemical class H_i as opposed to the final chemical compound structure h_{ij}, which is only considered during LO. Then, the information to be used by the selection operators to transit from HTS to H2L and from H2L to LO are the averages $\overline{P}, \overline{B}, \overline{T}$ defined at the level of the chemical class, aggregating property information of its constituent chemical structures.

Summarizing, to steer progress in the annotated adaptive system I will only use the selection operator ω_s from the set Ω, applied at the level of the chemical class in HTS and H2L, and applied at the level of the individual chemical structure at LO, hence;

$$\omega_{s_{HTS,H2L}} : \delta_{\overline{P}_{H_i}} \geq p, \delta_{\overline{B}_{H_i}} \geq b, \delta_{\overline{T}_{H_i}} \leq t \quad \textbf{(4-3)}$$

$$\omega_{s_{LO}} : \delta_{P_{h_{ij}}} \geq p, \delta_{B_{h_{ij}}} \geq b, \delta_{T_{h_{ij}}} \leq t \quad \textbf{(4-4)}$$

The cut-off levels p, b, and t depend on the discovery phase (HTS, H2L,LO) and the experimentation strategy (Old, Front-loaded, Early Front-loaded) used in the simulation.

In the next section I will expand more formally on the uncertainty involved in data capture steering the selection and optimization process of the annotated adaptive

system. This additional notation is needed to emulate the dynamics of the surrogate marker chains discovered in the Pharmaceutical Discovery case study.

4.3.3 Emulating surrogate marker chains

As mentioned before, the optimization process τ takes I(t) as an input to steer selection and optimization progress. The relevant detector values $\delta_{P_{hij}}, \delta_{B_{hij}}, \delta_{T_{hij}}$ measuring performance of the candidate compound chemical structures, and summarizing values $\delta_{\bar{P}_{H_i}}, \delta_{\bar{B}_{H_i}}, \delta_{\bar{T}_{H_i}}$ measuring performance of the chemical classes are used to steer progress through the process. In the adaptive system framework, performance on all three dimensions is formally modelled as

$$\mu_{EB,P,T} : \eta \rightarrow \text{Reals} \qquad \textbf{(4-5)}$$

mapping performance as a real value used in the selection process described above. However, considering the error involved in measuring values, it is more correct to state that the detected value will be drawn from a probability distribution, I will call *U*. So that

$$\mu_{EB,P,T} : \eta \rightarrow U \qquad \textbf{(4-6)}$$

This way, following each discovery phase the candidate structures from η(t-1) get a stochastic score drawing from *U* on properties B, P and T, and is used by the selection operator to generate the new set η(t). U is a complex function consisting of an exponential –the natural distribution of B, P, and T- and a normal –stochastic- component.

Now, to fit the model with the specificity of the pharmaceutical discovery research process, a *third annotation* needs to be formulated. In reality, the three properties are all measured in the different discovery phases. However, the scientific methods used differ for each phase. I chose to emulate this measurement process as depicted in Figure 4-2 below for the case of bio-availability.

The real value b for bio-availability –used in the simulation model as a compound property - is measured in the different phases using a surrogate marker method each having a specific measurement error. As described in Chapter 3, during HTS Lipinski's Rule of Five is used, in H2L it's PAMPA, and in LO it's FDP in animals.

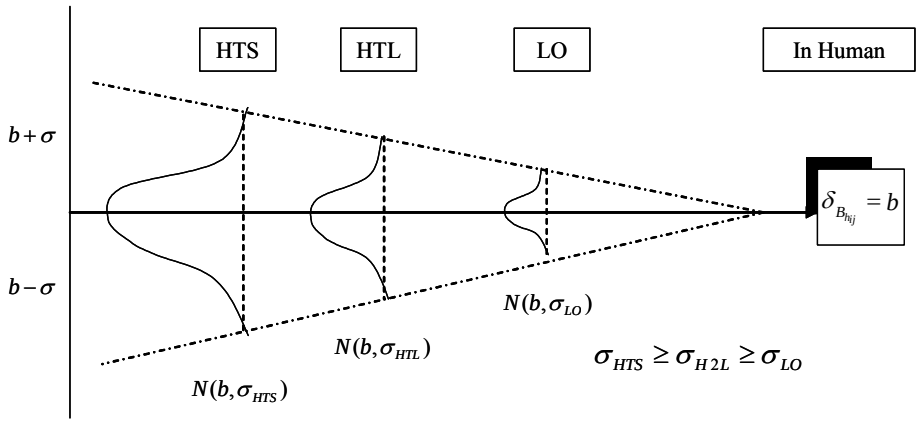


Figure 4-2: Emulating the Bio-availability surrogate marker chain

All these methods are surrogate measures for what the scientist is really interested in; bio-availability in humans. Methods become more accurate as one proceeds in the phases, which explains the funnel shape above. Their predictive power is given by the tightness between the successive measurement methods of the surrogate marker chain. I emulate this tightness by varying the measurement errors on (P), (B), and (T) properties in the different phases in such a way that a wanted correlation level between methods used in each phase is achieved[40]. The higher this correlation, the tighter the chain, and the better its predictive power will be.

4.3.4 The objective function

The adaptive system searches for an optimum in a multi-factorial fitness space; it optimizes the candidate structures for good pharmacological activity, bio-availability and selects for low toxicity. This implies that the simulation model needs to cope with multi-objective optimisation. Formally, an adaptive plan $\tau \in T$ searches for optimum performance;

$$\tau : I \times \eta \rightarrow \prod \qquad \textbf{(4-7)}$$

The adaptive plan receives only information on payoff, hence;

$$I(t) = \mu_{E_{P,B,T}}(\eta(t)) \qquad \textbf{(4-8)}$$

[40] I used the following definition of the correlation coefficient ρ;

$$\rho_{x,y} = \frac{Cov(x,y)}{\sigma_x.\sigma_y} = \frac{\frac{1}{n}\sum_{j=1}^{n}(x_j - \mu_x)(y_j - \mu_y)}{\sigma_x.\sigma_y}$$

To optimize the set η(t) based upon this performance information, to become the set η(t+1) an additional operator ω_M of the set Ω needs to be defined.

This leads me to formulate the *fourth annotation* to the adaptive system model. As mentioned before, progress steering in the adaptive systems model is done through a process of selection and optimization. A Genetic Algorithm implementation of an adaptive system uses the crossover operator to optimize the candidate structures. For the reasons set out above I cannot use this operator. Therefore, I chose to emulate the optimization process using a maximization operator ω_M that searches for the maximum within the search space available to the adaptive system at t. The performance function Π(t) and operator $\omega_{M(t)}$ used by the adaptive system to steer progress depends on the experimentation strategy and the Discovery research phase.

	HTS	H2L	LO
Old paradigm	$\Pi_{Old}(1) = \delta_{\overline{P_{H_i}}}$	$\Pi_{Old}(2) = \delta_{\overline{P_{H_j}}}$	$\Pi_{Old}(3) = \delta_{P_{h_{jk}}}$
	ω_{M_1} : Top m classes	ω_{M_2} : Top n classes	ω_{M_3} : Top p compounds in top n classes
Front-loaded paradigm	$\Pi_{FL}(1) = \delta_{\overline{P_{H_i}}}$	$\Pi_{EFL}(2) = \min(\delta_{\overline{P_{H_i}}}, \delta_{\overline{B_{H_i}}})$	$\Pi_{FL}(3) = \min(\delta_{P_{h_{jk}}}, \delta_{B_{h_{jk}}})$
	ω_{M_1} : Top m classes	ω_{M_2} : Top n classes	ω_{M_3} : Top p compounds in top n classes
Early Front Loading	$\Pi_{EFL}(1) = \min(\delta_{\overline{P_{H_i}}}, \delta_{\overline{B_{H_i}}})$	$\Pi_{EFL}(2) = \min(\delta_{\overline{P_{H_i}}}, \delta_{\overline{B_{H_i}}})$	$\Pi_{EFL}(3) = \min(\delta_{P_{h_{jk}}}, \delta_{B_{h_{jk}}})$
	ω_{M_1} : Top m classes	ω_{M_2} : Top n classes	ω_{M_3} : Top p compounds in top n classes

Table 4-1: Performance functions and maximization operator used by experimentation strategy and Discovery research phase

As can be verified in Table 4-1 above the optimization process starts at the level of the chemical classes to end at candidate compound level. The top performing classes and compounds are selected using payoff information on one or two properties. Depending on the experimentation strategy used, only biological activity information is used to find the best performer, or a multi-objective function combining payoff information on both biological activity and bio-availability dimensions is used to find the best performing compounds in the search space. All experimentation strategies are modelled to only deal with chemical class related performance in HTS and H2L, and with chemical compound related performance in LO. Both Front-loaded paradigm and Front-loaded strategies use multi-objective optimisation but start doing so at different moments in the process. The chosen multi-objective function conservatively takes the minimum of both (P) and (B) detector values as the calculated performance value for the chemical class or candidate compound, depending on the Discovery research phase.

4.3.5 Annotated memory less adaptive systems model summary

The annotated adaptive systems model serves as the theoretical model underlying the simulation study, in which each experimentation strategy to conduct 'Concept Selection' in a Pharmaceutical Discovery context is modelled to execute an adaptive plan. This adaptive plan τ starts from the universe of potential compounds representing the extant search space and initial uncertainty about the environment[41]. Only an extremely small fraction of this universe, I will further call the prevalence π, is composed of candidate compounds with therapeutic effect. Through the execution of successive selection and optimization loops the plan improves the fit of the solution with the target. The details of these η(t), I(t), Ω(t) adaptation loops can be found in the table cells of Figure 4-3 below detailing the conceptual specification of the annotated adaptive system emulating the various experimentation strategies used in Pharmaceutical discovery to execute the 'Concept selection' complexity-handling mode.

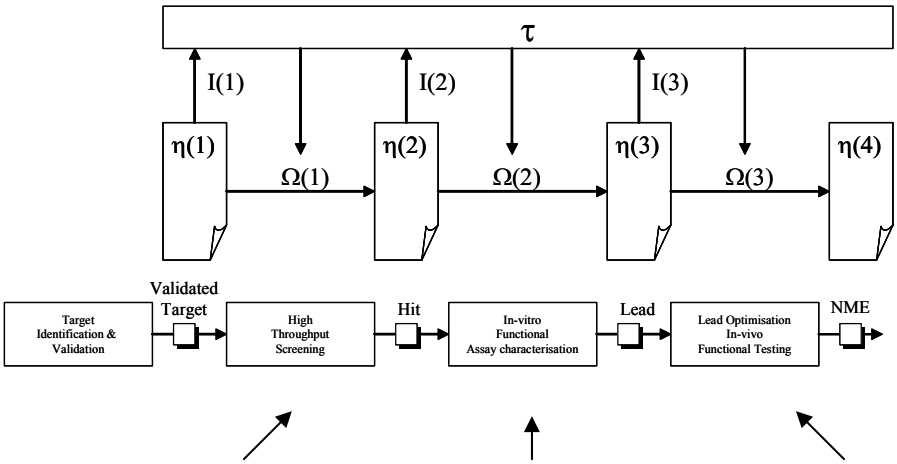
Summarizing, the conceptual specification of the annotated adaptive systems model above together with the Bayesian predictive performance criteria specified in my previous research project form a complete specification of a problem in adaptation (Holland, 1992: 28), since adaptive plan objects (η,Ω,I,τ) and evaluation criteria χ have been specified within the context ε of a chemical structure being adapted to optimally fit a biological target.

Finally, it should be noted that the annotated adaptive plan τ is modelled as a *memory less system*. As specified above the total information received by τ or input history up to time t is given by the sequence $\langle I(1), I(2),..., I(t-1)\rangle$ where no information is retained in the system. However, following Holland (1992: 23) one could think of η(t) being composed of $\eta_1(t)$, the set of structures tested against the environment at time t and the memory M(t) representing the retained parts of the input history. Then, more generally, since we know from above that the plan adapts the set of structures η(t) to become η(t+1) through the two-argument function $\tau : I \times \eta \to \eta$ we could say that t through $\tau : I \times (\eta \times M) \to (\eta \times M)$ also updates the memory from M(t) to M(t+1), where

$$\tau_M (I(t), \eta(t), M(t)) = M(t+1) \qquad \textbf{(4-9)}$$

is that part of the plan τ which updates the plan's memory. Thus, the conceptual framework is general enough to investigate mechanisms of memory update in greater detail. However, to limit the complexity of my theory building effort, I chose to model the annotated adaptive system as being memory less, only using information about the environment at hand at time t.

[41] A nontrivial problem of adaptation exists only when the adaptive plan is faced with an initial uncertainty about its environment (Holland, 1992: 25)



	HTS	H2L	LO
Resulting Sets	$\eta(2) = \underset{i=1}{\overset{m}{\Upsilon}} H_i$	$\eta(3) = \underset{j=1}{\overset{n}{\Upsilon}} H_j$	$\eta(4) = \{h_1^+, ..., h_p^+\}$
Old paradigm	$I_{\eta(1)} = \{\delta_{\overline{P_{H_i}}}\}$ $\Pi_{Old}(1) = \delta_{\overline{P_{H_i}}}$ $\omega_{s_1} : \delta_{\overline{P_{H_i}}} \geq p_{\min}$ ω_{M_1} : Top m classes to form η(2)	$I_{\eta(2)} = \{\delta_{\overline{P_{H_j}}}\}$ $\Pi_{Old}(2) = \delta_{\overline{P_{H_j}}}$ $\omega_{s_2} : \delta_{\overline{P_{H_j}}} \geq p_{\min}$ ω_{M_2} : Top n classes to form η(3)	$I_{\eta(3)} = \{\delta_{P_{h_{jk}}}\}$ $\Pi_{Old}(3) = \delta_{P_{h_{jk}}}$ $\omega_{s_1} : \delta_{\overline{P_{H_i}}} \geq p_{\min}$ ω_{M_3} : Top p compounds in top n classes: η(4)
Front-loaded paradigm	$I_{\eta(1)} = \{\delta_{\overline{P_{H_i}}}\}$ $\Pi_{FL}(1) = \delta_{\overline{P_{H_i}}}$ $\omega_{s_1} : \delta_{\overline{P_{H_i}}} \geq p_{\min}$ ω_{M_1} : Top m classes to form η(2)	$I_{\eta(2)} = \{\delta_{\overline{P_{H_j}}}, \delta_{\overline{B_{H_j}}}, \delta_{\overline{T_{H_j}}}\}$ $\Pi_{FL}(2) = \min(\delta_{\overline{P_{H_i}}}, \delta_{\overline{B_{H_i}}})$ $\omega_{s_2} : \delta_{\overline{P_{H_j}}} \geq p_{\min}$, $\delta_{\overline{B_{H_j}}} \geq b_{\min}, \delta_{\overline{T_{H_j}}} \leq t_{\max}$ ω_{M_2} : Top n classes to form η(3)	$I_{\eta(3)} = \{\delta_{P_{h_{jk}}}, \delta_{B_{h_{jk}}}, \delta_{T_{h_{jk}}}\}$ $\Pi_{FL}(3) = \min(\delta_{P_{h_{jk}}}, \delta_{B_{h_{jk}}})$ $\omega_{s_3} : \delta_{P_{h_{jk}}} \geq p_{\min}$, $\delta_{B_{h_{jk}}} \geq b_{\min}, \delta_{T_{h_{jk}}} \leq t_{\max}$ ω_{M_3} : Top p compounds in top n classes: η(4)
Early Front Loading	$I_{\eta(1)} = \{\delta_{\overline{P_{H_i}}}, \delta_{\overline{B_{H_i}}}, \delta_{\overline{T_{H_i}}}\}$ $\Pi_{EFL}(1) = \min(\delta_{\overline{P_{H_i}}}, \delta_{\overline{B_{H_i}}})$ $\omega_{s_1} : \delta_{\overline{P_{H_i}}} \geq p_{\min}$, $\delta_{\overline{B_{H_i}}} \geq b_{\min}, \delta_{\overline{T_{H_i}}} \leq t_{\max}$ ω_{M_1} : Top m classes to form η(2)	$I_{\eta(2)} = \{\delta_{\overline{P_{H_j}}}, \delta_{\overline{B_{H_j}}}, \delta_{\overline{T_{H_j}}}\}$ $\Pi_{EFL}(2) = \min(\delta_{\overline{P_{H_i}}}, \delta_{\overline{B_{H_i}}})$ $\omega_{s_2} : \delta_{\overline{P_{H_j}}} \geq p_{\min}$, $\delta_{\overline{B_{H_j}}} \geq b_{\min}, \delta_{\overline{T_{H_j}}} \leq t_{\max}$ ω_{M_2} : Top n classes to form η(3)	$I_{\eta(3)} = \{\delta_{P_{h_{jk}}}, \delta_{B_{h_{jk}}}, \delta_{T_{h_{jk}}}\}$ $\Pi_{EFL}(3) = \min(\delta_{P_{h_{jk}}}, \delta_{B_{h_{jk}}})$ $\omega_{s_3} : \delta_{P_{h_{jk}}} \geq p_{\min}$, $\delta_{B_{h_{jk}}} \geq b_{\min}, \delta_{T_{h_{jk}}} \leq t_{\max}$ ω_{M_3} : Top p compounds in top n classes: η(4)

Figure 4-3: Conceptual specification of the annotated adaptive systems model

The conceptual model formed the basis to build a computer simulation program in Excel Visual Basic Applications (VBA). Theory building results using simulation-based experiments with the *memory less annotated adaptive plans* formally described above emulating experimentation strategies during Concept Selection in a Pharmaceutical Discovery context will be discussed in the following paragraphs.

4.4 THE SIMULATION EXPERIMENT

4.4.1 Model parameters

Parameters used in the simulation model describe the shape of the extant search space, and the experimentation strategies used by the innovation team to conduct 'Concept Selection'. Table 4-2 below provides a summary of the value ranges of various parameters used in the simulation experiments.

Simulation Model Parameter Values				
Parameters	**Base simulation values**			
	Low	**Medium**	**High**	Additional Values Explored
Shape of Extant Search Space:				
Compound (P), (B), (T) value	1	4,5	9	Continuous across (1-9)
Exponential steepness a, b				range
Total # of virtual compounds	4	5 24K	6	
# of classes		80		
# of ref. compounds/class		15		
# of compounds/ref. comp.		20		
SD of ref. comp. in class	0,1	1	4	0,2 2
SD comp. around ref. comp	0,1	1	4	0,2 2
Marginal probability $p(H^+)$	0,05	0,1		
Experimentation Strategy				
# of classes in (HTS,H2L); m, n	(1,1)	(5,1)	(5,3)	(5,2)
# of compounds selected in (LO); p	5	10	15	1, 3, 10000
Surrogate Marker Chain*				
Tightness (HTS,H2L)	48%	70%	80%	50% 90%
Tightness (H2L,LO)	48%	70%	80%	
Detector Value ranges** $(p_{\min}, b_{\min}, t_{\max})$		(5,5,5)		

* specified correlation percentages have been translated into measurements errors on properties in the various discovery research phases

** values applied across all experimentation strategies

Table 4-2: Simulation Model Parameter Values

A virtual compound is described by a real value for its three properties (P), (B), (T) ranging from 1 to 9. This was done through random sampling of a negative exponential compound properties distribution for (P) and (B), given by;

$$P_{compound} = 9{,}5 - a.\ln\left[y(e^{9/a} - 1) + 1\right] \quad \textbf{(4-10)}$$

$$B_{compound} = 9{,}5 - b.\ln\left[y(e^{9/b} - 1) + 1\right] \quad \textbf{(4-11)}$$

and of the exponential distribution for (T);

$$T_{compound} = 10 - (0{,}25y + 0{,}75)P_{compound} \quad \textbf{(4-12)}$$

where y was chosen from a uniform [0,1] distribution, and a and b were used to parameterize distribution steepness. An interview with a toxicology specialist in PharmaCo indicated that highly potent compounds are less toxic than their less potent counterparts since toxicity is loosely connected to dosage needed; the higher the dosage needed to have an effect on the biological target, the higher the risk the drug will be toxic[42]. Therefore, in my virtual model the (T) distribution was modeled as the inverse of the (P) distribution adjusted with a light random effect. This led to a virtual compound population described by their (P,B,T) serving as input to the experimentation strategies.

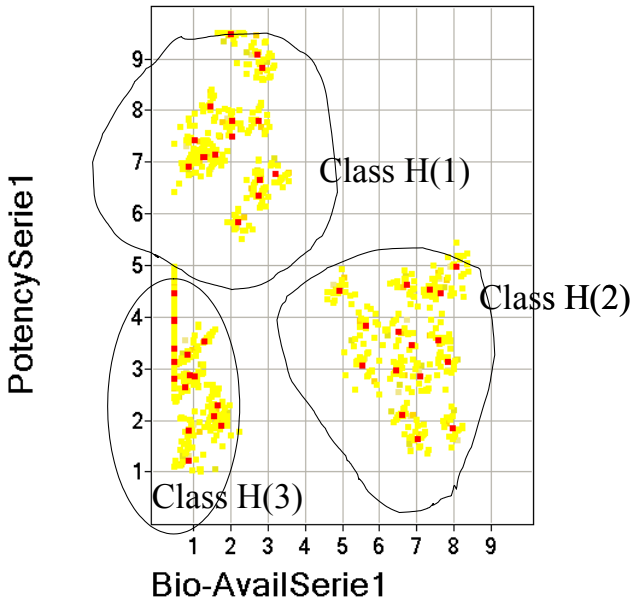


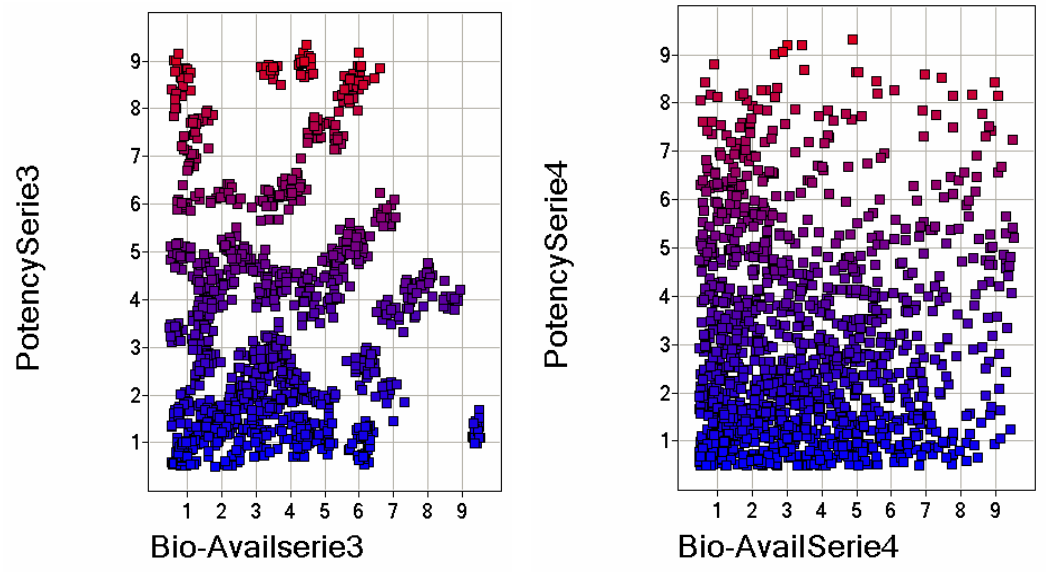
Figure 4-4 Virtual search space composition

The extant search space is described by the total number of virtual compounds it contains. It is generated in several steps. First, classes get assigned a value for (P), (B) and (T) through random sampling of the exponential distributions of these properties.

[42] See reference (16) in Appendix C.

Distribution steepness parameters a end b are used to modulate the prevalence of the population given a set of detector values $(p_{\min}, b_{\min}, t_{\max})$ for compound properties. Then, around these class values reference compounds are randomly generated following a normal distribution with a certain standard deviation (SD) chosen from the ranges described in Table 4-2 above. This is done while, in reality, classes are investigated by scientists in HTS and H2L by using a set of reference compounds describing the classes. Finally, virtual compounds are randomly generated around reference compounds using a normal distribution with a certain standard deviation.

Figure 4-4 gives an example of a virtual solution landscape composed of three classes H(1), H(2), and H(3). Each class consists of reference compounds (highlighted) and virtual compounds. Number of chemical classes, number of reference compounds used per class, and number of compounds per reference compounds, can be specified.



(a) **(b)**

Figure 4-5 Examples of concentrated (a) and rugged (b) virtual extant search spaces

Both numbers of objects and standard deviations can be used to vary the solution landscape shape from highly concentrated to highly rugged. Figure 4-5 above provides an example of a reference compound type representation of a concentrated and a rugged solution landscape, featuring a standard deviation for all descriptors (P,B,T) of respectively 0,2 for (a) and 1 for (b). Toxicity is represented here using a colour scale. It can be verified that red indicates low toxicity. In contrast, blue indicates high toxicity.

This top-down generation of classes emulates their relatedness on properties (P) and (B) of the virtual compounds they contain. On the other hand, toxicity for these virtual compounds is generated bottom-up since this is considered to be not class-related; virtual compound values are randomly generated through random sampling of a (T) distribution, which is loosely connected to the inverse of the (P) distribution. Then these compound values are averaged to get to the toxicity of a reference compound, which in turn is used to generate the average toxicity of a class.

The simulation experiment compares three experimentation strategies –*Old paradigm, Front-loaded (FL) paradigm, Early Front-loading (EFL)*- for predictive performance. Parameters activated for this study include the number of classes used in HTS and H2L –the latter determining the LO search space-, determining the resource effort spent in the different phases. The surrogate marker chain tightness can be varied at the level of discovery phase HTS-H2L and H2L-LO transitions. As explained before, surrogate marker chain tightness or correlation is expressed as a percentage in Table 4-2. This is translated in the simulation model by sets of measurement errors on the properties (P), (B), and (T) in the various phases. Detector values used for selection can be set for all discovery phases and experimentation strategies used.

Figure 4-6 below depicts the simulation framework described in my previous research project and connects it to the adaptive plan representation of experimentation strategies in the Pharmaceutical Discovery context. During Concept Selection, adaptive plans select and optimize candidate compounds and make predictions whether they are fit to be transferred to Concept Characterization (h^+), or not (h^-).

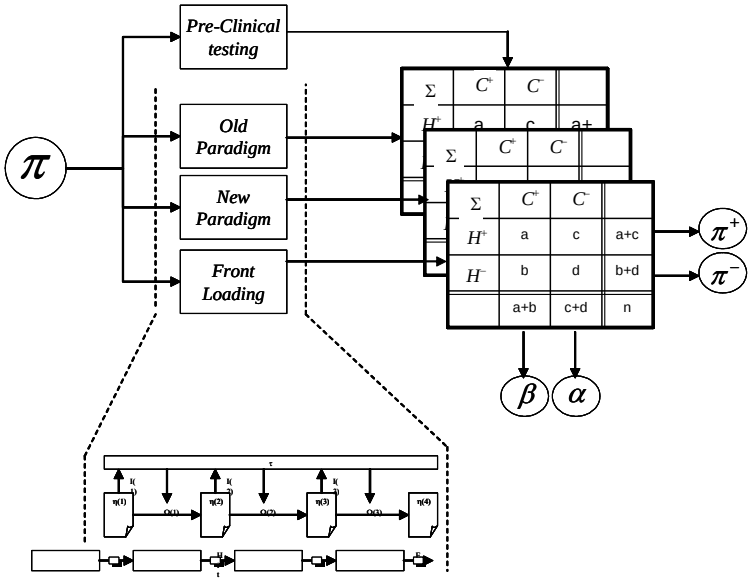


Figure 4-6: Predictive performance measurement concept

Positive and negative predictive value. As discussed before, Bayesian inference criteria π^+ and π^- were used to evaluate predictive performance of various experimentation strategies. The positive predictive performance π^+, denoted $p(C^+ \mid H^+)$ is read as the probability $p(C^+)$ that a compound will pass Concept Characterization and Concept Application testing, given it has been declared active $p(H^+)$ by the experimentation strategy used during Concept Selection. The fraction of really active compounds after Concept Characterization and Concept Application, although declared inactive by the experimentation strategy, is called π^- and denoted $p(C^+ \mid H^-)$. An experimentation strategy featuring high positive and negative predictive performance has a high value for π^+ and a low π^-, indicating

respectively a high number of confirmations of positive decisions made and a reduced number of lost opportunities. The latter category is not observable in pharmaceutical practice since, in reality, declined compounds are not transferred to later stages.

Overall quality selection power. Since this simulation study is an exploratory experiment I wanted to measure experimentation strategy effectiveness in a different way than through its predictive performance. Then, triangulation of *'Overall quality selection power'* with predictive performance results would increase confidence in comparative strategy effectiveness' findings. To test for the overall quality selection power of the experimentation strategies, all virtual compounds of the generated extant search space were given a rank. The latter was calculated using $\min(\delta_{P_h}, \delta_{B_h})$ as an overall quality indicator. The higher this number, the higher the *overall* quality –since considering *both* P and B as opposed to *only* P- of the generated compound. Then, the search space was ordered in the sense that the compound featuring the highest $\min(\delta_{P_h}, \delta_{B_h})$ was given rank R=1. Now, for each experimentation strategy the average rank of all selected compounds h^+ was calculated. Then, the experimentation strategy featuring the lowest $\overline{R}$ has the best *overall* quality selection power. Overall quality selection power is not observable in empirical practice.

Business performance. The business value distribution U(d) of an experimentation strategy was formally derived from the positive and negative predictive values π^+ and π^- following the formula derived in section 3.5.2. To calculate this business performance indicator for each experimentation strategy d, PharmaCo data needed to be provided for product Revenue R, project Clinical Development cost C_{Clin}, project Discovery Research cost C_D, and marginal probability p(H). Considering confidentiality fictitious numbers were used, which do reflect industrial reality.

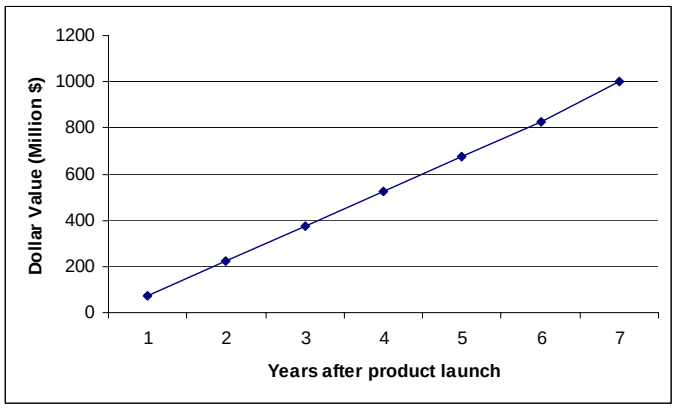


Figure 4-7 Blockbuster Revenue profile (typical)

As depicted in Figure 4-7 product revenue R was assumed to build up linearly to one billion dollar a year, a typical industry average of a blockbuster product (Duyck, 2003). C_D and C_{Clin} amount to \$20 and \$400 million dollar respectively. These numbers represent typical project costs excluding the contribution for attrition

which is usually taken into account when specifying development costs (see e.g. (Kennedy, 1997; Duyck , 2003)).

Marginal probability p(H) was characterized by taking the pharmaceutical R&D industry average for the discovery success rate; $p(H^+) = 0.1$ (Kennedy, 1997). Also a significantly lower value was tested[43].

4.4.2 The simulation

The simulation experiment was conducted in two steps, exploring the research questions and checking for robustness of the results under various resources, scientific, and solution landscape induced constraints; a screening experiment and a confirmatory experiment was executed.

# Runs	Landscape Shape	Funnel Shape	Surrogate Marker Chain Tigthness
Compare Frontloaded Strategies **1 - 9**	Tight landscape SD ref comp in class=0.1 SD comp around ref comp=0.1	Fixed at (HTS, H2L) = (5, 2)	Varying between 80%, 70%, 48%
10 – 18	Wide landscape SD ref comp in class=1 SD comp around ref comp=1	Fixed at (HTS, H2L) = (5, 2)	Varying between 80%, 70%, 48%
Compare Parallelism **19 – 30**	Tight landscape	Varying between (1,1), (5,1), (5,2), (5,3)	Varying between 80%, 70%, 48%
31 – 42	Wide landscape	Varying between (1,1), (5,1), (5,2), (5,3)	Varying between 80%, 70%, 48%

Table 4-3 Screening simulation experiment block design

As depicted in Table 4-3 above the screening experiment was composed of 42 simulations of 100 runs each, comparing in a first block (runs 1-18) the three Concept Selection experimentation strategies *–Old paradigm, Front-loading, Early Front-loading-* on their predictive performance and overall quality selection power. Three surrogate marker chains were used throughout this process. The funnel shape was held constant. The second block of simulations (runs 19-42) compared the impact of several parallel concept exploration types in an *Early Front-loaded* experimentation strategy on

[43] See Table 4-2

their predictive performance and overall quality selection power. More specifically, the impact of using more or less classes in HTS and H2L was explored. Again, three surrogate marker chains were used throughout the process.

The confirmatory experiment repeated specific simulation runs from the screening experiment (highlighted blocks in Table 4-3) using a wide solution landscape but now at 200 runs per simulation to get statistically significant results. Predictive performance and overall quality selection power distributions were calculated for each experimentation strategy. In addition to this, business performance distributions were derived from the simulation runs.

4.4.3 Data analysis

All experiments were executed within a Monte Carlo design. Random sampling of P, B, and T distributions led to a virtual compound population described by their potency, bio-availability and toxicity serving as input to the experimentation strategies. To generate statistically significant data, each simulation run was repeated 100 or 200 times –depending on whether it was a screening or confirmatory experiment- to obtain distributions of performance variables. Two-way and One-way ANOVA was used on these performance data series to test for significance of differences in predictive, business performance and overall quality selection power of experimentation strategies.

4.5 SIMULATION MODEL BEHAVIOUR

Two-way and One-way ANOVA results of the Monte Carlo simulation experiments indicate that *the level of front-loading* and the *number of parallel solution concept explorations* used, significantly influence predictive performance and overall quality selection power of the Concept Selection complexity-handling mode. These results are robust for varying levels of tightness of the surrogate marker chain.

Furthermore, screening experiments showed that the simulation method used only provides meaningful results within an operating window determined by (1) the ruggedness of the solution landscape, and (2) the number of compounds selected in LO (p). Maximum ruggedness possible was the 'Wide landscape' described above. When standard deviations were allowed to move beyond the SD=1 range[44], results became meaningless while remaining stable for all variations of input parameters. Also, when p was allowed to move below a certain threshold, results again became meaningless while remaining stable for all variations of input parameters. Therefore, confirmatory experiments were conducted at p=15 which does not reflect reality (where typically p=1 to 3) but was the lowest value possible to get meaningful results.

[44] SD=2, 3 as documented in Table 4-2 under 'Additional Values Explored' of 'Shape of Extant Search Space'

In the following, simulation model behaviour during confirmatory experiments is examined in greater detail. Then, research propositions will be formulated based upon the results of this simulation study. This will be the basis for subsequent top-down theory generation.

4.5.1 Influence of the level of front-loading on predictive, business performance and overall quality selection power

Overall quality selection power. A two-way between-groups analysis of variance was conducted to explore the impact of the level of frontloading and varying degrees of tightness of the surrogate marker chain on the overall quality selection power of the experimentation strategy used during Concept Selection. SPSS results are depicted in Figure 4-8 and analysed after.

2. STRAT

Dependent Variable: QUAL

			99% Confidence Interval	
STRAT	Mean	Std. Error	Lower Bound	Upper Bound
Old	157,253	5,572	142,868	171,638
FL	128,239	5,624	113,720	142,757
EFL	123,461	5,582	109,050	137,871

3. tightness

Dependent Variable: QUAL

			99% Confidence Interval	
tightness	Mean	Std. Error	Lower Bound	Upper Bound
80%	124,560	5,503	110,352	138,768
70%	131,209	5,630	116,673	145,744
48%	153,184	5,643	138,615	167,752

Figure 4-8 Experimentation Strategy Quality two-way ANOVA results

Descriptive Statistics

Dependent Variable: QUAL

STRAT	tightness	Mean	Std. Deviation	N
Old	80%	151,434523	95,6522092	97
	70%	152,559201	93,3541455	92
	48%	167,764105	148,3527104	96
	Total	157,298068	115,3397995	285
FL	80%	108,267404	46,4415822	97
	70%	126,235461	61,5801238	94
	48%	150,212852	124,3227385	89
	Total	127,632198	84,7048330	280
EFL	80%	113,977257	50,8200279	98
	70%	114,830853	70,7344476	93
	48%	141,573571	106,0995471	93
	Total	123,293600	79,5909649	284
Total	80%	124,523486	70,4109750	292
	70%	131,114154	77,6277360	279
	48%	153,383597	127,7452894	278
	Total	136,139396	95,7060045	849

Tests of Between-Subjects Effects

Dependent Variable: QUAL

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Corrected Model	338817,226[b]	8	42352,153	4,789	,000	,044	38,312	,990
Intercept	15762879,8	1	15762879,84	1782,421	,000	,680	1782,421	1,000
STRAT	190009,518	2	95004,759	10,743	,000	,025	21,486	,957
TIGHTNES	126617,078	2	63308,539	7,159	,001	,017	14,317	,813
STRAT * TIGHTNES	18877,744	4	4719,436	,534	,711	,003	2,135	,061
Error	7428556,902	840	8843,520					
Total	23502685,0	849						
Corrected Total	7767374,128	848						

a. Computed using alpha = ,01

b. R Squared = ,044 (Adjusted R Squared = ,035)

Figure 4-8 Experimentation Strategy Quality two-way ANOVA results (cont.)

Multiple Comparisons

Dependent Variable: QUAL

	(I) strategy	(J) strategy	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Old	Old					
		FL	29,665870*	7,9128943	,001	6,548915	52,782825
		EFL	34,004467*	7,8847353	,000	10,969777	57,039157
	FL	Old	-29,665870*	7,9128943	,001	-52,782825	-6,548915
		FL					
		EFL	4,338597	7,9197952	,848	-18,798518	27,475712
	EFL	Old	-34,004467*	7,8847353	,000	-57,039157	-10,969777
		FL	-4,338597	7,9197952	,848	-27,475712	18,798518
		EFL					
Scheffe	Old	Old					
		FL	29,665870*	7,9128943	,001	5,585411	53,746329
		EFL	34,004467*	7,8847353	,000	10,009701	57,999233
	FL	Old	-29,665870*	7,9128943	,001	-53,746329	-5,585411
		FL					
		EFL	4,338597	7,9197952	,861	-19,762863	28,440057
	EFL	Old	-34,004467*	7,8847353	,000	-57,999233	-10,009701
		FL	-4,338597	7,9197952	,861	-28,440057	19,762863
		EFL					

Based on observed means.

*. The mean difference is significant at the ,01 level.

Multiple Comparisons

Dependent Variable: Quality

	(I) Tightness	(J) Tightness	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	80%	80%					
		70%	-6,590668	7,8729431	,680	-29,590908	16,409572
		48%	-28,860111*	7,8801809	,001	-51,881496	-5,838725
	70%	80%	6,590668	7,8729431	,680	-16,409572	29,590908
		70%					
		48%	-22,269443	7,9692175	,015	-45,550942	1,012057
	48%	80%	28,860111*	7,8801809	,001	5,838725	51,881496
		70%	22,269443	7,9692175	,015	-1,012057	45,550942
		48%					
Scheffe	80%	80%					
		70%	-6,590668	7,8729431	,705	-30,549548	17,368212
		48%	-28,860111*	7,8801809	,001	-52,841017	-4,879204
	70%	80%	6,590668	7,8729431	,705	-17,368212	30,549548
		70%					
		48%	-22,269443	7,9692175	,021	-46,521304	1,982419
	48%	80%	28,860111*	7,8801809	,001	4,879204	52,841017
		70%	22,269443	7,9692175	,021	-1,982419	46,521304
		48%					

Based on observed means.

*. The mean difference is significant at the ,01 level.

Figure 4-8 Experimentation Strategy Quality two-way ANOVA results (cont.)

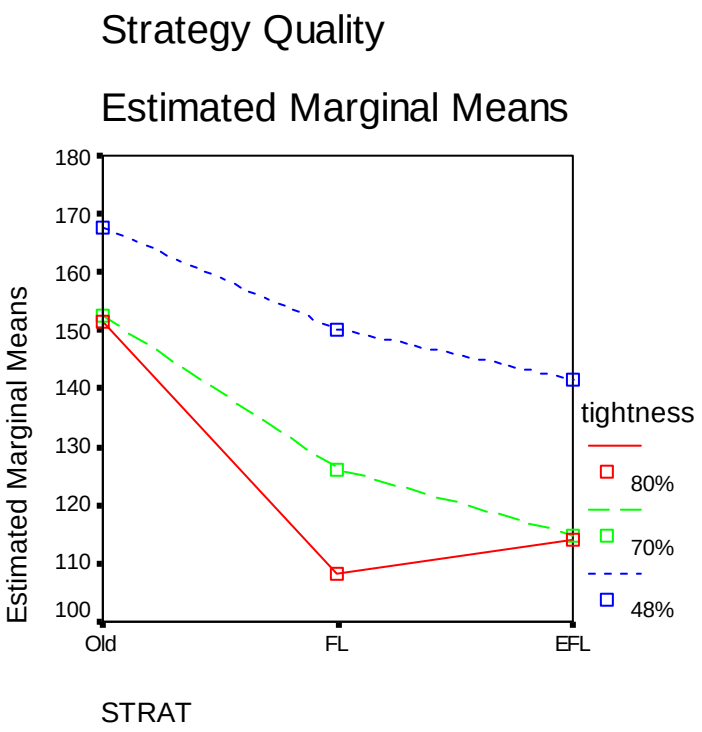


Figure 4-8 Experimentation Strategy Quality two-way ANOVA results (cont.)

For each strategy (Old, FL, EFL) three degrees of tightness were considered; 80%, 70%, and 48%. Referring to Figure 4-8 there was a statistically significant main effect for both levels of front-loading [F(2; 10.74), p= .000] and degrees of tightness [F(2; 7.16), p= .001]. However, the effect size was small (partial eta squared=.025 and .017 respectively). To increase confidence in simulation results both post-hoc comparisons using Tukey HSD and Scheffe's were performed. Both tests indicated that Old Paradigm strategy (M=157.3; SE=5.57) performs significantly worse than Front-loaded paradigm (M=128.2; SE=5.62) or Early Frontloaded (M=123.5; SE=5.58) experimentation strategies. However, Front-loaded and Early Front-loaded strategies do not differ significantly.

Surrogate marker chain tightness negatively influences overall quality selection power from a certain minimum level of tightness of the chain. Both Tukey HSD and Scheffe's tests indicate that 80% (M=124.6; SE=5.5) and 70% (M=131.2; SE=5.6) levels of chain tightness do not differ significantly. However, at 48% (M=153.2; SE=5.6) a statistically significant deterioration of performance is observed.

The interaction effect [F(4; 0.53), p= .71] did not reach statistical significance.

Summarizing, these simulation results indicate that front-loaded strategies select *overall* better candidate compounds. This is true for differing levels of surrogate marker chain tightness.

Predictive performance. A two-way between-groups analysis of variance was conducted to explore the impact of the level of frontloading and varying degrees of tightness of the surrogate marker chain on the positive and negative predictive performance of the experimentation strategy used during Concept Selection. SPSS results are depicted in Figure 4-9 and analysed after.

1. strategy

Dependent Variable: pi+

			99% Confidence Interval	
strategy	Mean	Std. Error	Lower Bound	Upper Bound
Old	,994	,002	,990	,998
FL	,995	,002	,991	,999
EFL	,993	,002	,989	,997

2. tightness

Dependent Variable: pi+

			99% Confidence Interval	
tightness	Mean	Std. Error	Lower Bound	Upper Bound
80%	,996	,001	,992	1,000
70%	,995	,002	,991	,999
48%	,991	,002	,987	,995

3. strategy * tightness

Dependent Variable: pi+

				99% Confidence Interval	
strategy	tightness	Mean	Std. Error	Lower Bound	Upper Bound
Old	80%	,994	,003	,987	1,001
	70%	,996	,003	,989	1,003
	48%	,992	,003	,985	,998
FL	80%	,995	,003	,989	1,002
	70%	,997	,003	,990	1,003
	48%	,994	,003	,987	1,001
EFL	80%	,998	,003	,991	1,005
	70%	,993	,003	,986	1,000
	48%	,988	,003	,981	,994

Figure 4-9 Predictive performance two-way ANOVA results (n=100)

Tests of Between-Subjects Effects

Dependent Variable: pi+

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Corrected Model	,007[b]	8	,001	1,379	,202	,013	11,028	,394
Intercept	837,961	1	837,961	1283919,786	,000	,999	1283919,786	1,000
STRAT	,001	2	,000	,690	,502	,002	1,380	,055
TIGHTNES	,004	2	,002	2,851	,058	,007	5,702	,322
STRAT * TIGHTNES	,003	4	,001	,987	,414	,005	3,948	,135
Error	,548	840	,001					
Total	839,280	849						
Corrected Total	,555	848						

a. Computed using alpha = ,01

b. R Squared = ,013 (Adjusted R Squared = ,004)

1. strategy

Dependent Variable: pi-

strategy	Mean	Std. Error	99% Confidence Interval	
			Lower Bound	Upper Bound
Old	,057	,001	,054	,060
FL	,055	,001	,052	,058
EFL	,054	,001	,051	,056

2. tightness

Dependent Variable: pi-

tightness	Mean	Std. Error	99% Confidence Interval	
			Lower Bound	Upper Bound
80%	,055	,001	,052	,058
70%	,057	,001	,054	,060
48%	,054	,001	,051	,057

3. strategy * tightness

Dependent Variable: pi-

strategy	tightness	Mean	Std. Error	99% Confidence Interval	
				Lower Bound	Upper Bound
Old	80%	,058	,002	,053	,062
	70%	,058	,002	,053	,063
	48%	,055	,002	,050	,060
FL	80%	,053	,002	,048	,058
	70%	,058	,002	,053	,062
	48%	,055	,002	,050	,059
EFL	80%	,054	,002	,049	,059
	70%	,055	,002	,050	,060
	48%	,053	,002	,048	,057

Figure 4-9 Predictive performance two-way ANOVA results (n=100) (cont.)

Tests of Between-Subjects Effects

Dependent Variable: pi-

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Corrected Model	,003[b]	8	,000	1,181	,307	,010	9,451	,316
Intercept	2,743	1	2,743	7891,973	,000	,899	7891,973	1,000
STRAT	,001	2	,001	2,048	,130	,005	4,096	,209
TIGHTNES	,001	2	,001	1,737	,177	,004	3,473	,168
STRAT * TIGHTNES	,001	4	,000	,470	,757	,002	1,882	,053
Error	,310	891	,000					
Total	3,056	900						
Corrected Total	,313	899						

a. Computed using alpha = ,01

b. R Squared = ,010 (Adjusted R Squared = ,002)

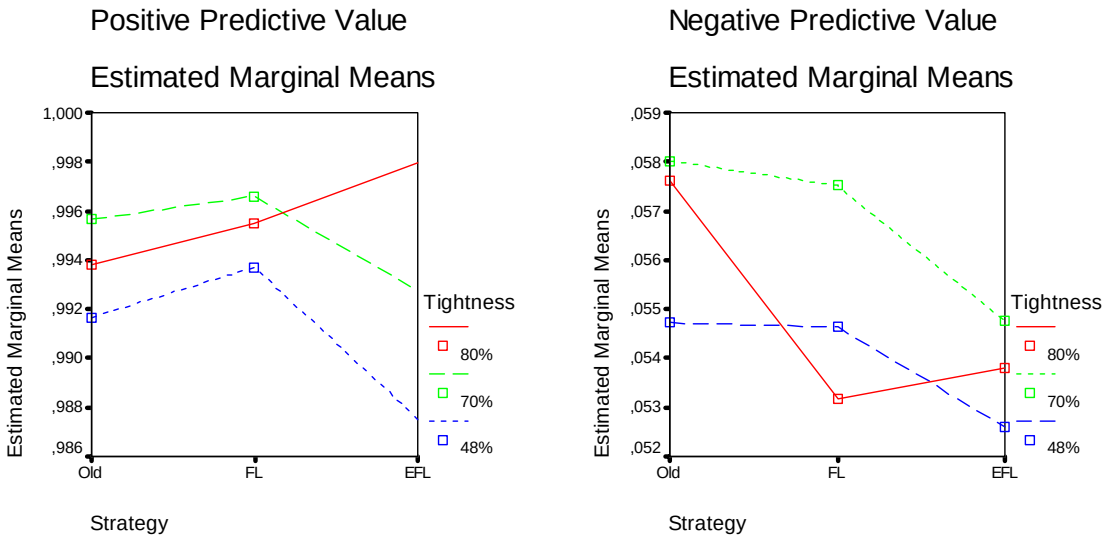


Figure 4-9 Predictive performance two-way ANOVA results (n=100) (cont.)

For each strategy (Old, FL, EFL) three degrees of tightness were considered; 80%, 70%, and 48%. Referring to Figure 4-9 during the screening experiment at n=100, neither of both main effects for both levels of front-loading nor degrees of tightness reached statistical significance; strategy type [F(2; 0.69), p= .5], tightness level [F(2; 2.85), p= .06], and interaction effect [F(4; 0.99), p= .41] for positive predictive performance, nor strategy type [F(2; 2.05), p= .13], tightness level [F(2; 1.74), p= .18] or interaction effect [F(2; 0.47), p= .76] for negative predictive performance.

Visual inspection of the estimated marginal means plots in Figure 4-9 above does indicate a slightly raising trend for positive predictive performance for front-loaded strategies, deteriorated by lowering levels of surrogate marker chain tightness. Also, a downward trend is observed for negative predictive value towards front-loaded strategies, impacted by the surrogate marker chain tightness. Also, from the descriptive statistics table in Figure 4-9 above, it becomes clear that for the positive predictive value the 95% confidence intervals for 80% (M=0.996; interval 0.993-0.999) and 70% (M=0.995; interval 0.992-0.998) are overlapping where the interval for 48% (M=0.991; interval 0.988-0.994) is below, indicating a deteriorating effect of surrogate marker chain tightness on positive predictive performance starting from a certain level. Conversely, the negative predictive value does not seem to be negatively impacted by

lowering levels of surrogate marker chain tightness since descriptors for 80% (M=0,055; interval 0.053-0.057), 70% (M=0.057; interval 0.055-0.059), and 48% (M=0.054; interval 0.052-0.056) all share about the same confidence interval.

Summarizing, the screening experiment indicates that positive predictive performance is negatively impacted by falling levels of surrogate marker chain tightness. Negative predictive performance does not seem to be impacted by the latter. However, due to the very small differences between Old Paradigm, Front-loaded paradigm and Early Front-loading simulation run descriptors, none of these effects indicate conclusive statistical significance. Therefore, since being conclusive about predictive performance of experimentation strategies is crucial for developing theory in my work, in the following one-way ANOVA confirmatory experiment the number of runs was elevated to n=200. Now, the solution landscape was fixed at wide and the surrogate marker chain tightness set to 70%. The latter tightness was preferred by the PharmaCo team since it represents best the capabilities of the present bio-chemical scientific reality. The choice for the wide landscape (SD reference compound in class; SD of compound around reference compound[45]) = (1; 1) was given by the limitations of the simulation method used. To the latter, the screening experiment had shown that very tight landscapes (0.1; 0.1) or very wide landscapes (4; 2) led to non-discriminating results between different experimentation strategies' performance variables.

So, finally a one-way between-groups analysis of variance was conducted in a confirmatory simulation experiment to probe for the impact of the level of frontloading on the business, positive and negative predictive performance of the experimentation strategy used during Concept Selection.

Descriptives

			N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	Between-Component Variance
							Lower Bound	Upper Bound			
PIPLUS	Old		200	,974752	,0504591	,0035680	,967716	,981788	,6711	1,0000	
	FL		199	,992351	,0266486	,0018891	,988626	,996077	,7522	1,0000	
	EFL		200	,992737	,0364334	,0025762	,987657	,997817	,5156	1,0000	
	Total		599	,986604	,0399337	,0016316	,983400	,989809	,5156	1,0000	
	Model	Fixed Effects			,0391058	,0015978	,983466	,989742			
		Random Effects				,0059344	,961071	1,012138			,0000980
PIMINUS	Old		200	,081718	,0355738	,0025154	,076758	,086679	,0131	,2201	
	FL		200	,080827	,0302083	,0021360	,076615	,085039	,0113	,1904	
	EFL		200	,080278	,0281459	,0019902	,076353	,084202	,0193	,1605	
	Total		600	,080941	,0314185	,0012827	,078422	,083460	,0113	,2201	
	Model	Fixed Effects			,0314654	,0012846	,078418	,083464			
		Random Effects				,0012846[a]	,075414[a]	,086468[a]			-,0000044

[a]. Warning: Between-component variance is negative. It was replaced by 0.0 in computing this random effects measure.

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
PIPLUS	23,189	2	596	,000
PIMINUS	3,542	2	597	,030

Figure 4-10 Predictive performance of Experimentation Strategies one-way ANOVA results (n=200)

[45] see Table 4-2

ANOVA

				Sum of Squares	df	Mean Square	F	Sig.
PIPLUS	Between Groups	(Combined)		,042	2	,021	13,794	,000
		Linear Term	Unweighted	,032	1	,032	21,151	,000
			Weighted	,032	1	,032	21,151	,000
			Deviation	,010	1	,010	6,437	,011
			Contrast					
	Within Groups			,911	596	,002		
	Total			,954	598			
PIMINUS	Between Groups	(Combined)		,000	2	,000	,107	,899
		Linear Term	Unweighted					
			Weighted					
			Deviation	,000	1	,000	,004	,950
			Contrast	,000	1	,000	,210	,647
	Within Groups			,591	597	,001		
	Total			,591	599			

Multiple Comparisons

Dependent Variable		(I) STRATEGY	(J) STRATEGY	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval	
							Lower Bound	Upper Bound
PIPLUS	Tukey HSD	Old	Old					
			FL	-,017599*	,0039155	,000	-,029051	-,006148
			EFL	-,017985*	,0039106	,000	-,029422	-,006548
		FL	Old	,017599*	,0039155	,000	,006148	,029051
			FL					
			EFL	-,000386	,0039155	,995	-,011837	,011066
		EFL	Old	,017985*	,0039106	,000	,006548	,029422
			FL	,000386	,0039155	,995	-,011066	,011837
			EFL					
	Scheffe	Old	Old					
			FL	-,017599*	,0039155	,000	-,029528	-,005670
			EFL	-,017985*	,0039106	,000	-,029899	-,006071
		FL	Old	,017599*	,0039155	,000	,005670	,029528
			FL					
			EFL	-,000386	,0039155	,995	-,012315	,011543
		EFL	Old	,017985*	,0039106	,000	,006071	,029899
			FL	,000386	,0039155	,995	-,011543	,012315
			EFL					
PIMINUS	Tukey HSD	Old	Old					
			FL	,000892	,0031465	,957	-,008311	,010094
			EFL	,001441	,0031465	,891	-,007762	,010643
		FL	Old	-,000892	,0031465	,957	-,010094	,008311
			FL					
			EFL	,000549	,0031465	,983	-,008654	,009752
		EFL	Old	-,001441	,0031465	,891	-,010643	,007762
			FL	-,000549	,0031465	,983	-,009752	,008654
			EFL					
	Scheffe	Old	Old					
			FL	,000892	,0031465	,961	-,008695	,010478
			EFL	,001441	,0031465	,901	-,008146	,011027
		FL	Old	-,000892	,0031465	,961	-,010478	,008695
			FL					
			EFL	,000549	,0031465	,985	-,009037	,010135
		EFL	Old	-,001441	,0031465	,901	-,011027	,008146
			FL	-,000549	,0031465	,985	-,010135	,009037
			EFL					

*. The mean difference is significant at the .01 level.

Figure 4-10 Predictive performance of Experimentation Strategies one-way ANOVA results (n=200) (cont.)

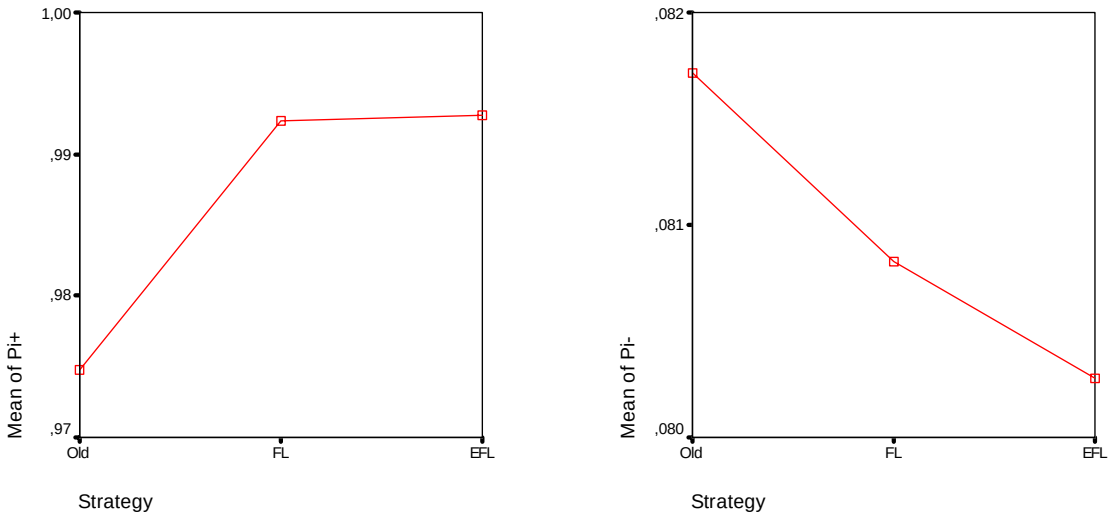


Figure 4-10 Predictive performance of Experimentation Strategies one-way ANOVA results (n=200) (cont.)

Referring to Figure 4-10, now there was a statistically significant difference at the p<.01 level for positive predictive performance of experimentation strategies. Although a downward trend towards front-loaded strategies was observed congruent with the screening experiment, negative predictive performance did not differ significantly [F(2, 597)=0.107; p=0.899] across the three strategies considered.

On positive predictive performance Old Paradigm, Front-loading and Early Front-loading differed significantly [F(2, 596)=13,79; p<.01] although the actual differences in mean scores are quite small. Also, this result needs to be moderated by the fact that the effect size, calculated using eta squared, was 0.044 indicating a small to medium impact of the independent variables on the outcome variable. Levene's test showed non-homogeneity of variance across the three groups –which I consider to be normal considering the different experimentation strategy types- implying the need for studying differences between groups at the p<.01 level of significance. Both Tukey's HSD and Scheffe's post-hoc comparisons indicated that Old Paradigm (M=0.974; SD=0.05) is significantly outperformed by Front-loaded paradigm (M=0.992; SD=0.026), and Early Front-loading (M=0.993; SD=0.036) at the p<.01 level of significance for positive predictive performance. Front-loaded paradigm and Early Front-loading did not differ significantly.

Business performance. A one-way between-group analysis of variance was conducted in a confirmatory simulation experiment to probe for the impact of the level of frontloading on the business performance of the experimentation strategy used during Concept Selection (see Figure 4-11 for results). There was a statistically significant difference at the p<.01 level for the different experimentation strategies. Financial assumptions used to perform ANOVA are discussed above.

Referring to Figure 4-11 Old Paradigm, Front-loaded paradigm and Early Front-loading differed significantly [F(2, 596)=12.9; p<.01] on business performance. This result needs to be moderated by the fact that the effect size, calculated using eta

squared, was 0.041 indicating a small to medium impact of the independent variables on the outcome variable. Levene's test showed non-homogeneity of variance across the three groups –which I consider to be normal considering the different experimentation strategy types- implying the need for studying differences between groups at the p<.01 level of significance.

Both Tukey's HSD and Scheffe's post-hoc comparisons indicated that Old Paradigm (M=2241.9; SD=137.8) is significantly outperformed by Front-loaded paradigm (M=2287.3; SD=78.2), and Early Front-loading (M=2290; SD=94) at the p<.01 level of significance for business performance. Front-loading and Early Front-loading did not differ significantly although the difference does amount to $2.7M in favour of Early Front-loading using the abovementioned assumptions.

Descriptives

Business Value ($)

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	Between-Component Variance
						Lower Bound	Upper Bound			
Old Paradigm		200	2241,963057	137,7652767	9,7414761	2222,753289	2261,172825	1452,3120	2406,0230	
FL		199	2287,393901	78,2267145	5,5453448	2276,458385	2298,329418	1733,9900	2440,0472	
EFL		200	2290,065872	94,0278671	6,6487742	2276,954778	2303,176966	1190,2434	2424,8277	
Total		599	2273,117149	108,5019251	4,4332685	2264,410480	2281,823817	1190,2434	2440,0472	
Model	Fixed Effects			106,4050028	4,3475906	2264,578689	2281,655609			
	Random Effects				15,6156595	2205,928389	2340,305909			674,8380740

Test of Homogeneity of Variances

Business Value ($)

Levene Statistic	df1	df2	Sig.
18,609	2	596	,000

ANOVA

Business Value ($)

			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		292128,636	2	146064,318	12,901	,000
	Linear Term	Unweighted	231388,079	1	231388,079	20,437	,000
		Weighted	231388,079	1	231388,079	20,437	,000
		Deviation	60740,557	1	60740,557	5,365	,021
Within Groups			6747926,675	596	11322,025		
Total			7040055,310	598			

Figure 4-11 Business Performance of Experimentation Strategies one-way ANOVA results (n=200)

Multiple Comparisons

Dependent Variable: Business Value ($)

	(I) STRATEGY	(J) STRATEGY	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Old Paradigm	Old Paradigm					
		FL	-45,430844*	10,6538594	,000	-76,590051	-14,271637
		EFL	-48,102815*	10,6405003	,000	-79,222951	-16,982679
	FL	Old Paradigm	45,430844*	10,6538594	,000	14,271637	76,590051
		FL					
		EFL	-2,671970	10,6538594	,966	-33,831177	28,487236
	EFL	Old Paradigm	48,102815*	10,6405003	,000	16,982679	79,222951
		FL	2,671970	10,6538594	,966	-28,487236	33,831177
		EFL					
Scheffe	Old Paradigm	Old Paradigm					
		FL	-45,430844*	10,6538594	,000	-77,889073	-12,972616
		EFL	-48,102815*	10,6405003	,000	-80,520343	-15,685286
	FL	Old Paradigm	45,430844*	10,6538594	,000	12,972616	77,889073
		FL					
		EFL	-2,671970	10,6538594	,969	-35,130199	29,786258
	EFL	Old Paradigm	48,102815*	10,6405003	,000	15,685286	80,520343
		FL	2,671970	10,6538594	,969	-29,786258	35,130199
		EFL					

*. The mean difference is significant at the .01 level.

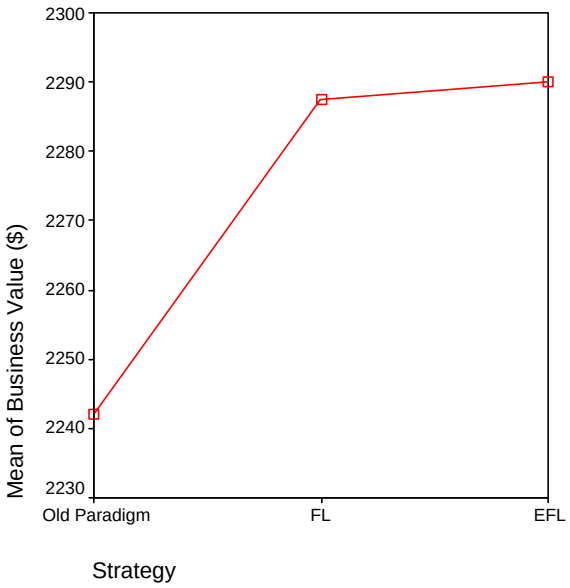


Figure 4-11 Business Performance of Experimentation Strategies one-way ANOVA results (n=200) (cont.)

4.5.2 Influence of concept exploration funnel shape on predictive, business performance and overall quality selection power

Overall quality selection power. A two-way between-groups analysis of variance was conducted to explore the impact of the funnel shape or level of parallelism within an early frontloaded (EFL) experimentation strategy and varying degrees of tightness of the surrogate marker chain on the overall quality selection power of the experimentation strategy used during Concept Selection. SPSS results are depicted in Figure 4-12.

1. CLASSES

Dependent Variable: QUAL

			99% Confidence Interval	
CLASSES	Mean	Std. Error	Lower Bound	Upper Bound
(1,1)	152,755	5,505	138,551	166,960
(5,1)	134,296	5,450	120,232	148,360
(5,2)	132,251	5,426	118,251	146,252
(5,3)	122,966	5,437	108,938	136,995

2. Tightness

Dependent Variable: QUAL

			99% Confidence Interval	
Tightness	Mean	Std. Error	Lower Bound	Upper Bound
80%	123,649	4,657	111,632	135,667
70%	133,668	4,657	121,652	145,684
48%	149,384	4,855	136,858	161,911

3. CLASSES * Tightness

Dependent Variable: QUAL

				99% Confidence Interval	
CLASSES	Tightness	Mean	Std. Error	Lower Bound	Upper Bound
(1,1)	80%	137,642	9,362	113,484	161,799
	70%	143,684	9,264	119,779	167,588
	48%	176,940	9,963	151,232	202,649
(5,1)	80%	117,166	9,169	93,507	140,825
	70%	144,030	9,412	119,743	168,317
	48%	141,692	9,732	116,582	166,802
(5,2)	80%	117,109	9,412	92,823	141,396
	70%	127,460	9,264	103,555	151,364
	48%	152,185	9,515	127,633	176,737
(5,3)	80%	122,680	9,313	98,650	146,710
	70%	119,499	9,313	95,469	143,528
	48%	126,720	9,622	101,893	151,546

Figure 4-12 Experimentation Strategy Quality Parallel Paths two-way ANOVA results (n=100)

Tests of Between-Subjects Effects

Dependent Variable: QUAL

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Corrected Model	301987,482[b]	11	27453,407	3,332	,000	,032	36,653	,975
Intercept	20357517,6	1	20357517,58	2470,844	,000	,692	2470,844	1,000
CLASSES	128092,145	3	42697,382	5,182	,001	,014	15,547	,800
TIGHTNES	121907,318	2	60953,659	7,398	,001	,013	14,796	,829
CLASSES * TIGHTNES	60853,500	6	10142,250	1,231	,288	,007	7,386	,261
Error	9046526,054	1098	8239,095					
Total	29587305,9	1110						
Corrected Total	9348513,536	1109						

a. Computed using alpha = ,01

b. R Squared = ,032 (Adjusted R Squared = ,023)

Multiple Comparisons

Dependent Variable: QUAL

	(I) Tightness	(J) Tightness	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval Lower Bound	99% Confidence Interval Upper Bound
Tukey HSD	80%	80%					
		70%	-10,027712	6,5851087	,281	-29,253338	9,197914
		48%	-25,376119*	6,7247378	,000	-45,009401	-5,742837
	70%	80%	10,027712	6,5851087	,281	-9,197914	29,253338
		70%					
		48%	-15,348407	6,7247378	,059	-34,981689	4,284875
	48%	80%	25,376119*	6,7247378	,000	5,742837	45,009401
		70%	15,348407	6,7247378	,059	-4,284875	34,981689
		48%					
Scheffe	80%	80%					
		70%	-10,027712	6,5851087	,314	-30,054540	9,999116
		48%	-25,376119*	6,7247378	,001	-45,827592	-4,924647
	70%	80%	10,027712	6,5851087	,314	-9,999116	30,054540
		70%					
		48%	-15,348407	6,7247378	,074	-35,799880	5,103065
	48%	80%	25,376119*	6,7247378	,001	4,924647	45,827592
		70%	15,348407	6,7247378	,074	-5,103065	35,799880
		48%					

Based on observed means.

*. The mean difference is significant at the ,01 level.

Figure 4-12 Experimentation Strategy Quality Parallel Paths two-way ANOVA results (n=100) (cont.)

Multiple Comparisons

Dependent Variable: QUAL

	(I) CLASSES	(J) CLASSES	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	(1,1)	(1,1)					
		(5,1)	17,885943	7,7341364	,096	-6,247854	42,019739
		(5,2)	19,656652	7,7204386	,054	-4,434401	43,747706
		(5,3)	28,828857*	7,7272660	,001	4,716499	52,941216
	(5,1)	(1,1)	-17,885943	7,7341364	,096	-42,019739	6,247854
		(5,1)					
		(5,2)	1,770710	7,6852046	,996	-22,210399	25,751818
		(5,3)	10,942915	7,6920633	,485	-13,059596	34,945425
	(5,2)	(1,1)	-19,656652	7,7204386	,054	-43,747706	4,434401
		(5,1)	-1,770710	7,6852046	,996	-25,751818	22,210399
		(5,2)					
		(5,3)	9,172205	7,6782904	,630	-14,787328	33,131739
	(5,3)	(1,1)	-28,828857*	7,7272660	,001	-52,941216	-4,716499
		(5,1)	-10,942915	7,6920633	,485	-34,945425	13,059596
		(5,2)	-9,172205	7,6782904	,630	-33,131739	14,787328
		(5,3)					
Scheffe	(1,1)	(1,1)					
		(5,1)	17,885943	7,7341364	,149	-8,225752	43,997637
		(5,2)	19,656652	7,7204386	,091	-6,408796	45,722101
		(5,3)	28,828857*	7,7272660	,003	2,740358	54,917357
	(5,1)	(1,1)	-17,885943	7,7341364	,149	-43,997637	8,225752
		(5,1)					
		(5,2)	1,770710	7,6852046	,997	-24,175783	27,717203
		(5,3)	10,942915	7,6920633	,568	-15,026734	36,912564
	(5,2)	(1,1)	-19,656652	7,7204386	,091	-45,722101	6,408796
		(5,1)	-1,770710	7,6852046	,997	-27,717203	24,175783
		(5,2)					
		(5,3)	9,172205	7,6782904	,699	-16,750945	35,095355
	(5,3)	(1,1)	-28,828857*	7,7272660	,003	-54,917357	-2,740358
		(5,1)	-10,942915	7,6920633	,568	-36,912564	15,026734
		(5,2)	-9,172205	7,6782904	,699	-35,095355	16,750945
		(5,3)					

Based on observed means.

*. The mean difference is significant at the ,01 level.

Figure 4-12 Experimentation Strategy Quality Parallel Paths two-way ANOVA results (n=100) (cont.)

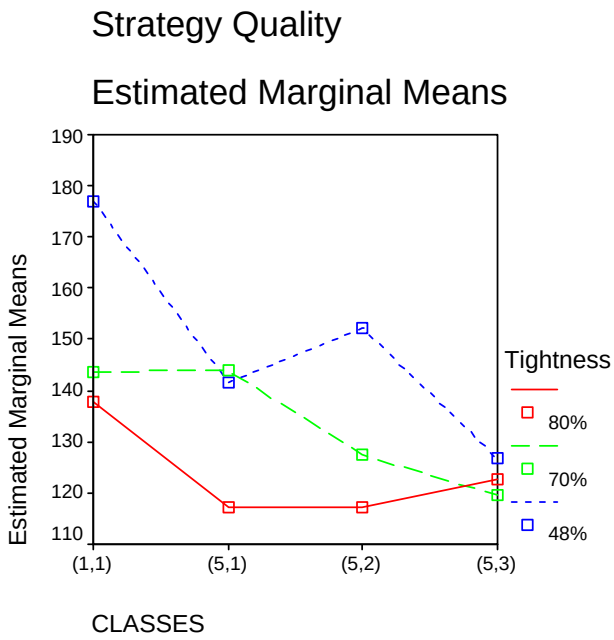


Figure 4-12 Experimentation Strategy Quality Parallel Paths two-way ANOVA results (n=100) (cont.)

Referring to Figure 4-12, for each funnel shaping strategy, modelled by the level of parallelism used as expressed by the number of classes evaluated during HTS and H2L-LO, three degrees of tightness were considered; 80%, 70%, and 48%. There was a statistically significant main effect for both levels of parallelism [F(3; 5.18), p= .001] and degrees of tightness [F(2; 7.39), p= .001]. However, the effect size was small (partial eta squared=.014 and .013 respectively). Post-hoc comparisons using Tukey HSD and Scheffe's tests both indicated that a (HTS; H2L-LO)=(1; 1) funnelling strategy for Early Front-loading (M=152.7; SE=5.5) performs significantly worse than a (5; 3) strategy (M=122.9; SE=5.4). There is no significant difference between a (5; 1), a (5; 2), and a (5; 3) funnelling strategy, meaning a minimum difference in the number of chemical classes assessed is necessary to increase significantly overall quality selection results of an experimentation strategy.

Surrogate marker chain tightness negatively influences overall quality selection power from a certain minimum level of tightness of the chain. Both Tukey HSD and Scheffe's tests indicate that 80% (M=123.6; SE=4.6) and 70% (M=133.7; SE=4.6) levels of chain tightness do not differ significantly. However, at 48% (M=149.4; SE=4.8) a statistically significant deterioration of performance is observed.

The interaction effect [F(6; 1.23), p= .29] did not reach statistical significance.

Summarizing, these simulation results indicate that increasing the number of classes assessed in HTS and H2L-LO in front-loaded strategies selects *overall* better candidate compounds. This is true for differing levels of surrogate marker chain tightness. Performance deteriorates significantly starting from a minimal level of surrogate marker chain tightness.

Predictive performance. A two-way between-groups analysis of variance was conducted to explore the impact of the funnel shape or level of parallelism within a frontloaded experimentation strategy and varying degrees of tightness of the surrogate marker chain on the predictive performance power of the experimentation strategy used during Concept Selection (see Figure 4-13).

1. CLASSES

Dependent Variable: pi+

			99% Confidence Interval	
CLASSES	Mean	Std. Error	Lower Bound	Upper Bound
(1,1)	,991	,001	,988	,995
(5,1)	,991	,001	,988	,994
(5,2)	,994	,001	,991	,998
(5,3)	,995	,001	,992	,999

2. Tightness

Dependent Variable: pi+

			99% Confidence Interval	
Tightness	Mean	Std. Error	Lower Bound	Upper Bound
80%	,996	,001	,993	,999
70%	,993	,001	,990	,996
48%	,990	,001	,987	,993

3. CLASSES * Tightness

Dependent Variable: pi+

				99% Confidence Interval	
CLASSES	Tightness	Mean	Std. Error	Lower Bound	Upper Bound
(1,1)	80%	,994	,002	,989	1,000
	70%	,993	,002	,987	,999
	48%	,987	,002	,980	,993
(5,1)	80%	,995	,002	,989	1,000
	70%	,990	,002	,984	,996
	48%	,989	,002	,982	,995
(5,2)	80%	,997	,002	,992	1,003
	70%	,994	,002	,988	1,000
	48%	,992	,002	,986	,998
(5,3)	80%	,997	,002	,991	1,002
	70%	,995	,002	,989	1,001
	48%	,994	,002	,988	1,000

Figure 4-13 Predictive performance Parallel Paths Strategies two-way ANOVA results (n=100)

Tests of Between-Subjects Effects

Dependent Variable: pi+

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Corrected Model	,010[b]	11	,001	1,957	,029	,019	21,526	,755
Intercept	1092,023	1	1092,023	2246711,766	,000	1,000	2246711,766	1,000
CLASSES	,004	3	,001	2,479	,060	,007	7,436	,377
TIGHTNES	,006	2	,003	6,124	,002	,011	12,247	,730
CLASSES * TIGHTNES	,001	6	,000	,413	,870	,002	2,481	,057
Error	,534	1098	,000					
Total	1095,078	1110						
Corrected Total	,544	1109						

a. Computed using alpha = ,01

b. R Squared = ,019 (Adjusted R Squared = ,009)

Multiple Comparisons

Dependent Variable: pi+

	(I) Tightness	(J) Tightness	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	80%	80%					
		70%	,002947	,0015994	,156	-,001722	,007617
		48%	,005618*	,0016333	,002	,000849	,010387
	70%	80%	-,002947	,0015994	,156	-,007617	,001722
		70%					
		48%	,002671	,0016333	,231	-,002098	,007439
	48%	80%	-,005618*	,0016333	,002	-,010387	-,000849
		70%	-,002671	,0016333	,231	-,007439	,002098
		48%					
Scheffe	80%	80%					
		70%	,002947	,0015994	,184	-,001917	,007812
		48%	,005618*	,0016333	,003	,000651	,010585
	70%	80%	-,002947	,0015994	,184	-,007812	,001917
		70%					
		48%	,002671	,0016333	,263	-,002297	,007638
	48%	80%	-,005618*	,0016333	,003	-,010585	-,000651
		70%	-,002671	,0016333	,263	-,007638	,002297
		48%					

Based on observed means.

*. The mean difference is significant at the ,01 level.

Figure 4-13 Predictive performance Parallel Paths Strategies two-way ANOVA results (n=100) (cont.)

1. CLASSES

Dependent Variable: pi-

CLASSES	Mean	Std. Error	99% Confidence Interval	
			Lower Bound	Upper Bound
(1,1)	,056	,001	,053	,058
(5,1)	,055	,001	,053	,058
(5,2)	,054	,001	,052	,057
(5,3)	,055	,001	,052	,058

2. Tightness

Dependent Variable: pi-

Tightness	Mean	Std. Error	99% Confidence Interval	
			Lower Bound	Upper Bound
80%	,056	,001	,054	,059
70%	,055	,001	,053	,058
48%	,054	,001	,051	,056

3. CLASSES * Tightness

Dependent Variable: pi-

CLASSES	Tightness	Mean	Std. Error	99% Confidence Interval	
				Lower Bound	Upper Bound
(1,1)	80%	,054	,002	,050	,059
	70%	,057	,002	,052	,062
	48%	,055	,002	,051	,060
(5,1)	80%	,060	,002	,055	,065
	70%	,054	,002	,049	,058
	48%	,053	,002	,048	,058
(5,2)	80%	,055	,002	,050	,059
	70%	,055	,002	,050	,060
	48%	,053	,002	,049	,058
(5,3)	80%	,057	,002	,052	,062
	70%	,056	,002	,051	,061
	48%	,053	,002	,048	,058

Figure 4-13 Predictive performance Parallel Paths Strategies two-way ANOVA results (n=100) (cont.)

Tests of Between-Subjects Effects

Dependent Variable: pi-

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Corrected Model	,004[b]	11	,000	1,185	,292	,011	13,035	,421
Intercept	3,648	1	3,648	10718,964	,000	,900	10718,964	1,000
CLASSES	,000	3	,000	,296	,828	,001	,888	,030
TIGHTNES	,002	2	,001	2,215	,110	,004	4,430	,232
CLASSES * TIGHTNES	,003	6	,000	1,286	,261	,006	7,717	,278
Error	,404	1188	,000					
Total	4,057	1200						
Corrected Total	,409	1199						

a. Computed using alpha = ,01

b. R Squared = ,011 (Adjusted R Squared = ,002)

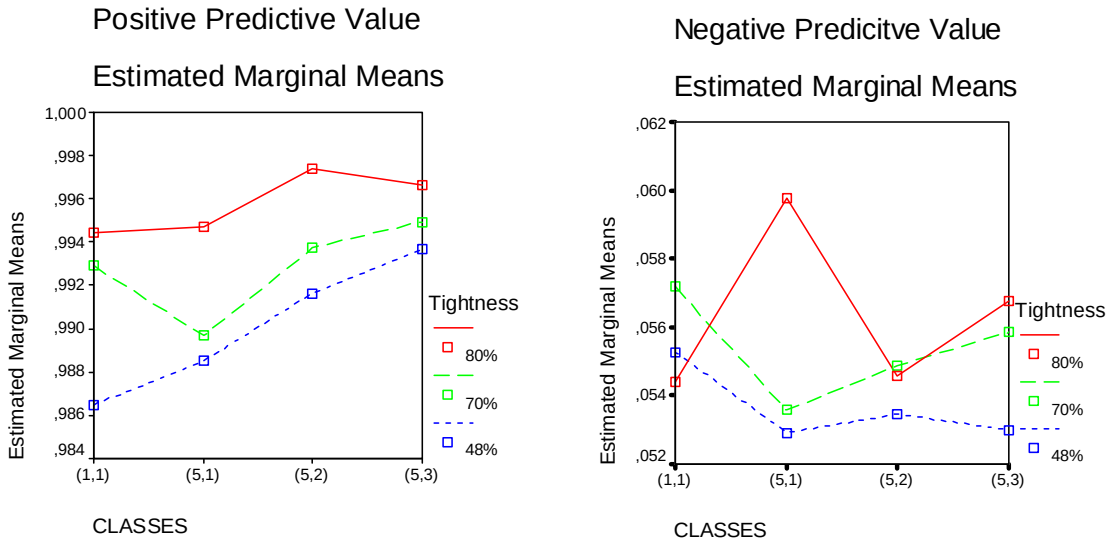


Figure 4-13 Predictive performance Parallel Paths Strategies two-way ANOVA results (n=100) (cont.)

Referring to Figure 4-13 above, for each funnel shaping strategy three degrees of tightness were considered; 80%, 70%, and 48%. During the screening experiment at n=100, only the tightness level [F(2; 6.12), p= .002] for positive predictive performance reached statistical significance. Number of classes [F(3; 2.48), p= 0.06], nor interaction effects [F(6; 0.413), p= 0.87] were significant for positive predictive performance. Number of classes [F(3; 0.29), p= .83], nor tightness level [F(2; 2.2), p= 0.11] or interaction effect [F(6; 1.29), p= 0.26] were significant for negative predictive performance.

Post-hoc comparisons using Tukey HSD and Scheffe's tests indicated tightness levels of 80% (M=0.996, SE=0.001) and 48% (M=0.99, SE=0.001) were significantly different impacting positive predictive performance across all front-loaded funnel shaping strategies.

Summarizing, the screening experiment indicates that positive predictive performance across all funnel shaping strategies is negatively impacted by falling levels of surrogate marker chain tightness. Negative predictive performance does not seem to

be impacted by the latter. However, due to the very small differences between the various simulation run descriptors, none of these effects indicates conclusive statistical significance. Therefore, since being conclusive about predictive performance of experimentation strategies is crucial for developing theory in my work, in the following one-way ANOVA confirmatory experiment the number of runs was elevated to n=200. Now, the solution landscape was fixed at wide and the surrogate marker chain tightness set to 70%. The latter tightness was preferred by the PharmaCo team since it represents best the capabilities of the present bio-chemical scientific reality. The choice for the wide landscape (SD reference compound in class; SD of compound around reference compound[46]) = (1; 1) was given by the limitations of the simulation method used. To the latter, the screening experiment had shown that very tight landscapes (0.1; 0.1) or very wide landscapes (4; 2) led to non-discriminating results between different experimentation strategies' performance variables.

So, finally a one-way between-groups analysis of variance was conducted in a confirmatory simulation experiment to probe for the impact of various frontloaded funnel shaping strategies on the business, positive and negative predictive performance of the experimentation strategy used during Concept Selection. SPSS results are depicted in Figure 4-14.

Descriptives

pi+

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	Between-Component Variance
						Lower Bound	Upper Bound			
(1,1)		173	,968315	,0402066	,0030569	,962281	,974348	,8255	1,0000	
(5,1)		200	,974378	,0249307	,0017629	,970901	,977854	,9243	1,0000	
(5,2)		200	,954662	,0367375	,0025977	,949539	,959785	,8469	1,0000	
(5,3)		200	,949688	,0295143	,0020870	,945573	,953803	,8650	,9805	
Total		773	,961532	,0345609	,0012431	,959091	,963972	,8255	1,0000	
Model	Fixed Effects			,0331224	,0011913	,959193	,963870			
	Random Effects				,0058301	,942978	,980086			,0001298

Test of Homogeneity of Variances

pi+

Levene Statistic	df1	df2	Sig.
8,260	3	769	,000

Figure 4-14 Predictive performance of Front-loaded Funnel shaping strategies one-way ANOVA results (n=200)

[46] see Table 4-2

ANOVA

pi+

			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		,078	3	,026	23,837	,000
	Linear Term	Unweighted	,053	1	,053	48,671	,000
		Weighted	,057	1	,057	51,617	,000
		Deviation	,022	2	,011	9,947	,000
Within Groups			,844	769	,001		
Total			,922	772			

Multiple Comparisons

Dependent Variable: pi+

			Mean			99% Confidence Interval	
	(I) CLASSES	(J) CLASSES	Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Tukey HSD	(1,1)	(1,1)					
		(5,1)	-,006063	,0034390	,292	-,016805	,004679
		(5,2)	,013653*	,0034390	,000	,002911	,024394
		(5,3)	,018627*	,0034390	,000	,007885	,029368
	(5,1)	(1,1)	,006063	,0034390	,292	-,004679	,016805
		(5,1)					
		(5,2)	,019716*	,0033122	,000	,009370	,030061
		(5,3)	,024689*	,0033122	,000	,014344	,035035
	(5,2)	(1,1)	-,013653*	,0034390	,000	-,024394	-,002911
		(5,1)	-,019716*	,0033122	,000	-,030061	-,009370
		(5,2)					
		(5,3)	,004974	,0033122	,437	-,005372	,015320
	(5,3)	(1,1)	-,018627*	,0034390	,000	-,029368	-,007885
		(5,1)	-,024689*	,0033122	,000	-,035035	-,014344
		(5,2)	-,004974	,0033122	,437	-,015320	,005372
		(5,3)					
Scheffe	(1,1)	(1,1)					
		(5,1)	-,006063	,0034390	,376	-,017685	,005560
		(5,2)	,013653*	,0034390	,001	,002030	,025275
		(5,3)	,018627*	,0034390	,000	,007004	,030249
	(5,1)	(1,1)	,006063	,0034390	,376	-,005560	,017685
		(5,1)					
		(5,2)	,019716*	,0033122	,000	,008522	,030909
		(5,3)	,024689*	,0033122	,000	,013496	,035883
	(5,2)	(1,1)	-,013653*	,0034390	,001	-,025275	-,002030
		(5,1)	-,019716*	,0033122	,000	-,030909	-,008522
		(5,2)					
		(5,3)	,004974	,0033122	,522	-,006220	,016168
	(5,3)	(1,1)	-,018627*	,0034390	,000	-,030249	-,007004
		(5,1)	-,024689*	,0033122	,000	-,035883	-,013496
		(5,2)	-,004974	,0033122	,522	-,016168	,006220
		(5,3)					

*. The mean difference is significant at the .01 level.

Figure 4-14 Predictive performance of Front-loaded Funnel shaping strategies one-way ANOVA results (n=200) (cont.)

Descriptives

pi-

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	Between-Component Variance
						Lower Bound	Upper Bound			
(1,1)		200	,126785	,0180144	,0012738	,124273	,129297	,0932	,1642	
(5,1)		200	,125469	,0252276	,0017839	,121951	,128986	,0911	,2140	
(5,2)		200	,118608	,0176580	,0012486	,116146	,121071	,0836	,1581	
(5,3)		200	,111824	,0215595	,0015245	,108818	,114830	,0635	,1599	
Total		800	,120671	,0216457	,0007653	,119169	,122174	,0635	,2140	
Model	Fixed Effects			,0208421	,0007369	,119225	,122118			
	Random Effects				,0034511	,109688	,131654			,0000455

Test of Homogeneity of Variances

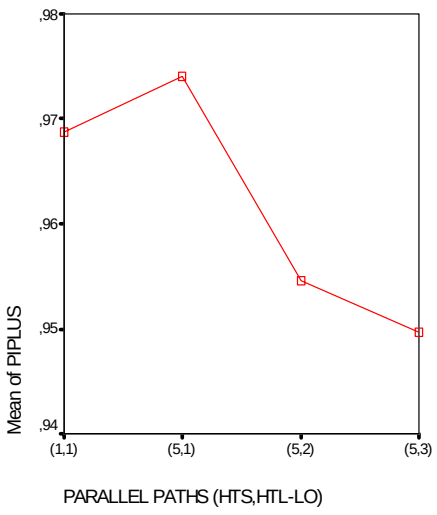
pi-

Levene Statistic	df1	df2	Sig.
5,602	3	796	,001

ANOVA

pi-

			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		,029	3	,010	21,935	,000
	Linear Term	Contrast	,027	1	,027	61,635	,000
		Deviation	,002	2	,001	2,084	,125
Within Groups			,346	796	,000		
Total			,374	799			



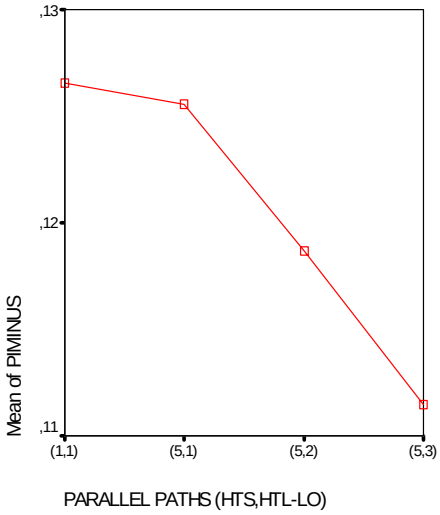


Figure 4-14 Predictive performance of Front-loaded Funnel shaping strategies one-way ANOVA results (n=200) (cont.)

Multiple Comparisons

Dependent Variable: pi-

	(I) CLASSES	(J) CLASSES	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval Lower Bound	99% Confidence Interval Upper Bound
Tukey HSD	(1,1)	(1,1)					
		(5,1)	,001316	,0020842	,922	-,005193	,007826
		(5,2)	,008176*	,0020842	,001	,001667	,014686
		(5,3)	,014961*	,0020842	,000	,008452	,021470
	(5,1)	(1,1)	-,001316	,0020842	,922	-,007826	,005193
		(5,1)					
		(5,2)	,006860*	,0020842	,006	,000351	,013369
		(5,3)	,013645*	,0020842	,000	,007135	,020154
	(5,2)	(1,1)	-,008176*	,0020842	,001	-,014686	-,001667
		(5,1)	-,006860*	,0020842	,006	-,013369	-,000351
		(5,2)					
		(5,3)	,006785*	,0020842	,006	,000275	,013294
	(5,3)	(1,1)	-,014961*	,0020842	,000	-,021470	-,008452
		(5,1)	-,013645*	,0020842	,000	-,020154	-,007135
		(5,2)	-,006785*	,0020842	,006	-,013294	-,000275
		(5,3)					
Scheffe	(1,1)	(1,1)					
		(5,1)	,001316	,0020842	,940	-,005727	,008359
		(5,2)	,008176*	,0020842	,002	,001134	,015219
		(5,3)	,014961*	,0020842	,000	,007918	,022004
	(5,1)	(1,1)	-,001316	,0020842	,940	-,008359	,005727
		(5,1)					
		(5,2)	,006860	,0020842	,013	-,000183	,013903
		(5,3)	,013645*	,0020842	,000	,006602	,020688
	(5,2)	(1,1)	-,008176*	,0020842	,002	-,015219	-,001134
		(5,1)	-,006860	,0020842	,013	-,013903	,000183
		(5,2)					
		(5,3)	,006785	,0020842	,015	-,000258	,013828
	(5,3)	(1,1)	-,014961*	,0020842	,000	-,022004	-,007918
		(5,1)	-,013645*	,0020842	,000	-,020688	-,006602
		(5,2)	-,006785	,0020842	,015	-,013828	,000258
		(5,3)					

*. The mean difference is significant at the .01 level.

Figure 4-14 Predictive performance of Front-loaded Funnel shaping strategies one-way ANOVA results (n=200) (cont.)

Referring to Figure 4-14, now positive [$F(3, 769)=23{,}84$; $p<.01$] and negative [$F(3, 796)=21.93$; $p<.01$] predictive performance for the various front-loaded funnel shaping strategies differed significantly although the actual differences in mean scores are quite small. Effect size, calculated using eta squared, was 0.084 and 0.077 for positive and negative predictive performance respectively indicating a medium impact of the independent variables on the outcome variables. Levene's test showed non-

homogeneity of variance across the three groups –which I consider to be normal considering the different experimentation strategy types- implying the need for studying differences between groups at the p<.01 level of significance.

Both Tukey's HSD and Scheffe's post-hoc comparisons for positive predictive performance indicated that a (1; 1) and a (5; 1) funnel shaping strategy both differ significantly from a (5; 2) and (5; 3) strategy at the p<.01 level of significance. A (5; 2) and a (5; 3) strategy do not differ significantly from each other.

Both Tukey's HSD and Scheffe's post-hoc comparisons for negative predictive performance indicated that a (1; 1) and a (5; 1) funnel shaping strategy both differ significantly from a (5; 2) strategy and (5; 3) strategy at the p<.01 level of significance. Also, (5; 2) and (5; 3) strategies differ significantly from each other at the p<.01 level of significance.

Business performance. A one-way between-group analysis of variance was conducted in a confirmatory simulation experiment to probe for the impact of various frontloaded funnel shaping strategies on the business performance of the experimentation strategy used during Concept Selection (see for results). There was a statistically significant difference at the p<.01 level for the different funnel shaping strategies. Financial assumptions used to perform ANOVA are discussed above.

Descriptives

Business Value ($)

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	Between-Component Variance
						Lower Bound	Upper Bound			
(1,1)		173	2134,807781	99,1509775	7,5383092	2119,928273	2149,687289	1803,6454	2259,3901	
(5,1)		200	2157,477173	75,6095024	5,3463992	2146,934306	2168,020040	1978,6831	2264,1233	
(5,2)		200	2120,860256	91,4701157	6,4679139	2108,105811	2133,614701	1824,7294	2263,8506	
(5,3)		200	2121,273133	79,2522790	5,6039824	2110,222324	2132,323943	1881,0194	2238,9648	
Total		773	2133,562560	87,5736503	3,1498053	2127,379361	2139,745758	1803,6454	2264,1233	
Model	Fixed Effects			86,4218164	3,1083767	2127,460649	2139,664470			
	Random Effects				8,7582073	2105,690035	2161,435084			267,1987929

Test of Homogeneity of Variances

Business Value ($)

Levene Statistic	df1	df2	Sig.
2,030	3	769	,108

Figure 4-15 Business Performance of Front-loaded Funnel shaping strategies one-way ANOVA results (n=200)

ANOVA

Business Value ($)

			Sum of Squares	df	Mean Square	F	Sig.
Between Groups	(Combined)		177125,699	3	59041,900	7,905	,000
	Linear Term	Unweighted	55717,506	1	55717,506	7,460	,006
		Weighted	62802,819	1	62802,819	8,409	,004
		Deviation	114322,880	2	57161,440	7,653	,001
Within Groups			5743453,639	769	7468,730		
Total			5920579,338	772			

Multiple Comparisons

Dependent Variable: Business Value ($)

						99% Confidence Interval	
	(I) CLASSES	(J) CLASSES	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Tukey HSD	(1,1)	(1,1)					
		(5,1)	-22,669391	8,9730432	,057	-50,696721	5,357938
		(5,2)	13,947526	8,9730432	,406	-14,079804	41,974855
		(5,3)	13,534648	8,9730432	,433	-14,492681	41,561977
	(5,1)	(1,1)	22,669391	8,9730432	,057	-5,357938	50,696721
		(5,1)					
		(5,2)	36,616917*	8,6421816	,000	9,623035	63,610799
		(5,3)	36,204039*	8,6421816	,000	9,210157	63,197922
	(5,2)	(1,1)	-13,947526	8,9730432	,406	-41,974855	14,079804
		(5,1)	-36,616917*	8,6421816	,000	-63,610799	-9,623035
		(5,2)					
		(5,3)	-,412878	8,6421816	1,000	-27,406760	26,581005
	(5,3)	(1,1)	-13,534648	8,9730432	,433	-41,561977	14,492681
		(5,1)	-36,204039*	8,6421816	,000	-63,197922	-9,210157
		(5,2)	,412878	8,6421816	1,000	-26,581005	27,406760
		(5,3)					
Scheffe	(1,1)	(1,1)					
		(5,1)	-22,669391	8,9730432	,095	-52,994422	7,655639
		(5,2)	13,947526	8,9730432	,491	-16,377505	44,272556
		(5,3)	13,534648	8,9730432	,518	-16,790383	43,859679
	(5,1)	(1,1)	22,669391	8,9730432	,095	-7,655639	52,994422
		(5,1)					
		(5,2)	36,616917*	8,6421816	,000	7,410056	65,823778
		(5,3)	36,204039*	8,6421816	,001	6,997178	65,410901
	(5,2)	(1,1)	-13,947526	8,9730432	,491	-44,272556	16,377505
		(5,1)	-36,616917*	8,6421816	,000	-65,823778	-7,410056
		(5,2)					
		(5,3)	-,412878	8,6421816	1,000	-29,619739	28,793983
	(5,3)	(1,1)	-13,534648	8,9730432	,518	-43,859679	16,790383
		(5,1)	-36,204039*	8,6421816	,001	-65,410901	-6,997178
		(5,2)	,412878	8,6421816	1,000	-28,793983	29,619739
		(5,3)					

*. The mean difference is significant at the .01 level.

Figure 4-15 Business Performance of Front-loaded Funnel shaping strategies one-way ANOVA results (n=200) (cont.)

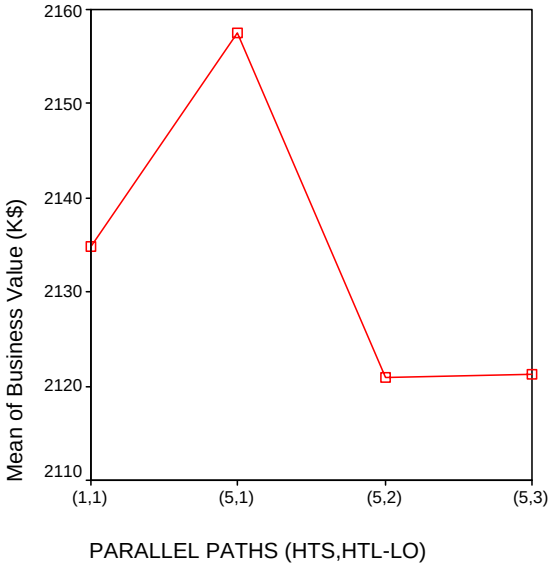


Figure 4-15 Business Performance of Front-loaded Funnel shaping strategies one-way ANOVA results (n=200) (cont.)

Referring to Figure 4-15 business performance [F(3, 769)=7.9; p<.01] for the various front-loaded funnel shaping strategies differed significantly. This result needs to be moderated by the fact that the effect size, calculated using eta squared, was 0.031 indicating a small to medium impact of the independent variables on the outcome variable. Levene's test showed homogeneity of variance across the three groups. Still, differences between groups were studied at the p<.01 level of significance.

Both Tukey HSD and Scheffe's post-hoc comparisons for business performance indicated that a (5; 1) funnel shaping strategy (M=2157.5, SD=75.6) outperforms a (1; 1) strategy (M=2134.8, SD=99.1) by $22M and both a (5; 2) strategy (M=2120.9, SD=91.5) and (5; 3) strategy (M=2121.3, SD=79.3) by $36M, at the p<.01 level of significance. A (1; 1) strategy does not differ significantly from the others and a (5; 2) and (5; 3) strategy do not differ significantly from each other.

4.6 DISCUSSION

4.6.1 Key Monte Carlo simulation study findings

Impact of front-loaded experimentation strategies. Both Front-loaded strategies –Front-loaded paradigm and Early Front-loading- significantly outperform the Old Paradigm experimentation strategy on positive predictive performance. However, the difference between both Front-loaded strategies is insignificant, meaning in the context of pharmaceutical Discovery that in-silico characterization of ADME-T does not significantly increase positive predictive value as compared to ADME-T characterization starting in H2L, as done in Front-loaded paradigm. Second, negative predictive performance differences are insignificant across experimentation strategies.

These results lead me to formulate the following propositions relating Concept Selection experimentation strategy to its predictive performance, supporting conjecture C2-1 formulated in the previous Chapter under section 3.6.1:

> *Proposition 4-1: There is an inverse relationship between the chosen level of residual ambiguity at the end of Concept Selection and its positive predictive value.*

Proposition 4-1 proposes that the more proof of concept critical variables are characterized during Concept Selection, the higher will be its positive predictive value. Or, in terms of conditional probabilities; given they were found positive during Concept Selection, the probability that solution concepts will be successfully introduced into the market will be higher if more proof of concept variables are characterized during Concept Selection.

Conversely, since simulation results show no significant difference between Front-loaded paradigm and Front-loaded experimentation strategies, conjecture C2-2 formulated in the previous Chapter proposing that the sooner proof of concept critical variables are characterized during Concept Selection, the higher will be its negative predictive value, is not supported.

In contrast to some practitioner views, my simulation results show non-significant differences in negative predictive performance for the various strategies. This indicates that front-loaded strategies do not decrease the probability of missed opportunities as compared to an Old Paradigm strategy. They only increase its positive predictive value, increasing the chances of surviving Concept Characterization and Concept Application testing once declared active at the end of Concept Selection.

Furthermore, simulation results indicate that front-loaded experimentation strategies significantly select *overall* better quality compounds than an Old Paradigm strategy. In other words, taking into consideration all solution variables necessary for delivering Proof of Concept, leads to better solutions selected during Concept Selection. This leads me to formulate the following proposition:

> *Proposition 4-2: Front-loaded experimentation strategies select overall better quality concepts than non-front-loaded strategies during Concept Selection.*

Also, simulation results indicate that front-loaded strategies applied during Concept Selection lead to higher business value than Old Paradigm strategies. Although not statistically significant, applying front-loading earlier did lead to more business value. Hence, the following proposition:

> *Proposition 4-3: Front-loaded experimentation strategies lead to higher business value than non-front-loaded strategies.*

Finally, my simulation results provide no support for conjecture C2-3 formulated in the previous Chapter proposing that a minimum tightness level is required

for front-loading to outperform Old Paradigm strategies. Instead, front-loaded strategies consistently outperformed Old Paradigm strategies on quality selection power and predictive performance, regardless of the tightness level of the surrogate marker chain. However, a minimum tightness level was found to exist below which positive predictive performance significantly deteriorated. Negative predictive performance was not influenced by the chain's tightness level.

> *Proposition 4-4: Front-loaded experimentation strategies used during Concept Selection will feature better overall quality selection power, and higher positive and negative predictive value than old paradigm strategies regardless of the tightness level of the predictive surrogate marker chain.*

Impact of parallelism. Simulation results indicate significant positive effects of broadening the solution concept funnel on quality selection power, predictive and business performance of Concept Selection experimentation strategies.

> *Proposition 4-5: Broadening the funnel in a Front-loaded experimentation strategy during Concept Selection increases its negative predictive power, significantly decreasing the chances of missed opportunities in subsequent development. A minimum number of parallel concept explorations are required to gain effect.*

> *Proposition 4-6: Broadening the funnel in a Front-loaded experimentation strategy during Concept Selection selects better overall quality concepts. A minimum number of parallel concept explorations are required to gain effect.*

> *Proposition 4-7: Broadening the funnel to an optimum point in a Front-loaded experimentation strategy during Concept Selection leads to optimal business performance.*

However, simulation results did show that broadening the concept exploration funnel has a significantly deteriorating effect on positive predictive performance during Concept Selection.

> *Proposition 4-8: Broadening the funnel in a Front-loaded experimentation strategy during Concept Selection decreases its positive predictive power, significantly decreasing the chances of surviving Concept Characterization and Concept Application testing once declared active at the end of Concept Selection.*

Finally, simulation results indicate that no minimum tightness level is required for broader funnel shaping strategies to outperform leaner strategies. However, a minimum tightness level was found to exist below which positive predictive performance significantly deteriorated. Negative predictive performance was not influenced by the chain's tightness level.

Proposition 4-9: Broader funnel shaping strategies used during Concept Selection will feature better overall quality selection power, and higher positive and negative predictive value than old paradigm strategies regardless of the tightness level of the predictive surrogate marker chain.

4.6.2 Validity considerations and study limitations

Computer simulation outperforms other research methods with respect to reliability. However, the key caveat in simulation-based research is validity. Validation is assessing whether a specific simulation model is an acceptable representation of the corresponding real system, given the goal of the simulation model (Kleijnen et al. 2001). Models for simulation purposes cannot be shown to be true or valid in any absolute sense. What can be said is that the model can be valid under certain assumptions. In this study key assumptions were formulated to reduce the complexity of representing the innovation process without endangering the fulfilment of the research goal of theory development. In summary, three simplifying assumptions were made concerning the representation of the solution landscape, and for optimization and selection conducted during the innovation process.

First, the solution landscape was represented using three compound properties aggregated into reference compounds and chemical classes. This oversimplification of reality was necessary to make the implementation of the conceptual model possible in a VBA environment. Second, the innovation environment was represented as a mechanistic process of optimization stages concluded by a number of decision gates where candidate solution concepts were promoted to the next stage or terminated if they did not fulfil the selection criteria. Respecting the garbage can philosophy, the complex scientist optimization behaviour was conservatively reduced to a simple multi-factorial function taking the minimum of the (P) and (B) values as an input to a search for maximum performance in an extant search space. White box validation (Pidd, 1992) of this process with PharmaCo scientists, showing that the model behaves in a reasonable fashion, depicting a familiar universe of organizing the discovery research process confirmed the face validity of this complexity reduction of reality, provided it served the purpose of theory development.

However, considering the complexities of the probabilistic modelling of the innovation process and the hard-to-unravel nature of the model's inner working, face validation was insufficient to claim internal validity of the simulation model. Therefore, validation ex negatio (Masuch and Lapotin, 1989) of key assumptions was done, showing that these assumptions do little harm to the model's predictive power. This is why comparative performance conclusions about experimentation strategies were checked for robustness by varying the most basic parameters in the model. Thus, simulation results showed to be robust for changes in solution landscape ruggedness up to a certain level, down to a certain level for varying numbers of the set of compounds declared active in LO, and for changes in the marginal probability p(H).

Finally, external validity must be gained through empirical observation of the model's predictions. 'The behaviour of the "real" system is observed under specified

conditions and the model is then run under conditions which are as close as possible to these. If the model is valid in a black box sense, then the observations of the model should be indistinguishable from those of the "real" system' (Pidd, 1992: 106). Recent experience at PharmaCo[47] suggests external validity of the findings that Front-loading outperforms other experimentation strategies on positive predictive performance. Further longitudinal research data on compounds declared active by Discovery and succeeding Clinical Development at PharmaCo and other pharmaceutical companies, would be required to gain sufficient empirical support for the model's predictions. It should be noted at this place that the model's predictions on negative predictive performance are not verifiable in practice since the key conditional probability involved is not observable in practice.

4.7 CONCLUSION

The memory less annotated adaptive systems model of PharmaCo's discovery research process produced reliable, internally valid results that could be used for theory generation. A theoretical contribution was made quantifying predictive performance of alternative experimentation strategies for Concept selection in the specific context of Pharmaceutical Discovery.

More specifically, simulation results indicated that Front-loaded strategies in this context outperform other strategies on positive predictive performance, irrespective of the tightness of the surrogate marker chain. The number of classes used influences significantly the positive and negative predictive performance of experimentation strategies. This practically means that conducting parallel explorations of concepts during Concept Selection in a pharmaceutical context significantly reduces the probability of missed opportunities in Concept Characterisation. These results were shown to be robust for varying levels of tightness of the surrogate marker chain.

Finally, as computer simulation was used as a technique of theorizing, further empirical validation of results is necessary to gain sufficient support for the model's predictions. Most probably, falsification attempts in other pharmaceutical or research-intensive contexts will lead to modifications of the model's present version which only underlines the purpose of the model as a suitable adaptive theory generator.

[47] As mentioned in the previous project confirmatory case study an internal PharmaCo study revealed that in the 1995-2000 period about 50 compounds failed in pre-clinical and clinical development programs due to poor 'drug-likeness' of NME's, meaning they showed too low performance on PK/PD, toxicology, or could not be suitably packed in a drug delivery vehicle. The same study reveals that, in retrospect, the present discovery process, taking into consideration both biological activity and pharmacokinetic property classes, would have 'caught' about 20 compounds being unrightfully promoted to NME status.

5 Towards a theory explaining experimentation behaviour in the fuzzy front-end innovation process

5.1 INTRODUCTION

This journey exploring ways to increase predictive performance of the fuzzy front-end innovation process was conducted in three research projects. The previous chapters discussed their results. Starting from the radical versus incremental dichotomy proposed in the innovation literature, and grounded in exploratory case studies in pharmaceutical Development, the purpose of the first research project was to develop a more fine-grained understanding of how and why radical innovation experimentation differs from incremental innovation management practice. The second project, conducted in a pharmaceutical Discovery radical innovation context, literally replicated the findings of the first project. Furthermore, it developed the Bayesian perspective and empirical process grounding required for the last, Monte Carlo computer simulation-based project. The latter resulted in a propositional model explaining the experimentation process-predictive performance link.

The first research project, described in Chapter 2, concluded that incremental or radical innovations are still intuitive, ill-defined concepts, making them unsuitable as constructs to develop this required deeper understanding of managing experimentation for performance in the context of innovation projects. Therefore, I suggested shifting paradigms and focusing on the collaborative problem solving behaviour conducted by the innovation team. An Interpretative view of Complexity Theory was used as the theoretical paradigm to focus on the problem solving done by the innovation team, the latter making sense of the *experienced complexity* they are facing in their –be it radical or incremental- innovation projects. Now, by focusing on the complexity experienced by the innovation team and its ensuing mental modelling process of the solution to the innovation problem, my case study results made me propose a prescriptive framework proposing experienced complexity, characterized into types by specifying levels of ambiguity and uncertainty facing the innovation team, to drive their choice for a specific experimentation and project management approach, I called *complexity-handling mode*. Then, regardless of the type of innovation -radical or incremental- they are working on, but contingent upon the type of experienced complexity, they were proposed to choose between three modes of complexity-handling experimentation behaviour; *Concept Selection, Concept Characterisation, and Concept Application*.

The second research project, described in Chapter 3, was focused around the research questions whether the pharmaceutical Discovery process could be used for literal replication of the proposed model on complexity-handling, and on how it could provide the process grounding for the subsequent simulation-based study. This confirmatory case study analysis successfully replicated part of the findings pertaining to my proposed model relating experienced complexity to choices for a specific complexity-handling mode. More specifically, the Concept Selection mode and its transition to the subsequent Concept Characterisation complexity-handling mode could

be replicated to the pharmaceutical Discovery context. However, the pharmaceutical Discovery case provided no evidence to corroborate the Concept Characterisation and Concept Application parts of the model. In addition, this case study documented alternative experimentation strategies used in pharmaceutical Discovery, all specific applications of the Concept Selection complexity-handling mode. Various forms of front-loaded experimentation strategies were found to be used by Discovery management. All manage the build-up of a mental model of the innovative solution to a level of *residual ambiguity* the innovation team feels comfortable with to start Concept Characterization. Then, a rationale and Bayesian methodology was proposed to evaluate predictive performance of front-loaded experimentation strategies using Monte Carlo simulation. Finally, research conjectures were formulated to guide subsequent process simulations, linking these Concept Selection experimentation strategies to predictive and business performance.

The third research project, described in Chapter 4, discussed the results of top-down simulation-based theory development on predictive performance of front-loaded experimentation strategies. Following a review of theoretical models representing the complexity of dynamic experimentation and decision-making processes, it was argued that an annotated *adaptive system paradigm* is the best choice to emulate Concept Selection experimentation behaviour. Simulation results indicated that front-loaded strategies in a pharmaceutical Discovery context slightly but significantly outperform other strategies on positive predictive performance. The degree of parallelism used to explore the solution space influences significantly the negative predictive performance of experimentation strategies. The latter means that conducting parallel explorations of solution concepts during Concept Selection, in a pharmaceutical Discovery context, significantly reduces the probability of missed opportunities in subsequent Concept Characterisation or Concept Application. Finally, although taking into account a set of simplifying assumptions, from a business value standpoint, simulation results indicate that front-loaded experimentation strategies outperform other strategies.

In this final Chapter formal teleological process theory will be developed, inductively grounded in the research results described above. The formal process theory will integrate the case study-based empirical work with the simulation-based theorizing effort. Process grounding conducted in the first and second research projects, probing into the *meaning* of the innovation process for the team, will lead to a prescriptive framework using the three complexity-handling modes described above, holistically explaining the *mechanisms* used by the team as a socio-technical system handling the complexity of its mission to solve the innovation problem. Conversely, computer-based simulation, prepared for in the second project and conducted in the third research project will lead to a propositional model, explaining the impact of front-loaded experimentation strategies on the predictive performance of Concept Selection, being the complexity-handling mode best describing the work in the fuzzy-front end innovation process as carried out in pharmaceutical Discovery.

Congruent with the requirements of mode 2 research (Gibbons et al. 1994) the constructed formal theory should stand the test of reality by adequately describing innovation team's complexity-handling behaviour and by explaining its relationship to predictive performance, both in such a way that practitioners active in pharmaceutical or

other technology-intensive R&D sectors should be able to recognize themselves and others in the foreign concepts of the proposed formal theory explaining their experimentation behaviour.

In the remainder of this Chapter I will first construct formal process theory relating experimentation behaviour to predictive performance. This will be followed by a discussion of the contribution of the developed theory to the innovation and experimentation management body of knowledge, and of its implications for R&D management practice. Finally, study limitations and areas for further research of this theory development effort will be indicated.

5.2 DEVELOPING FORMAL PROCESS THEORY

This thesis resulted in a formal teleological process theory relating experimentation behavior conducted in the fuzzy front-end innovation process to predictive performance, an outcome parameter neglected in previous innovation process research conducted in pharmaceutical or other R&D contexts. By formal theory I mean theory developed for the conceptual area of experimentation behavior, as opposed to substantive being developed for a specific empirical area of inquiry, such as experimentation conducted in pharmaceutical Discovery. The formal theory must be general enough to be applicable to a number of substantive areas. Also, it should complement and guide empirical work to understand experimentation behaviour in radical innovation settings.

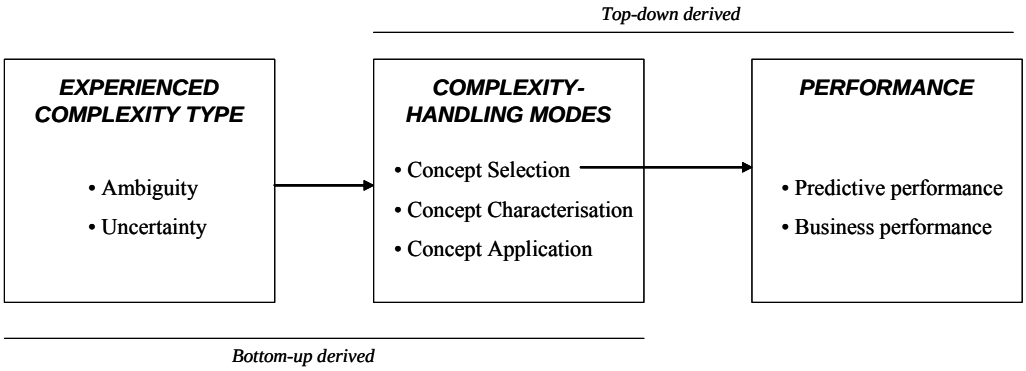


Figure 5-1 High level teleological process theory representation

The end result is a *bottom-up top-down* developed formal process theory, consisting of two parts. As depicted in the high level representation of the teleological process theory in Figure 5-1 above, experienced complexity type is proposed to lead to a preferred complexity-handling mode whose experimentation strategy is proposed to lead to predictive and business performance. The proposed relationship between the two first concepts is the bottom-up grounded part of the theory. It is an inductively derived prescriptive framework explaining complexity-handling contingent upon the type of complexity experienced by the innovation team. The proposed relationship between the two last concepts is the top-down simulation-derived part of the theory. The scope of the latter is confined to the experimentation strategy carried out during Concept

Selection –only one of the complexity-handling modes- and its proposed simulation-derived relationship to predictive and business performance.

In the following I will fill in the prescriptive framework and propositional model that constitute the developed theory before discussing its contribution to knowledge and practice, and its limitations and potential areas for further research.

5.2.1 Prescriptive framework explaining complexity-handling behaviour

As discussed in Chapter 2 and 3, the results of my first two research project lead me to propose that experienced complexity type explains the complexity-handling mode dynamically chosen by the innovation team through the course of their project. First, I will define and describe the complexity-handling modes emerging from the case studies conducted in pharmaceutical Discovery and Development. Then, I will propose a prescriptive framework in which experienced complexity dynamically explains the complexity-handling modes followed by the innovation team.

Complexity-handling mode definitions and descriptions. A visual mapping and temporal bracketing analysis (Langley, 1999) of the experimentation process conducted in the seven case studies of my first two research projects led me to define three complexity-handling modes emerging as three brackets from the case data; Concept Selection, Concept Characterisation, and Concept Application. The chronological overviews of project phases depicted in Figure 5-2 are further detailed in the empirical results sections of Chapters 2 and 3. In all cases innovation problem definition preceded these identified problem solving phases.

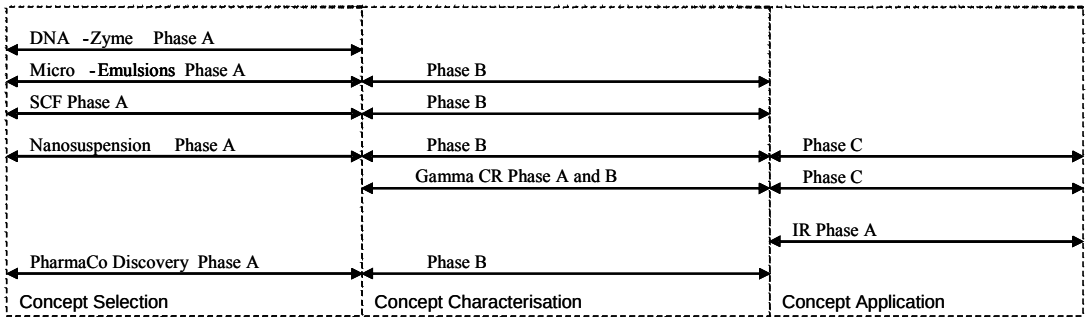


Figure 5-2 Temporal brackets identified across all PharmaCo innovation project cases

Concept Selection is defined as a complexity-handling mode leading to a mental model characterising the innovative system's core, depicting the critical variables and their relationships affecting solution proof of concept level performance. It starts with generating candidate solutions for the innovation problem at hand. Then, working towards meeting solution critical requirements it gradually resolves ambiguity and brings focus by ruling out as soon as possible in the process candidate solutions that don't work, and by organising work primarily around the key questions to be solved,

operationalised in the solution critical requirements. The latter are defined at proof of concept level. Hence, following Boisot and Child's (Boisot and Child, 1999) typology, experienced complexity is predominantly absorbed by generating and selecting solution options.

Typical Concept Selection experimentation process examples include, amongst others, Phase A of the DNA-Zyme (Figure 2-2), Micro-emulsion (Figure 2-3), and the Discovery process (Figure 3-4) cases. In all these cases solution critical requirements were specified as a target against which alternative solutions were tested. Then, several competing solutions were tried in parallel to match the target, managed through a stage-gate process introducing continuously raising hurdles for the competing candidate solutions; delivery systems for macromolecular therapeutics for DNA-Zyme, various candidate polymers to be used as micro-emulsions enhancing bioavailability of therapeutic agents, and different chemical classes coming out of High Throughput Screening, serving as chemical scaffolds, to be further screened and optimized during the pharmaceutical Discovery process.

Concept Characterisation is defined as a complexity-handling mode leading to a mental model characterising the full innovative solution's application domain, depicting all relevant variables and their relationships affecting application system performance. Starting from proof of concept the full application domain gets characterised through the definition of uncertainty areas in which a more structured adaptive learning process gradually resolves ambiguity beyond proof of concept level down to all variables and their functional relationships affecting solution performance. Also, critical value ranges get defined for all variables to indicate the boundaries of the application domain, as compared to other drug delivery technologies. Several application domains can be characterized sequentially or in parallel as evidenced by the case examples below. Hence, following Boisot and Child's (1999) typology, experienced complexity is both absorbed and reduced. Experienced complexity is absorbed by the creation of application domains and uncertainty areas. Experienced complexity is reduced by the definition of critical value ranges.

Concept Characterisation experimentation process examples include, amongst others, Phase B of the Micro-emulsion (Figure 2-3), Supercritical Fluids (Figure 2-4), Nanosuspension (Figure 2-5), and the Discovery process (Figure 3-4) cases. In the first case, after a selection of polymers with good micro-emulsifying properties was made, parenteral and oral/transdermal formulations domains were characterised for further usage of the advanced technology. Once the proof of concept of the use of Supercritical Fluids was delivered to extend the meltextrusion manufacturing process to include thermosensitive elements like peptides or proteins, two problem domains were characterized, preparing the renewed meltextrusion process for its introduction into manufacturing. Also, as a whole this was only the first application domain for the technology delivered at proof of concept level, since afterwards it has been characterized as a platform technology for applications as diverse as micronization, human tissue replacement, or even as a technology accelerating the joining of human bones. Another example is provided by the Nanosuspension project where the initial proof of concept had been delivered before. However, now it had to be used to deliver a new clinical candidate to market. Therefore, first proof of Concept needed to be

delivered for this specific clinical candidate using a Concept Selection complexity-handling mode where a range of candidate formulations was tested and one was ultimately chosen (during Phase A in Figure 2-5). Then, two uncertainty domains needed to be characterized (during Phase B in Figure 2-5); the reproduction of the nano-milling process to deliver company-specific compounds, and the sterilization of the manufacturing process using nano-technology. Finally, Phase B of the pharmaceutical Discovery process case is the pre-clinical testing phase where the new chemical candidate is further characterised through in-vivo tests in animals and humans.

Concept Application is defined as a complexity-handling mode leading to a mental model characterising the full innovative drug delivery system's application to the level of all relevant variables and their relationships, with a specification of the values delivering an effective application system. Hence, following Boisot and Child's (1999) typology, experienced complexity is predominantly reduced by specifying all value ranges and problem-solving mechanisms for one application domain.

Gamma Controlled Release (Figure 2-6) Phase C and Alfa Immediate Release tablet (Figure 2-7) Phase A are two typical examples of a Concept Application type experimentation behaviour. In both examples the characterised application domain is fully optimized for commercial usage using highly formalized development procedures and efficiency-enhancing stage-gate driven decision-making.

Table 5-1 below describes the three identified complexity-handling modes along the dimensions target setting, experimentation, learning, and coordination approaches followed by the innovation team. It combines the descriptors as found in the exploratory project in pharmaceutical Development (see Table 2-2) adapted by the findings from the confirmatory Discovery case documented in Table 3-1 to form the definitive empirical descriptor set for the complexity-handling modes emerging from my seven case studies.

	Complexity-handling modes		
	Concept Selection	***Concept Characterisation***	***Concept Development***
Predominant complexity-handling	• Complexity absorption	• Complexity absorption (application domains and uncertainty areas) • Complexity reduction (definition of value ranges)	• Complexity reduction
Target setting approach	• *Target defined up-front* as minimal system critical requirements to pass	• *Target moving* toward feasible application domain requirements to pass	• *Target defined up-front* as product/process requirements to pass. Only modifiable at formal stage-gates
Experimentation approach	• Define uncertainty areas to characterise critical variables and to anticipate problems to be solved • *Run parallel experiments to characterise* critical variables affecting response for different candidate system solutions • System solution selection • Show Proof of Concept by spelling out assumptions about the set of relevant variables and their functional relationships	• *Run parallel/concurrent experiments to characterise* all variables affecting response for different uncertainty areas within proof of concept delivered solution • Integrate uncertainty areas into limited / characterised application domain	• *Run parallel/concurrent experiments to optimise* variables values within a solution • Followed by concurrent engineering driven by QFD derived product definition • FMEA based process design
Learning approach	• *External explicit:* Mainly at project start learning from published science. Later ad-hoc for problem solving. • *External tacit:* On-going knowledge transfer by interaction between teams and external technology suppliers. • *Internal tacit:* Use of pockets of previous knowledge	• *External explicit:* Mainly at project start learning from published science. Later ad-hoc for problem solving. • *External tacit:* Knowledge-transfer mainly at project start between external technology supplier and team. Later for problem solving. • *Internal tacit:* Use of pockets of previous knowledge	• *Internal explicit:* Learning based on formal procedure-based inquiry • *Internal tacit:* Learning by doing, based on previous internal experience
Coordination approach	• Define milestone targets that are reached if results can be shown • Estimate work package effort/ timeline based on first experiments or expert knowledge • Focus on experiments capable of selecting as quickly as possible solutions that meet all system critical requirements • Through close monitoring of progress: Eliminate as quickly as possible candidate system solutions not meeting one of the system critical requirements	• Define uncertainty areas and assumptions to be tested per area • Define milestone targets for uncertainty areas that are reached if results can be shown. • Use DoE experimental guides to systematize the testing process • Guide progress through real-time coordination of concurrent results of different uncertainty areas • Bring focus through adaptive learning i.e.; assumptions testing, learning, continue/redirect efforts to characterise the feasible application domain • Possible go back to previous mode if application domain cannot be delivered or if new application domain emerges	• Define product/process requirements to be met • Use pre-designed DoE experimental guides to systematize the experimentation process • Use GANTT type plans & schedules for cross-activity programming and tracking task completion • Monitor plan variation and act accordingly by executing contingency plans • Use of standard approaches and documented best practices to problem-solving

Table 5-1 Empirical complexity-handling modes descriptor set

Experienced complexity typology. Complexity-handling modes describe the way the innovation team handles the complexity it is experiencing. For the sake of the following argumentation I define e*xperienced complexity* as a property of the interaction between problem solver and problem. Various e*xperienced complexity types* exist where each type is characterized by differing levels of uncertainty and ambiguity chosen by the innovation team. Hence, in my Interpretative view of complexity the levels of uncertainty and ambiguity facing the innovation team are not exogenously given problem variables as proposed by the logico-scientific view of complexity (see Section 2.3 for more details), determined by the properties of the complex system to be designed. Instead, in my theorizing effort I follow Schrader *et al.* (1992) that it are two dissimilar components of problem framing whose level the innovation team dynamically chooses. Therefore, I propose experienced complexity to explain the complexity-handling mode chosen dynamically through the course of the innovation project, regardless of the latter being radical or incremental.

Proposition 1: The experienced complexity type facing the innovation team determines the choice for a complexity-handling mode

The innovation team handles experienced complexity by constructing a *mental model* of the problem-solving situation. Mental models specify the problem-solving mechanisms, the solution variables, their relationships and their value ranges that are relevant for understanding and describing the problem and provide the solution space within which the problem will be solved. Key to understanding the subjective nature of experienced complexity-handling through mental models is the acknowledgement that ambiguity and uncertainty are not exogenously given characteristics of the problem to be solved. Instead, mental models are constructed through a creative problem framing and solving process in which levels of ambiguity and uncertainty are dynamically chosen and reduced by the innovation team as the mental modelling process progresses (Schrader et al. 1992). Case evidence documented in Appendix B provides empirical indicators for this endogenous view of complexity.

Experienced Complexity Type	***Proof of Concept Ambiguity-based***	***Ambiguity-based***	***Uncertainty-based***
Ambiguity level	Proof of Concept (PoC) critical variables and their functional relationships, and/or problem solving mechanisms unknown to the team	Only PoC critical variables and their functional relationships, and relevant problem solving mechanisms known to the team	All application domain specific variables and their functional relationships, and problem solving mechanisms known to the team
Uncertainty level	Value ranges of critical variables unknown to the team	Not all value ranges of critical variables known to the team	Value ranges for all application domain specific variables known to the team
Case phases where experienced complexity type dominated	DNA-Zyme Phase A Microemulsions Phase A SCF Phase A Nanosuspension Phase A Discovery Phase A	Microemulsions Phase B SCF Phase B Nanosuspension Phase B Gamma CR Phases A and B Discovery Phase B	Nanosuspension Phase C Gamma CR Phase C Alfa IR tablet Phase A

Table 5-2: Experienced complexity typology with relevant case examples

In the following I characterize experienced complexity using levels of ambiguity and uncertainty facing the innovation team. Table 5-2 provides the definitions for the different constructs emerging from the cases. It is a summary of Table 2-3 and the Discovery case process data and classifies all relevant cases of Chapters 2 and 3.

Empirical indicators of uncertainty and ambiguity during Concept Selection evidence that the complexity experienced by the team is Proof of Concept (PoC) ambiguity-based. Ambiguity level is PoC-based since the mental model containing relevant variables to solve the problem and/or the problem solving mechanisms needed to get the solution at PoC level is not yet formed within the team. The innovation team goes through a creative search process among potential candidate solutions, composed of variable sets, to discover the set of variables that might generate a system that has the required solution critical functionality. Since during this process the winning mental model and its related variable sets is not yet known, by definition neither could the variable value ranges be known. Hence, although uncertainty level is also high this is not relevant for the choice made by the innovation team. First, the winning mental model must be found, showing PoC-level performance on the solution critical requirements. Therefore, I call this experienced complexity type *Proof of Concept ambiguity-based*.

In the DNA-Zyme, SCF, Micro-emulsions and PharmaCo's Discovery process cases this mode was chosen while no evidence existed to the team that proof of concept had been delivered and documented before for these new-to-world technologies. In the Nanosuspension case this mode was chosen while proof concept had not been delivered yet for PharmaCo's own compounds. This leads me to formulate the following proposition;

> *Proposition 1a: A 'Concept Selection' complexity-handling mode will be chosen if the innovation team experiences Proof of Concept ambiguity-based complexity.*

Secondly, case evidence reported in Appendix A suggests that the choice to manage the innovation project following the Concept Characterization complexity-handling mode is made by the team whenever it is facing an ambiguity type sense-making opportunity. Ambiguity is experienced since only a mental model containing relevant variables and problem-solving mechanisms to solve the problem at proof of concept level has been formed within the team. Since neither problem-solving mechanisms, nor problem relevant variables nor their functional relationships are known beyond proof of concept, by definition neither could the variable value ranges be known beyond this level. Although uncertainty level is still high this is not relevant for the choice made by the innovation team. Therefore, I call this experienced complexity type ambiguity-based.

> *Proposition 1b: A 'Concept Characterization' complexity-handling mode will be chosen if the innovation team experiences ambiguity-based complexity.*

Concept application domain characterisation will gradually resolve ambiguity and reduce uncertainty to a level where the innovation team's mental model of the system to be designed contains all variables with their value ranges relevant to start application development. The Gamma CR project faced an initially ambiguous situation but variables and their functional relationships relevant for proof of concept performance were accessible to the team, which led them to the choice of Concept Characterisation as complexity-handling mode. Once PoC delivered for PharmaCo's own compounds, for the Nanosuspension project the innovative technology needed to be made sterile and ready to accept company specific compounds. In the Gamma CR project the right choices in a number of uncertainty areas were still needed to get to an optimal kinetic profile and to make the technology ready for implementation in an application to be delivered to the market. After delivery of the initial proof of concept that certain polymers in interaction with $SCCO_2$ can lower the required meltextrusion temperature hence extending its application range to thermolabile active ingredients, the SCF project team chose this complexity-handling mode to further develop the technology to a level were it can be used for application development. In PharmaCo's Discovery process case, Phase B (see Figure 3-4) is the pre-clinical characterisation of a compound transferred from Discovery. Here, Concept Characterisation was only chosen as complexity-handling mode from the moment proof of concept was delivered.

Finally, I propose that the choice to manage the innovation project following the Concept Application complexity-handling mode is made by the team whenever it is facing an uncertainty type sense-making opportunity without any residual ambiguity. Only uncertainty, no or few residual ambiguity is experienced since for a specific application domain a mental model containing all relevant problem-solving mechanisms and solution variables with their respective value ranges has been formed within the team's collective mind. Since the level of experienced ambiguity is negligible it is not relevant for the choice made by the innovation team. Therefore, I call this experienced complexity type uncertainty-based. This leads me to formulate the following proposition which is in line with the proposition formulated by Schrader *et al.* (1992) that problems will be framed involving little ambiguity if the problem-solver has successfully solved apparently isomorphic or related problems previously:

Proposition 1c: A 'Concept Application' complexity-handling mode will be chosen if the innovation team experiences uncertainty-based complexity.

The Alfa IR Tablet project started with a full mental model of the problem-solving situation available. Residual ambiguity was negligible since the project could be developed using platform technologies like concept formulations and documented best practices. The innovation team chose this mode while they had successfully solved apparently isomorphic problems before leading them to rule out ambiguity and to choose for a complexity-handling mode that is strong in delivering results at pre-planned stage-gates, using rigorous experimentation techniques. Likewise, in the Gamma CR case (see Figure 2-6) the innovation team chose this complexity-handling mode as soon as the combined IR/CR beads solution was characterized.

Experienced complexity dynamics. The incapacity of project exogenous characteristics cited in the literature –like radical, fuzzy front-end or incremental innovation project- to explain complexity-handling mode transitions is evidenced by a number of my cases. The Micro-emulsions project will soon change from Concept Selection mode to Concept Characterisation mode, although it will still be a 'Super-High-Tech' technology while still based on non-existent technologies at project initiation. The same holds for the SCF project that has changed complexity-handling mode but is still a 'High-Tech' project. Also, it remains to have a 'project uncertainty profile' (Loch et al. 2000; De Meyer et al. 2002) characterised by ambiguity, variation and risk.

Instead, by taking an interpretative approach my case data lead to a better understanding how and why experimentation approaches vary dynamically over the course of the project. By taking the perspective of the mental model of the problem-solving situation, one can see that the team gradually builds up understanding of the innovative system to be designed, continuously reducing levels of *residual ambiguity*, the latter being defined as the problem-solving mechanisms and solution variables remaining to be characterized[48]. The decision to transit to a new complexity-handling mode, then, is driven by the perceived completeness of the team's mental model. As long as ambiguity has not been sufficiently resolved or uncertainty sufficiently reduced, the team will stay in a certain complexity-handling mode.

> *Proposition 2: The decision to change complexity-handling mode is determined by the innovation team' perceived completeness of their mental model of the problem-solving situation at hand, operationalised in the proven delivery of minimum solution requirements.*

The decision to transit from Concept Selection to Concept Characterisation mode is made by the team as soon as solution critical requirements are met. High initial ambiguity must be resolved to the level that the emerged mental model contains all problem-solving mechanisms, solution critical variables and their relationships necessary to deliver Proof of Concept.

> *Proposition 2a: The decision to transit from 'Concept Selection' to 'Concept Characterisation' complexity-handling mode is made by the innovation team as soon as a mental model has emerged that contains all critical problem-solving mechanisms, solution variables and their relationships necessary to deliver Proof of Concept.*

Nor the DNA-Zyme, nor the micro-emulsions project have reached this transition since no delivery solution has been reached in the former nor has a definitive set of self-emulsifying polymers been found yet in the latter. The team does not know the definitive set of critical variables nor their relationships delivering proof of concept yet. Mid 2000 proof of concept was delivered in the Supercritical fluids project when polymers in interaction with $SCCO_2$ showed that they can decrease the meltextrusion

[48] The concept of residual ambiguity was introduced in section 3.3.6 as a construct emerging from the Discovery case data and used to formalize front-loaded experimentation strategy types.

temperature and that active substance can be made soluble in $SCCO_2$ under certain critical conditions. This led to the project team's decision to make the transition to the Concept Characterisation mode where the design of the melt-extruder using $SCCO_2$ is further characterised. In PharmaCo's Discovery process case, the transition from the first to the second mode is formally made as soon a chemical candidate fulfils the solution critical requirements of a NME, to be further characterised and applied to a specific project in pharmaceutical Development.

Secondly, the decision to transit from Concept Characterisation to Concept Application mode can be made by the team as soon as the application domain for the concept is characterised and solution critical requirements are met. Ambiguity must be resolved to the level that the emerged mental model contains now all problem-solving mechanisms, solution variables, their relationships and value ranges necessary to deliver an application within the characterised application domain.

> *Proposition 2b: The decision to transit from 'Concept Characterisation' to 'Concept Application' complexity-handling mode can be made by the innovation team as soon as a mental model has emerged that contains all problem-solving mechanisms, solution variables, their relationships and value ranges necessary to deliver an application.*

In the Controlled Release project the transition decision to go for application development was made after the concept showed performance following a kinetic release profile that was hypothesized by the team. Then the team could switch to a project management approach tailored to manage 'incremental' projects. No other case projects made this transition yet since the mental model depicting the innovative delivery system has not crystallised to a level that it can be used to develop an application. In the Nanosuspension case the transition from Phase B to C (see Figure 2-5) was planned and ready to be made. Only, the clinical candidate on which the new technology could be applied got halted in the Development process, therefore preventing the advanced technology project to reach Concept Application status.

Finally, during Concept Characterisation and Concept Application modes a decision can be made by the team to go back to one of the previous modes if a fundamental problem arises or a new situation emerges preventing moving forward.

> *Proposition 2c: The decision to move back from 'Concept Characterisation' or 'Concept Application' complexity-handling modes can be made by the innovation team as soon as a situation arises where the present complexity-handling mode does not lead to further absorption and/or reduction of experienced complexity.*

An example of a problem where the innovation team decided to go back from Concept Application to Concept Characterisation is provided by the Gamma CR project where a drug candidate was in Development featuring a specific kinetic profile. However, the latter showed to be insufficient to meet the needs of the patient population. Therefore, the team was forced to revisit the problem and characterize a CR solution that was known to work at proof of concept level but that had not been applied

before in the project-specific context. The SCF case is an example of a technology that had shown proof of concept and was characterized for the meltextrusion application domain. However, after this characterisation the team could have decided to go for a human bone-joining application, in which case it would have to go back to Concept Selection to find ways to get to proof of concept for this new application domain.

A prescriptive framework explaining experimentation behaviour. In the previous, following Doty and Glick (1994) I derived inductively a set of complexity-handling modes and types of experienced complexity and proposed a description using the same set of dimensions. Then, I proposed a relationship between experienced complexity types, the explaining construct, and complexity-handling modes, the explained construct, chosen by the team to manage their project. Finally, I asserted that the delivery of minimal solution requirements, a specific event within the mental model emergence process, could be used to explain complexity-handling mode transitions. Figure 5-3 below provides a prescriptive framework of the propositions discussed so far.

Experienced complexity type is proposed to drive the choice for a specific complexity-handling mode (P1a, P1b, P1c). The delivery of minimum system requirements is proposed to drive the emergent mental model (P2a, P2b, P2c).

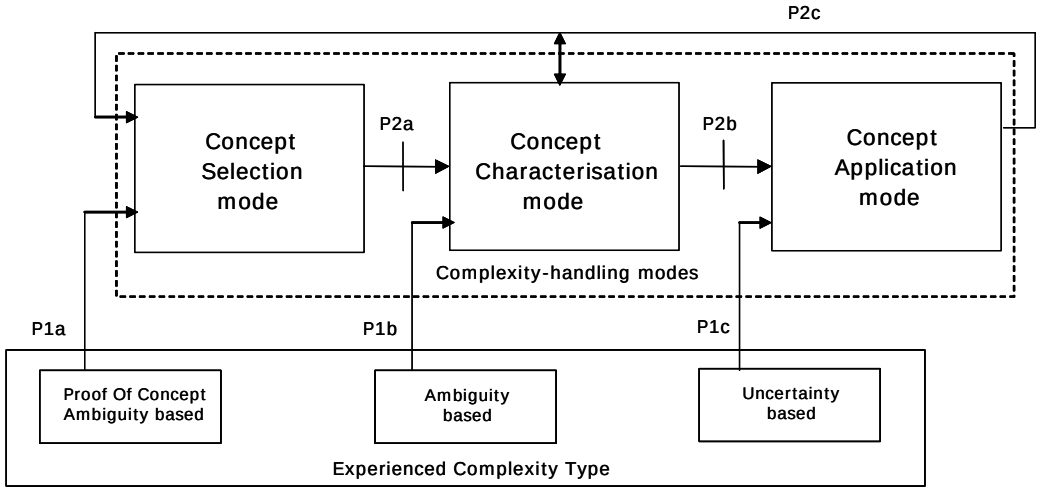


Figure 5-3 Prescriptive framework explaining complexity-handling behaviour

The formulated framework focuses on complexity-handling as a result of complexity experienced in finding a solution to a defined innovation problem. The PharmaCo Discovery case has shown that significant efforts are spent defining the innovation problem –finding a relevant biological target-, before one gets to chemical problem-solving. It should be made clear to the reader that this model assumes product or at least innovation problem definition has been carried out before and is not part of the ambiguity component of the complexity experienced by the team.

5.2.2 Propositional model relating experimentation strategies to predictive performance

The top-down developed part of the formal teleological process theory explains the relationship between Concept Selection, one of the complexity-handling modes chosen by the innovation team, and performance. A Bayesian framework to quantitatively evaluate predictive performance was developed in Chapter 3 and used in Chapter 4 to compare experimentation strategies used during Concept Selection. The PharmaCo Discovery process case study conducted in Chapter 3 allowed me to document various experimentation strategies used in one of the technology-intensive front-end innovation environments where this complexity-handling mode is typically used. Simulation model design, behavior and discussion of the results were further documented in Chapter 4.

Two key dimensions of experimentation strategies for Concept Selection were used for theorizing using a top-down computer simulation-based model; (1) the shape of the solution concept funnel, and (2) the number of solution variables and problem-solving mechanisms characterized at various points during the Concept Selection process. As depicted in Figure 5-4, the first is indicated with a vertical arrow, the second by the level of shading used for each activity. The more variables characterized, the more intense the shading.

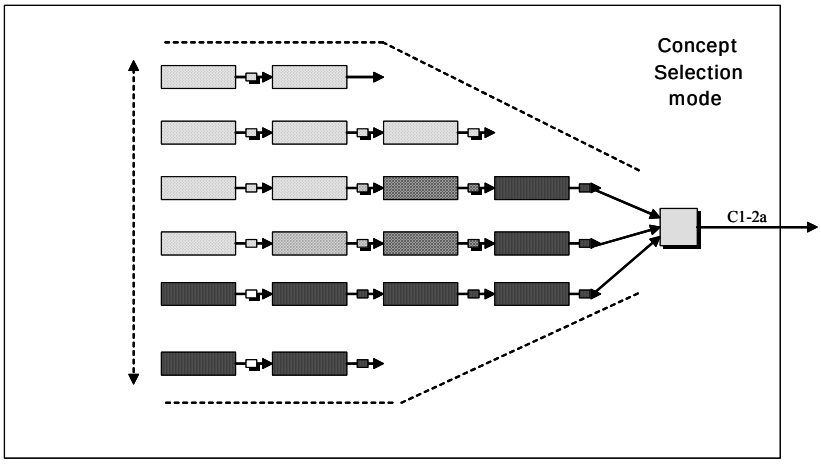
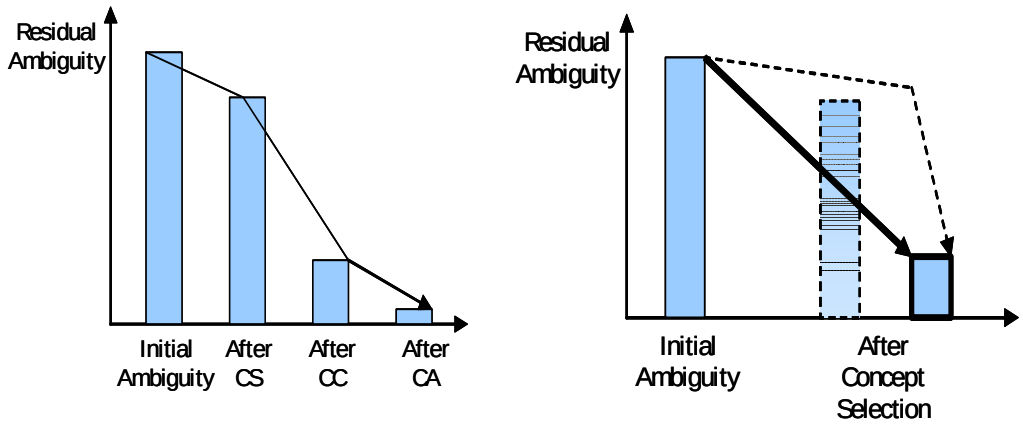


Figure 5-4 A summary of Concept Selection experimentation strategies

The effect of front-loading. As previously discussed, the effect of front-loading on predictive performance has not been studied before. To do so, the concept of residual ambiguity needs to be revisited and put in a dynamic perspective as depicted in Figure 5-5(a). As soon as the problem is defined, problem solving can start. Initially, the residual ambiguity component of the complexity experienced by the innovation team is high. It will gradually be lowered as the team builds its mental model of the solution, leaving less and less solution variables and problem-solving mechanisms to be modeled.

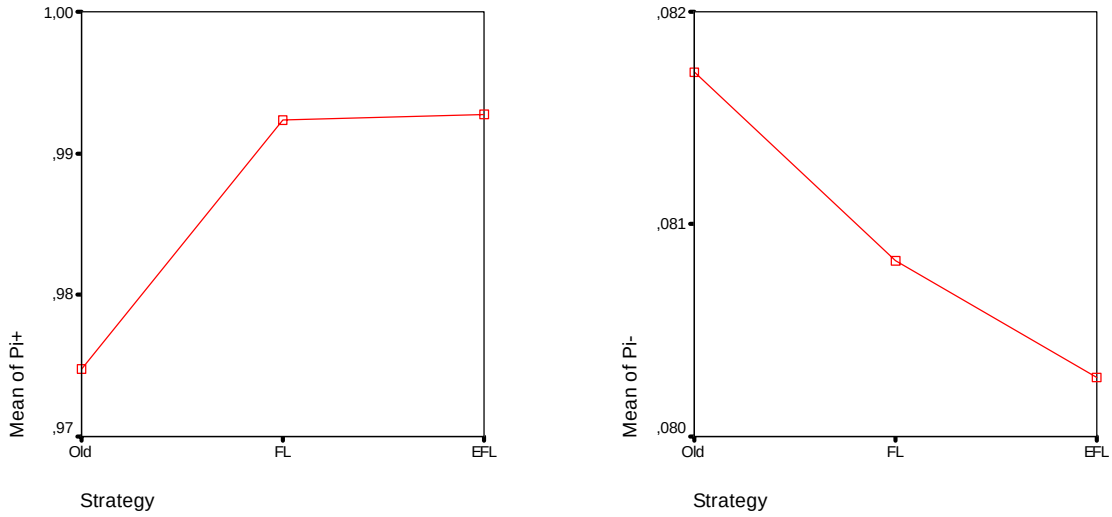
My simulation results indicate that two elements of this residual ambiguity reduction dynamic could play a role in explaining predictive and business performance of front-loaded strategies used in Concept Selection.



(a) **(b)**

Figure 5-5 A dynamic perspective of residual ambiguity

Figure 5-5 (b) depicts these two elements; (1) the steepness of the path leading from initial ambiguity to the residual ambiguity after Concept Selection (CS), and (2) the level of residual ambiguity after CS. In words, this picture graphically portrays that the earlier and the more ambiguity get resolved during Concept Selection, the higher will be its predictive performance (indicated by the full arrow and emphasized bar).



(a) **(b)**

Figure 5-6 Positive and negative predictive performance of Concept Selection experimentation strategies

Figure 5-6 summarizes my simulation results for predictive performance of Concept Selection experimentation strategies showing a slightly better positive and negative predictive performance for front-loaded strategies. A detailed discussion of two-way and one-way ANOVA simulation results is provided in Chapter 4. Both Front-

loaded strategies (Early Frontloading and Frontloading) significantly outperform the Old Paradigm experimentation strategy on positive predictive performance. However, the difference between both Front-loaded strategies is insignificant, meaning in the context of pharmaceutical Discovery that in-silico characterization of ADME-T does not significantly increase positive predictive value as compared to ADME-T characterization starting in H2L, as done in Front-loaded paradigm. Second, although Figure 5-6 (b) shows a downward trend in negative predictive value towards the Front-loaded paradigm, the differences are insignificant.

These results lead me to formulate the following propositions relating Concept Selection experimentation strategy to its predictive performance:

> *Proposition 3: There is an inverse relationship between the chosen level of residual ambiguity at the end of Concept Selection and its positive predictive value.*

In contrast to some practitioner views, my simulation results show non-significant differences in negative predictive performance for the various strategies. This indicates that front-loaded strategies do not decrease the probability of missed opportunities as compared to an Old Paradigm strategy. They only increase its positive predictive value, increasing the chances of surviving Concept Characterization and Concept Application testing once declared active at the end of Concept Selection.

Furthermore, simulation results indicated that front-loaded experimentation strategies significantly select *overall* better quality compounds than an Old Paradigm strategy. In other words, taking into consideration all solution variables necessary for delivering Proof of Concept, leaving less variables uncharacterized, leads to better solutions selected during Concept Selection. This leads me to formulate the following proposition:

> *Proposition 4: Front-loaded experimentation strategies select overall better quality concepts than non-front-loaded strategies during Concept Selection.*

Finally, in my simulation front-loaded strategies consistently outperformed Old Paradigm strategies on overall quality selection power and predictive performance, regardless of the tightness level of the surrogate marker chain. However, a minimum tightness level was found to exist below which positive predictive performance significantly deteriorated. Negative predictive performance was not influenced by the chain's tightness level. Surrogate marker chain tightness was a proxy used in the specific pharmaceutical Discovery context to measure the predictive power of the problem-solving mechanisms used during experimentation. Therefore, translating these substantive results to a more general applicable level leads me to formulate the following proposition, derived from P 4-4 under previous section 4.5.1:

Proposition 5: Front-loaded experimentation strategies used during Concept Selection will feature better overall quality selection power, and higher positive and negative predictive performance than non front-loaded strategies regardless of the predictive power of the problem-solving mechanisms used during experimentation. Lower levels of predictive power deteriorate performance of overall quality selection power and positive predictive performance variables.

The effect of parallelism. Previous studies in technology-intensive industries indicate the benefits of broadening the concept testing funnel (Sobek II et al. 1999) or at least propose to optimize the shape of the concept funnel (Dahan and Mendelson, 2001). However, the impact of concept funnel shaping strategies on predictive performance has not been studied before.

Figure 5-7 below illustrates my simulation results conducted in a context emulating the pharmaceutical Discovery experimentation and decision-making process. Broadening the solution concept funnel was found to have significant positive effects on quality selection power and predictive performance of experimentation strategies during Concept Selection. In words, this picture graphically portrays that opening up the funnel improves performance (indicated by the full emphasized lines).

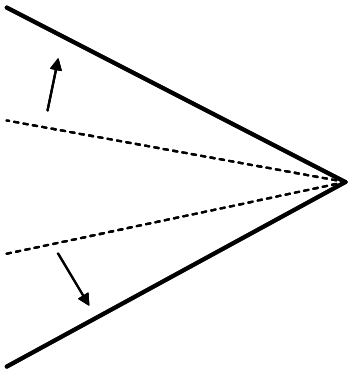


Figure 5-7 Broadening of solution concept funnel during Concept Selection

This leads me to formulate the following propositions relating funnel shaping strategies to quality selection power and predictive performance:

Proposition 6: Broadening the funnel in a Front-loaded experimentation strategy during Concept Selection increases its negative predictive power, significantly decreasing the chances of missed opportunities in subsequent development. A minimum number of parallel concept explorations are required to gain effect.

Proposition 7: Broadening the funnel in a Front-loaded experimentation strategy during Concept Selection selects better overall quality concepts. A minimum number of parallel concept explorations are required to gain effect.

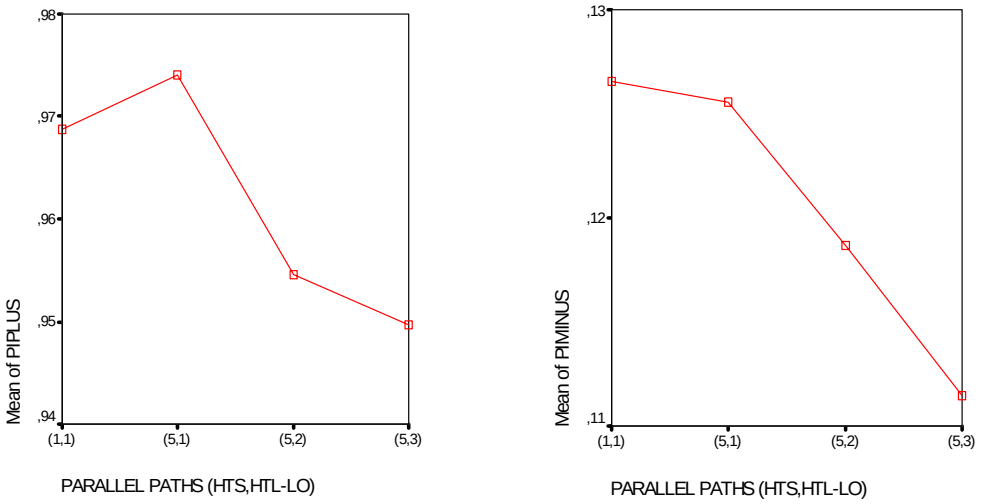


Figure 5-8 Positive and negative predictive performance impact of solution concept broadening strategies

However, simulation results did show that broadening the concept exploration funnel has a significantly deteriorating effect on positive predictive performance during Concept Selection as evidenced in Figure 5-8.

Proposition 8: Broadening the funnel in a Front-loaded experimentation strategy during Concept Selection decreases its positive predictive power, significantly decreasing the chances for a candidate solution concept of surviving Concept Characterization and Concept Application testing once promoted at the end of Concept Selection.

Finally, simulation results indicate that no minimum tightness level is required for broader funnel shaping strategies to outperform leaner strategies. However, a minimum tightness level was found to exist below which positive predictive performance significantly deteriorated. Negative predictive performance was not influenced by the chain's tightness level. Surrogate marker chain tightness was a proxy used in the specific pharmaceutical Discovery context to measure the predictive power of the problem-solving mechanisms used during experimentation. Therefore, translating these substantive results to a more general applicable level leads me to formulate the following proposition, derived from P 4-9 under previous section 4.5.1:

Proposition 9: Broader funnel shaping strategies used during Concept Selection will feature better overall quality selection power, and higher positive predictive performance than leaner strategies regardless of the predictive power of the problem-solving mechanisms used during

experimentation. Lower levels of predictive power deteriorate performance of these outcome variables.

5.2.3 Propositional model relating experimentation strategies to business performance

This study's simulation results indicate that front-loaded strategies applied during Concept Selection lead to higher business value than Old Paradigm strategies. Although not statistically significant, applying front-loading earlier did lead to more business value. Also, broadening the concept funnel to a certain optimum point had a positive impact on business performance. Hence, the following propositions, visualized in Figure 5-9:

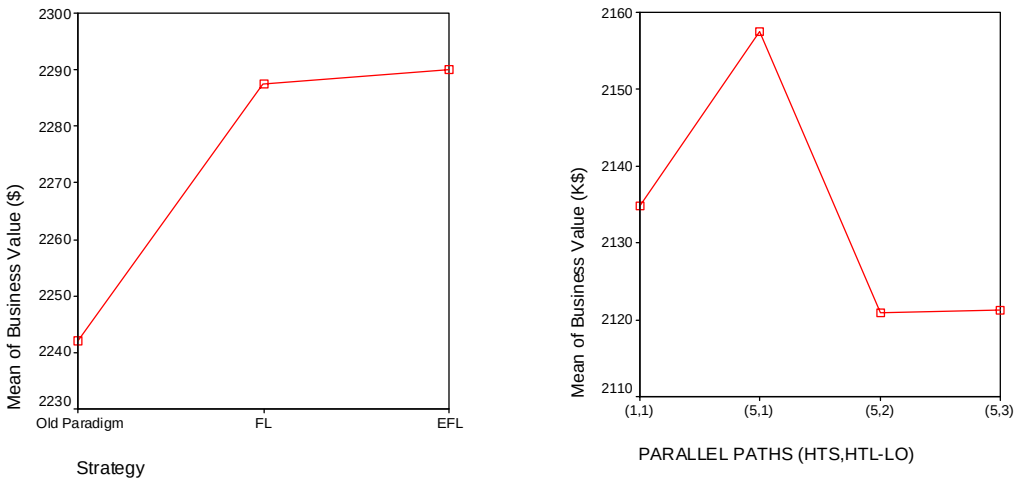


Figure 5-9 Business performance of front-loaded experimentation strategies

Proposition 10: Front-loaded experimentation strategies lead to higher business value than non-front-loaded strategies.

Proposition 11: Broadening the funnel to an optimum point in a Front-loaded experimentation strategy during Concept Selection leads to optimal business performance.

Figure 5-10 below summarizes all abovementioned propositions derived from the simulation-based study into a propositional model linking experimentation strategy related decisions, made to carry out Concept Selection, to predictive and ultimately to business performance. The signs on the edges indicate how the ending nodes are proposed to change given a change in the starting node. A plus sign (+) indicates a move in the same direction. A negative sign (-) indicates a move in the opposite direction. Proposition numbers are also indicated along the edges.

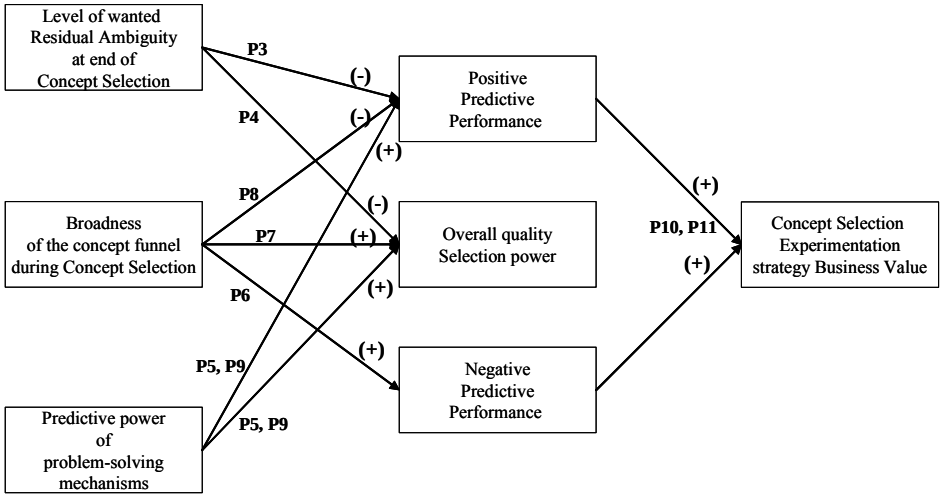


Figure 5-10 Propositional model built from simulation-based case study

As an example of how to read the propositional model set out above; the level of Residual Ambiguity wanted at end of Concept Selection is clearly determined by the choice for one of the front-loaded experimentation strategies. Front-loaded strategies increase positive predictive performance by opting for lower levels of Residual Ambiguity at the end of Concept Selection than an Old Paradigm strategy. Taking the example further this consequently leads to increased business performance given the (+) relationship between positive predictive performance and business performance. Conversely, broadness of the concept funnel during Concept selection has opposite impact on respectively positive (-) and negative (+) predictive performance, at their turn both positively influencing business performance. Therefore, from this signed directed graphs propositional model, the impact of broadening the concept funnel on business performance cannot be determined.

5.2.4 Predictive model linking experimentation strategy to business performance

In my theory-building effort I would like to go further than a propositional model and propose a framework *predicting* business performance given a set of choices made along the lines of the above mentioned propositions. Therefore, in the following I will argue that to build such a predictive model of business performance of experimentation strategies, the decision tree-based model developed in section 3.5.2 including conditional probabilities needs to be augmented (Bielza et al. 1999) with the decisions related to the propositions set out above.

The decision tree depicted in Figure 3-11 of section 3.5.2 above was developed to evaluate and compare the business performance of experimentation strategies. The R&D process was modeled as a sequence of experimentation phases concluded by a

decision to carry forward or to terminate the solution concept. Two pharmaceutical R&D key decisions were taken into consideration in the tree; the decision to transfer a compound from Discovery into Development, and the decision to take a compound to market, after declaring it *really* active.

To construct a predictive model starting from the propositions set out above, linking business performance to experimentation strategies carried out for Concept selection, I chose to convert the decision tree relating business value to predictive performance (see Figure 3-10b) in a *Sequential Decision Diagram* (Covaliu and Oliver, 1995) also called *Bayesian belief network* (Pelikan and Goldberg, 2003). The latter is a directed acyclic graph with nodes corresponding to the variables of interest and edges between the nodes representing dependencies between the variables. The advantage of using this representation instead of the decision tree representation used in Chapter 3 is that it allows me to easily represent conditional probabilities tied to decisions made in the context of an experimentation policy. Augmenting the decision tree with the various decisions ensuing from the propositions set out above would make the tree very large and complicated to handle (Smith, 1989; Shenoy, 1992; Covaliu and Oliver, 1995), which obviously I want to avoid. A transformation of the augmented decision tree would lead to the sequential decision diagram or Bayesian belief network depicted in Figure 5-11 below.

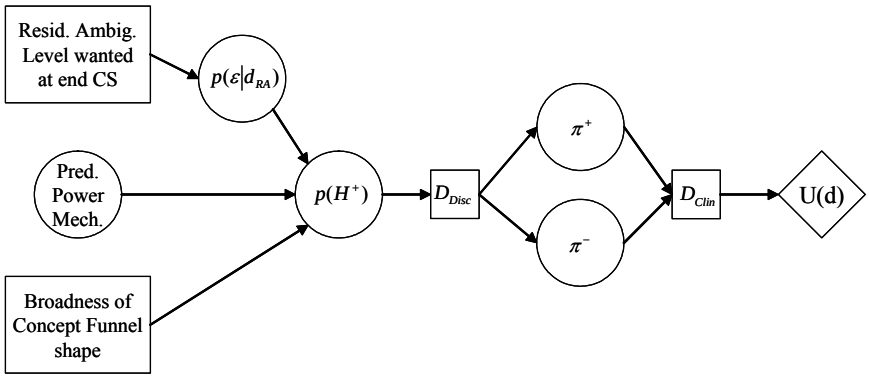


Figure 5-11 Bayesian belief network representation of the pharmaceutical R&D process

Since the decision tree is symmetrical –meaning scenarios always result in the same sequence of value realizations (Jensen, 2001)- it can be transformed in the sequential decision diagram above (see for details Covaliu and Oliver, 1995; Bielza et al. 1999; Jensen, 2001). As a drawing convention, the directed edges relate two variables so that the terminal or ending node depends on the initial or starting node. The character of the dependencies is then specified by a table of conditional probabilities, including a probability of each value of the variable conditioned on each possible configuration of its parents. The structure and the probability tables fully determine a probability distribution, which can be written as a chain rule:

$$p(X) = \prod_{i=1}^{n} p(X_i \mid \Pi_i) \quad \textbf{(5-1)}$$

where $X = (X_1, X_2, ..., X_n)$ is a vector of random variables, Π_i is the set of parents of X_i (Pelikan and Goldberg, 2003). The basic decision tree is now augmented with the two decisions studied and one uncertainty; (1) the level of residual ambiguity wanted at the end of Concept Selection, (2) the wanted broadness of the concept funnel shape, and (3) the uncertainty about the predictive power of problem-solving mechanisms used. The level of residual ambiguity allowed for by the experimentation strategy influences the uncertainty distribution $p(\varepsilon \mid d_{RA})$ caused by the level of variables left uncharacterized (ε) for reaching Proof of Concept. Applied to the pharmaceutical Discovery case the decision to implement a front-loaded experimentation strategy will leave less variables uncharacterized than an Old Paradigm strategy, which will positively impact positive and negative performance, hence business performance. Predictive power of problem-solving mechanisms has been proposed to have a moderating effect on positive predictive performance hence on business performance.

The Bayesian belief network set out above is the second prescriptive framework of my proposed teleological theory of experimentation behavior in the fuzzy front-end innovation process. By using the chain rule (5-1) set out above the utility U(d) or business value of a particular configuration of decisions made in an experimentation strategy can be calculated and a choice for optimal performance can be made. Or, formally;

$$\begin{aligned} U(d) &= f(u, d_i, p(H^+), \pi^+, \pi^-) \\ &= f(u, p(H^+ \mid d_{RA}, d_{funnel}, \varepsilon, p_{psmech}), \pi^+, \pi^-) \end{aligned} \quad \textbf{(5-2)}$$

Hence, based on this belief network an optimal experimentation strategy can be determined. Applied to the pharmaceutical Discovery case it can be verified in Figure 5-9 above that, given the utilities u –revenues and costs documented in section 4.4.1 above describing the simulation model parameters used, and given a predictive surrogate marker chain of 70%, the choice for a front-loaded experimentation strategy complemented with the decision to apply a (5,1) funnel shaping strategy could lead to maximum business performance. As to the latter funnel shaping strategy, this example shows that a predictive model can go a step further than the propositional model above that was incapable of showing what the funnel broadness was optimizing business performance.

Concluding, the previous sections formulated a teleological process theory linking the experimentation and decision-making process during Concept Selection to business performance. The proposed theory consists of a bottom-up developed prescriptive framework explaining the choice for a complexity-handling mode given a type of complexity experienced by the innovation team. Second, a top-down developed propositional and a predictive model was proposed linking the execution of one of these

complexity-handling modes –Concept Selection- to its predictive and business performance.

In the following sections the proposed innovation process theory will be situated in the various literature domains to which it claims to make a theoretical and practical contribution.

5.3 DISCUSSION

The primary aim of this thesis was to advance new theory explaining the mechanisms of experimentation behaviour that increase predictive and business performance of the fuzzy front-end innovation process. Empirical evidence of this study complemented with simulation results allowed me to build teleological process theory proposing a causal link between levels of front-loading and parallelism used during fuzzy front-end experimentation and predictive and business performance of the innovation process.

Overall, findings support the viewpoint developed in other industries that front-loaded and parallelized experimentation strategies enhance innovation process performance. However, my study is the first to extend current thinking on the benefits of these experimentation strategies to include their contribution to increasing predictive performance as an outcome variable of the innovation process. Also, it raises questions about the way project management for various types of innovations is proposed by the literature.

First, I will position the proposed theory in its relevant literature domains. Then, I will elaborate on both theoretical and practical contributions of my research leading to the proposed teleological process theory advanced above.

5.3.1 Positioning the contribution of the proposed process theory in the literature

The teleological process theory developed above is positioned at the cross-roads of three bodies of knowledge (see Figure 5-12); (1) innovation as the dynamic capability of building and embedding routines to manage *internal innovation processes* like experimentation and project management approaches (Tidd et al. 2001; De Meyer et al. 2002; Thomke, 2003), (2) innovation as a *complex responsive process* drawing upon both the Interpretive e.g. (Stacey, 1995; Stacey et al. 2000; Stacey, 2001; Fonseca, 2002) and Complex Adaptive System e.g. (Holland, 1992; 1998; Goldberg, 2000) sides of the Complexity Sciences to make sense of experimentation behaviour in situations of high ambiguity and uncertainty, and (3) Decision Theory as a branch of a behavioural theory of the firm (Cyert and March, 1992) and its application to *causal thinking and decision analysis* under conditions of high uncertainty and ambiguity (Howard, 1988; Spirtes et al. 2000; Pearl, 2000).

The bottom-up developed prescriptive framework used the Interpretive view of the Complexity Sciences to *provide empirical evidence and develop emerging theory* in the area of routines used to manage innovation projects (Schrader et al. 1992; Pich et al. 2002). Also, *it raises questions* about existing empirical classification schemes used to suggest project management approaches to innovation project managers contingent upon the degree of technical uncertainty and complexity they are facing at the outset of their project (Pich et al. 2002). My case evidence suggests that the technological uncertainty evaluation made at the outset of the project is not a guarantee for a unique project management approach followed along the project. Instead, a dynamic view is required.

The top-down developed propositional model and Bayesian belief network used the Complex Adaptive Systems view of the Complexity Sciences augmented with causal thinking from Decision Theory to *propose new process theory* explaining predictive and business performance of front-loaded e.g. (Thomke, 2003) and set-based or parallelized experimentation e.g. (Sobek II et al. 1999) in fuzzy front-end innovation processes.

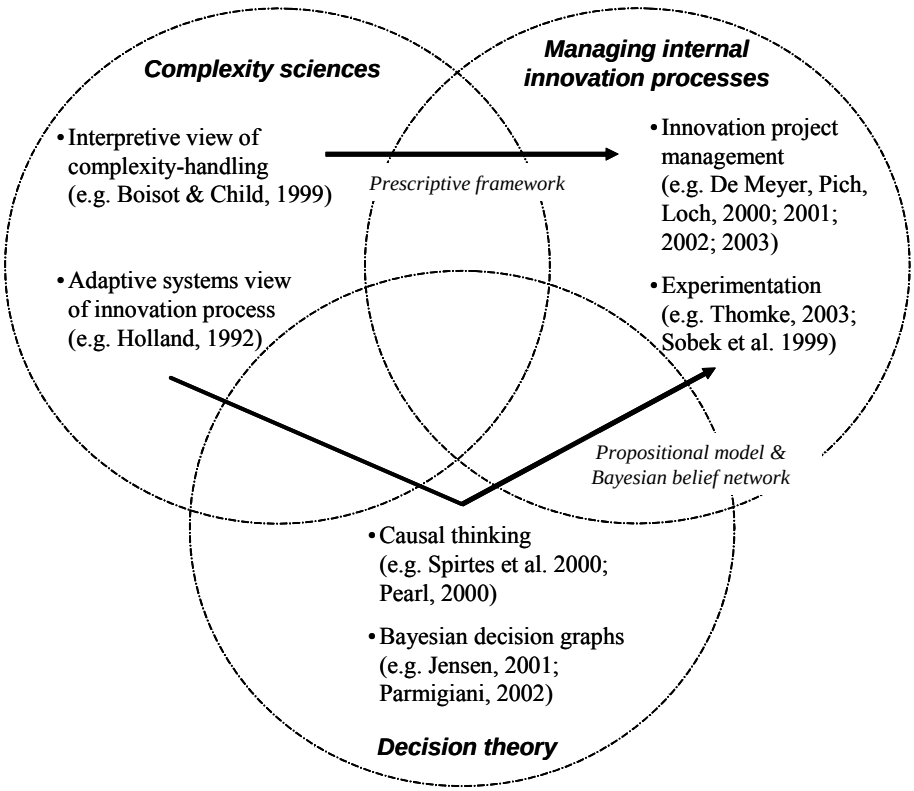


Figure 5-12 Situating my research in the literature domains

In the next section I will discuss the contributions made by both bottom-up and top-down approaches to theory development and their implications for managerial practice.

5.3.2 Theoretical contributions

Prescriptive framework. In Chapter 2, a review of the innovation routine literature focussing on the differences between managing radical and incremental innovations led me to conclude that exogenous project characteristics are poor predictors for explaining the type of experimentation approach used by the innovation team. My empirical case evidence provided in Chapters 2 and 3 showed that evaluating at the outset whether an innovation project is radical or incremental (Leifer et al. 2000), low-, medium- or high-tech (Shenhar and Dvir, 1996) is not sufficient to explain the project management approach and experimentation strategy dynamically chosen during the course of the project. Instead, I argued that to get to a more fine-grained understanding of experimentation behaviour a change in paradigm is required; attention needs to be turned to the dynamic –changing over time- nature of the endogenous problem framing and problem solving going on in the innovation project, and to the nature of the complexity experienced by the innovation team when trying to solve the innovation problem at hand.

Hence, my prescriptive framework, making the complexity-handling mode dynamically chosen by the innovation team contingent upon the type of complexity they are experiencing, supports the view proposed by Schrader *et al.* that levels of uncertainty and ambiguity are not exogenously given but are rather determined in the problem-framing process. This choice is proposed by the authors to be contingent upon context characteristics (Schrader et al. 1992) that I would specify as the type of complexity experienced by the innovation team. Also, it supports the fundamental view held by De Meyer *et al.* that an evolving innovation project uncertainty profile can be the starting point for a dynamic choice between *instructionist* and *learning-selectionist* approaches (Pich et al. 2002; De Meyer et al. 2002).

However, although in my prescriptive framework I used Schrader *et al.*'s definition of uncertainty, my research purpose leads me to use a different categorization of the ambiguity construct. More specifically, Schrader *et al.* distinguish between a situation where the variables are known but ambiguity exists in the relationships and problem-solving mechanisms, versus the worse situation where also the relevant variable set to solve the problem is in need of determination. The problem with this categorization is that it doesn't solve the need for a contingency framework explaining which project management approach to dynamically follow depending on the nature of the complexity experienced. Since in both Concept Selection and Concept Characterization modes discovered in my exploratory and confirmatory cases both types of ambiguity were prevalent and clearly a different project management approach was used, this indicated to me that a different categorization of ambiguity was required to use as an entry point to the prescriptive complexity-handling framework. Distinguishing between 'proof of concept ambiguity' and 'ambiguity', using proof of concept as a minimum threshold of the solution mental model completeness, allowed me to come to a definition of ambiguity fit for the purpose of a contingency framework, prescribing the type of project management approach to be followed, and which also stood the test of an adequate representation of reality by pharmaceutical Discovery scientists.

Second, my case evidence in pharmaceutical Discovery and Development provides empirical support for the *explained side* of the theoretical framework proposed by Pich *et al.* (2002), explaining project management behaviour under various conditions of uncertainty, ambiguity, and complexity. More specifically, my exploratory case study results found their 'Instructionism' to be the strategy used during my proposed 'Concept Application' complexity-handling mode. Also, I found their 'Selection' and 'Learning and Selection' strategy to be used during my 'Concept Selection' cases, and their proposed 'Learning' strategy during 'Concept Characterization'. However, my prescriptive framework develops their thinking further by (1) providing clarity of concept on the *explaining side* of the explanatory framework, and (2) by proposing a more explicit dynamic component to the explanatory model.

As to the first point, it is acknowledged by the authors that precise rules of when to use which approach are currently unknown, and that 'more research is needed to refine the suggested management approaches and determine when to use which approach' (Pich et al., 2002: 1021). Meanwhile, project urgency, amount of learning that can be achieved, costs, complexity and ambiguity are cited as inputs to choose for the right mix of instructionism, learning, and selectionism. Also, in a more practitioner-oriented paper a project uncertainty profile is proposed as a possible entry point to choose amongst the various abovementioned strategies (De Meyer et al. 2002). In my opinion, the latter ads to the confusion by not distinguishing between ambiguity and uncertainty, and introducing new concepts like variation and chaos. Conversely, drawing upon the work of Interpretative complexity theorists (Boisot, 1995; Boisot and Child, 1999) distinguishing between absorbed and reduced complexity, following the ambiguity- or uncertainty-related nature of the complexity experienced by the innovation team, and grounded in my case evidence, my prescriptive framework provides a clear input to the project manager facing the choice for a specific complexity-handling mode. Instead of exogenously defining complexity as 'an inability to define the effects of actions because too many variables interact' (Pich et al., 2002: 1009) I modelled experienced complexity as a construct consisting of two components – uncertainty and ambiguity as related to innovation problem solution variables- as a unique entry point to my framework.

As to the second point, in my prescriptive framework I strived for conceptual clarity of the complexity-handling dynamics describing mode transitions by introducing the concept of *residual ambiguity*. As discussed above, Schrader *et al.* (1992) distinguish between various levels of ambiguity depending on the knowledge of relevant variable sets, their causal relationships, and problem-solving mechanisms used. Pich *et al.* (2002) refer to information inadequacy arising from both project ambiguity, an endogenously defined problem-framing variable, and project complexity, the latter defined as an exogenous variable describing the complexity of the rugged solution fitness landscape (Kaufmann, 1993: 40-67). In contrast, to increase clarity in my prescriptive framework I summarized all non-characterized solution variables, their causal relationships, and/or related problem-solving mechanisms, at a specific moment in time during the innovation project, in a concept I called residual ambiguity. The latter is supposed to decrease during the course of the innovation project. This endogenously defined Interpretative complex systems concept, characterizing the actual state of the mental modelling process, is used to identify when the transition between Concept

Selection, Concept Characterization, and Concept Application can be made in the minds of the innovation team and R&D decision-makers; as an example, only if the solution concept mental model is sufficiently complete to show proof of concept, a transition between Concept Selection and Concept Characterisation can be made. The required level of residual ambiguity to accept solution proof of concept is, congruent with Schrader et al. (1992), a *choice* made by R&D management.

Summarizing, my qualitative prescriptive framework explaining complexity-handling behaviour contingent upon the type of experienced complexity, contributes to an emerging research agenda formulating prescriptive models dynamically explaining innovation project management and experimentation approaches contingent upon endogenously defined drivers of complexity, uncertainty, and ambiguity.

Propositional model and Bayesian belief network. The propositional model (Figure 5-10) and Bayesian belief network (Figure 5-11) developed above are first of a kind attempts to conceptually and quantitatively relate policy variables used to manage the fuzzy front-end innovation process to predictive and business performance.

Previous studies related frontloaded and parallelized R&D strategies to time-to-market, quality and cost efficiency performance variables. To study fuzzy front-end decision-making policy, Verganti (1999) used 'integrated product development performance', defined as a relative measure indicating the strategic priority researched companies gave to time-to-market and product quality and the level of which they outperformed their competitors on these performances. A later study (MacCormack and Verganti, 2003) corroborating these results measured product quality performance using a panel of experts participating in a Delphi evaluation process. Thomke and Fujimoto (2000) used development cost, project lead time, and number of prototypes built, to measure the proposed positive effect of shifting problem identification and problem-solving to the early phases. Terwiesch *et al.* (2002) build upon Set Based Concurrent Engineering (SBCE) practices documented at Toyota (Sobek II et al. 1999) and use development cost as a criterion to evaluate which R&D policy to follow. Finally, recent studies on parallelization (Dahan and Mendelson, 2001; Loch et al. 2001) also use cost as a criterion to optimize the level of parallelism envisaged in an experimentation policy. So clearly, my research adds predictability as a performance dimension to evaluate the fuzzy front-end policy-performance link.

In the previous I argued that to study predictive performance of fuzzy front-end discovery processes conditional probability thinking had to be applied to an adaptive search process. Complex adaptive system's *efficiency* had been studied before for genetic algorithm-based search processes (Goldberg, 1989; 2000; Holland, 1992; Mitchell, 2001) On the other hand, Bayesian algorithm-based methods to dynamically assess *effectiveness* of pharmaceutical clinical development (Parmigiani, 2002) and product launch strategies (Lee et al. 2003) were known. However, this study is the first of a kind providing a quantitative method to assess the *relative effectiveness* of various adaptive experimentation search processes based on Bayesian conditional probability thinking. Doing so, this thesis makes a modest contribution to a growing research agenda designing optimized experimentation strategies in highly ambiguous and uncertain solution spaces (Pelikan and Goldberg, 2003; Callan, 2003).

5.3.3 Practical contributions and implications

Conducted in the spirit of Mode 2 research (Gibbons et al. 1994; Romme, 2003; Van de Ven and Johnson, 2004; van Aken, 2004a; 2004b) this thesis was not conceived to only advance new theory. Instead, my study findings also have practical implications for innovation managers of technology-intensive companies, responsible for defining fuzzy front-end experimentation policies. Not only will pharmaceutical R&D managers benefit from the specific simulation study-based challenges this study raises to the recommendations made by the practitioner-based literature on discovery experimentation policy. Also, more generally this thesis contributes to good innovation practice by providing a prescriptive framework to dynamically organize the experimentation process in the face of the complexity type experienced by the innovation team. Finally, the proposed Bayesian belief network is a first step towards a decision support system supporting the design and optimization of fuzzy front-end experimentation policies in R&D organisation-specific situations.

First, my simulation results corroborate practitioner findings on experimental design in drug discovery claiming that a too high level of residual ambiguity at transfer into clinical development, leaving ADME-T characteristics uncharacterized, results in poor predictive performance (Oprea, 2002; DeWitte, 2002). Furthermore, my simulation data support the idea of being more generous in promoting compounds for further study in discovery (DeWitte, 2002). It increases significantly the experimentation strategy's negative predictive performance, which decreases the chances of missed opportunities in subsequent clinical development at a negligible cost. Also, my results confirm the deteriorating impact of the "noisiness" of the experiments conducted in discovery acclaimed by DeWitte (2002). Lower levels of tests' predictive power were shown to negatively impact positive predictive performance of experimentation strategies. Conversely, my simulation data do not support the idea that in-silico ADME-T significantly outperforms experimentation strategies where ADME-T only starts from H2L with in-vitro models (Pickering, 2001; Coty, 2002; Yu and Adedayo, 2003). Although both performing better than an Old paradigm experimentation strategy, my results show insignificant differences in predictive performance between Front-loaded and Early Front-loaded strategies.

Second, there seems to be a desire on the part of management to understand how to manage effectively the development of radically new products if, of course, it can be managed (O' Connor, 1998; Veryzer, 1998). Previous studies focussing on the difference of the radical versus incremental innovation process conclude that the development of disruptively innovative products does not seem to follow conventional stage-gate processes and find a degree of informality with respect to how this development process is managed (Veryzer, 1998). Now, the prescriptive framework presented above offers project managers in practice a diagnostic tool they can use to dynamically choose over the course of the project for a specific complexity-handling mode, contingent upon the type of complexity they're experiencing. Present diagnostic tools fall short against this framework while they are not dynamic, only considering the situation at the outset of the project, and take into account project exogenous characteristics as opposed to characteristics related to the complexity experienced by the innovation team (Shenhar and Dvir, 1996; De Meyer et al. 2001). Or, they don't

consider the role of ambiguity and only focus on uncertainty types to guide the choice for a specific project management approach (De Meyer et al. 2001).

Finally, the Bayesian belief network developed above fits the definition of a technological rule, defined by van Aken as '[a] chunk[s] of general knowledge, linking an intervention or artefact with a desired outcome or performance in a certain field of application'(van Aken, 2004b: 228). When used in the context of a specific pharmaceutical discovery organisation the network is able to calculate the relative predictive and business impact of various experimentation policies. Thus, it serves its intended purpose as a quantitative design tool supporting organizational and innovation process redesign.

5.4 LIMITATIONS AND AREAS FOR FURTHER RESEARCH

The developed teleological process theory is firmly grounded in several case studies conducted on an extended spectrum of one major pharmaceutical company's R&D operation. However, this means that in order to have the proposed models evolve into more generally applicable formal theory, a research agenda needs to be set up gradually increasing their explanatory power beyond the present scope of experimentation behaviour conducted in pharmaceutical R&D. Also, in a second effort the proposed theory should be expanded to include the investigation of the role of problem definition, prior knowledge and cross-project learning (see Figure 5-13). Finally, the Bayesian belief network developed for the 'Concept Selection' part of the prescriptive framework is only a first step towards an integrated simulation-based methodology to study the effect of R&D policy variables on predictive and business performance during the various complexity-handling modes.

Therefore, in the following I will argue for a future research agenda building upon this work consisting of a confirmatory path deepening and widening the application reach of the developed theory in respectively pharmaceutical and other technology-intensive sectors, an exploratory path extending the reach to innovation problem definition in pharmaceutical discovery, and a quantitative operations research path further exploring the role of predictability on R&D and business performance.

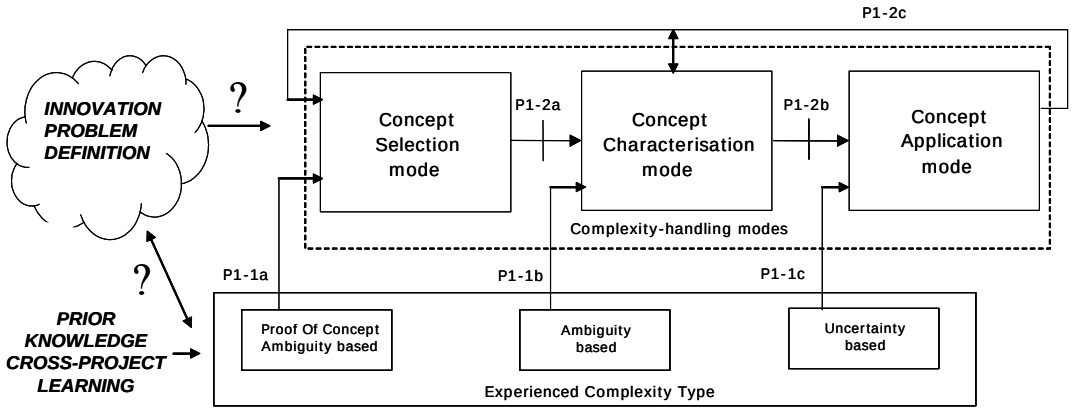


Figure 5-13 Areas of further research

First, although firmly grounded in the seven cases of this study, the present prescriptive framework needs to be replicated to other technology-intensive contexts to increase its external and internal validity. Since my exploratory results could only be partially replicated in the confirmatory case of pharmaceutical discovery, a case study-based *confirmatory research path* would enhance internal and construct validity by replicating the exploratory findings of the 'Concept Characterisation' and 'Concept Application' parts of the framework to other pharmaceutical and biotech R&D companies. After this, the prescriptive framework will be replicated to other research-intensive sectors like electronics, aerospace, or software development enhancing its external validity and possibly refining its propositions and constructs.

Secondly, a case study-based *exploratory research path* could extend the prescriptive framework to explore the role of experienced complexity on complexity-handling in the problem definition domain prior to problem-solving. This path is not only theoretically relevant. It is even essential for the future of managerial practice in pharmaceutical R&D since here problem definition can only be enhanced by a better understanding of disease pathways leading to more 'druggable' hence more potentially business relevant biological targets. As mentioned by an influential discovery practitioner; 'The attempt to replace the quality of scientific arguments by the sheer quantity of data as expressed in HTS or ultra-HTS[49] in the past has failed. An approach that is based on a much broader understanding of biochemical and genetic mechanisms of diseases appears to represent the necessary correction' (Drews, 2003: 416). How and why to come to this better understanding? These should be the research questions guiding an exploratory research agenda into pharmaceutical fuzzy front-end experimentation behaviour for target identification and validation.

Also, it is generally (Leonard-Barton, 1992; 1995; Verganti, 1997) and specifically (Henderson and Cockburn, 1994; Henderson, 1994a; 1994b; Horrobin, 2003; Duyck, 2003) acknowledged that increased disease understanding and economies of scope (Henderson and Cockburn , 1996) can only be realised by sustained capability building, which implies the need for leveraging prior knowledge and cross-project learning mechanisms. This implies that further case study-based research should investigate the role of corporate memory mechanisms underlying experimentation behaviour. However, as mentioned before in the previous Chapter, in this study I modelled a memory-less system to limit the complexity of the simulation model to emulate the adaptive search & optimization process. Hence, future simulation models supporting the exploratory research agenda should clearly have the capability to emulate the role of memory on experimentation behaviour and decision-making. A first step to explore this field was taken (Smart et al. 2004).

Finally, a quantitative *operations research path* should focus on the pivotal role predictability will play in managing pharmaceutical research and development. Since a ten percent improvement in predicting failures before clinical trials could save $100

[49] What is meant here is that increasing candidate compound libraries to be tested in High Throughput Screens against multiple poorly understood biological targets is not going to help to quickly and efficiently identify and validate targets. This view is supported by various influential publications (see a.o. (Horrobin, 2003; Duyck, 2003)

million in development costs per drug (Food and Drug Administration, 2004) it will be clear that managing predictive performance of R&D strategies will be high on any life sciences' company agenda. Also, from an academic perspective, further research in this area will lead to increased understanding of the impact of various R&D policy variables on predictive and business performance.

This thesis produced an internally valid simulation model emulating in a simplified mechanistic way the noisy search and optimization process followed by pharmaceutical discovery scientists to select a candidate chemical structure for subsequent clinical development. Further case-based process research could focus on studying external validity of the propositions resulting from the theory-generating simulation study. Although it should be noted that the model's predictions on negative performance are not empirically verifiable since the conditional probability involved cannot be observed in practice, at least partial empirical validation of the model's predictions can be performed in longitudinal studies of pharmaceutical company's product portfolio management.

However, considering the complexities of the discovery process I tend to apply the reasoning developed by Masuch and Lapotin (1989) that moving in one sweep from an internally valid but extremely simplified model of reality to the empirical reality of the real world is just too complex and bound to be problematic. Therefore, I follow their suggestion to first build more specific versions of the model to test the robustness of the operational assumptions made. Simulation results of these specific models may then lead to modifications or additions to the initial one developed in this thesis, which then in its refined form can be used for empirical validation.

Applying this thinking to the realities of the business needs in a life sciences R&D context I would propose to investigate the design of two specific models focussing on the transition between 'Concept Selection' and 'Concept Characterisation'. This being driven by the acknowledgement that 'we are falling woefully short of defining clear chains of causality that would effectively "link genetics to physiology" in a manner that could form the basis for robust, reliable models of complex biological processes' (Duyck, 2003: 604). In a first model I would extend the Bayesian belief network defined above to include policy variables used in 'Concept Characterisation' and study their effect on the tightness of the causal chain and its predictive and business performance. A second specific model could start anew and use a different process emulation engine. One possibility is to leave the present idea of emulating the discovery process as a noisy hierarchical search at successive levels of disaggregation, and build instead a "hierarchy machine" combining a genetic algorithm search process with Bayesian conditional logic as recently suggested by Pelikan and Goldberg (2003). If both newly built models of the discovery process provide results corroborating the initial model's results, then literal replication can be claimed, which further improves the model's internal validity and increases the robustness of its operational assumptions, making it now ready for empirical validation.

A final avenue of further operations research is given by the need acclaimed by the US Food and Drug Administration for applied research in the "industrialization" process to increase predictability of its results (Food and Drug Administration, 2004).

The latter process translates a characterized concept into a manufacturable product ready for mass customization. Here, the model would have to include policy variables in 'Concept Application' that potentially influence the predictability of the design process hence reducing the risk of failure in pharmaceutical production.

5.5 A FINAL CONCLUSION

David Horrobin (2003) uses an intriguing metaphor to provide a strategic perspective on the future of discovery and development in the life sciences industry. In his review paper he compares modern biomedical and pharmaceutical research with the Castalian 'glass bead game' as described in Herman Hesse's (1943) book of an isolated fantasy world called Castalia that recruits the brightest young people, educates them outstandingly, and persuades them that the ultimate achievement of the human mind is to play the complex, fully refined and internally consistent glass bead game. The only problem is that the Castalian game makes no contribution to the issues of the real world. In effect, the Castalians avoid 'wasting' their brains on real issues.

The fundamental issue of modern pharmaceutical discovery research is its lack of congruence with the real world of medical illness. 'The charge that we may be building a vast and internally consistent medical research game that has lost touch with patients is a serious one, and deserves serious attention'…'What needs to be done to reduce the risk of isolated self-consistency?' (Horrobin, 2003: 153). Duyck responds; 'in summary, an emerging challenge for life science research is to unify the fields of genetics and physiology, resulting in a more comprehensive and predictable picture of biology while enhancing the translational research process. The current lack of predictability not only represents a deficit in our knowledge base, but results in substantial opportunity cost, increased financial cost for therapeutic development, and limits on the potential impact of our basic research enterprise on public health' (Duyck, 2003: 605).

Therefore, the main contribution of my thesis to this debate is situated in its endogenous focus on the process of emerging innovative *solution understanding* and on the role of the latter on *predictability*; a performance variable shown to be not only relevant to the scientific debate but that also should be taken into consideration when studying business performance of research and development operations.

More specifically, it was argued that studying project management of innovation projects necessitates *an endogenous look* focusing on complexity-handling. A shift is needed from the exogenous view of various prescriptive project management models starting from the nature of the innovative product to an endogenous view of the innovation process. A prescriptive framework resulted from this initial work, dynamically connecting experienced complexity to complexity-handling modes. The latter build an emergent mental model of the innovative solution and continuously reduce its residual ambiguity before transferring it into the market.

Furthermore, a propositional model was derived using top-down simulation, relating fuzzy front-end policy variables to predictive and business performance thereby *extending the concept of front-loading in scope and breadth.* Not only did the discovery case show the relevance of front-loading as a concept underlying pharmaceutical fuzzy front-end policy enhancing solution understanding. Simulation results also quantified its positive influence on predictive performance.

The innovation literature has neglected so far *predictive performance as a performance indicator* for experimentation policies or innovation strategies. To my knowledge, this thesis was the first to provide a method to quantitatively compare fuzzy front-end experimentation policies on their predictive and resulting business performance.

As a contribution to practice, this thesis not only developed the prescriptive framework set out above guiding innovation managers to dynamically choose an approach for their projects. It also developed a solution-oriented research product under the form of a Bayesian belief network defined as a technological rule to evaluate business performance of company-specific configurations of policy variables defining fuzzy front-end experimentation.

Concluding, and looping back to the Castalian metaphor above one could say that this thesis was about researching ways to avoid playing internally consistent discovery games with no relation to practice. Instead, to improve the translation of the solution concept to practice, tools and concepts were developed to study the predictive performance of the connection to reality. Future research, then, should predominantly turn its attention from the solution side to the innovation problem definition side, focussing on the business relevance of better understanding human disease, connecting management science to the natural sciences.

REFERENCES

Abernathy, W. and Clark, K.B. (1985) Innovation: Mapping the winds of creative destruction. *Research Policy* **4**, 3-22.

Abernathy, W. and Rosenbloom, R. (1968) Parallel and sequential R&D strategies: Application of a simple model. *IEEE Transactions on Engineering Management* **EM 15(1)**, 2-10.

Abetti, P.A. (2000) Critical success factors for radical technological innovation: A five case study . *Creativity and Innovation Management* **9(4)**, 208-221.

Adler, P. (1995) Interdepartmental interdependence and coordination: The case of the design/manufacturing interface. *Organization Science* **6(2)**, 147-167.

Allen, P.M. (2001a) A complex systems approach to learning in adaptive networks. *International Journal of Innovation Management* **5(2)**, 149-180.

Allen, P.M. (2001b) The dynamics of knowledge and ignorance: Learning the new systems science. In: Matthies, M., Malchow, H. and Kriz, J., (Eds.) *Integrative approaches to natural and social dynamics*, Berlin: Springer Verlag.

Allen, P.M. (2001c) Knowledge, ignorance, and learning. *Emergence* **2(4)**, 78-103.

Allen, T.J. (1966) Studies of the problem-solving process in engineering design. *IEEE Transactions on Engineering Management* **EM-13(2)** , 72-83.

Allen, T.J., Tushman, M.L. and Lee, D.M.S. (1979) Technology transfer as a function of position in the spectrum from research through development and technical services. *Academy of Management Journal* **22(4)**, 694-708.

Anderson, Ph., Meyer, A., Eisenhardt, K.M., Carley, K. and Pettigrew, A.M. (1999) Introduction to the special issue: Applications of complexity theory to organization science. *Organization Science* **10(3)**, 233-236.

Arlington, S., Barnett, S., Hughes, S. and Palo, J. (2002) Pharma 2010: The threshold of innovation. *IBM Business Consulting Services Futures Series*.

Baker, W.E. and Sinkula, J.M. (2002) Market orientation, learning orientation and product innovation: Delving into the organization's black box. *Journal of Market-focused Management* **5(1)**, 5-23.

Barrett, F.J. (1998) Creativity and improvisation in Jazz and organizations: Implications for organizational learning. *Organization Science* **9(5)**, 605-622.

Basa, F. (1996) Project management and organization structure in drug development. *Drug Information Journal* **30**, 621-636.

Bergeron, B. (2003) *Bioinformatics computing*, Upper Saddle River, NJ: Prentice Hall.

Bhaskar, R. (1986) *Scientific realism and human emancipation*, London: Verso.

Bielza, C., Müller, P. and Rios-Insua, D. (1999) Decision analysis by augmented probability simulation. *Management Science* **45(7)**, 995-1007.

Blaikie, N. (1993) *Approaches to social enquiry*, Cambridge: Polity Press.

Boisot, M. (1995) *Information space: A framework for learning in organizations, institutions and culture*, London: Routledge.

Boisot, M. and Child, J. (1999) Organizations as adaptive systems in complex environments: The case of China. *Organization Science* **10(3)** , 237-252.

Burgelman, R.A. and Maidique, M.A. (1988) *Strategic management of technology and innovation*, Boston, MA: Irwin.

Callan, R. (2003) *Artificial intelligence*, New York: Palgrave Macmillan.

Campbell, D.J. (1988) Task complexity: A review and analysis. *Academy of Management Review* **13(1)**, 40-52.

Carrero, V., Peiro, J.M. and Salanova, M. (2000) Studying radical organizational innovation through grounded theory. *European Journal of Work and Organizational Psychology* **9**(**4**), 489-514.

Case, R.H. (1998) The structure of high-performing project management organizations. *Drug Information Journal* **32**, 577-607.

Cavalla, D. (1997) *Modern strategy for preclinical pharmaceutical R&D: Towards the virtual research company*, Chichester: John Wiley & Sons.

Chandy, R.K. and Tellis, G.J. (1998) Organizing for radical product innovation. *Journal of Marketing Research* **35**, 474-487.

Cheng, Y.T. and Van de Ven, A.H. (1996) Learning the innovation journey: Order out of chaos. *Organization Science* **7(6)**, 593-614.

Chesbrough, H. (2003a) The era of open innovation. *Sloan Management Review* **(Spring)**, 35-41.

Chesbrough, H. (2003b) *Open innovation: The new imperative for creating and profiting from technology*, Boston, MA: Harvard Business School Press.

Chesbrough, H. (2004) Managing open innovation. *Research and Technology Management* **(January-February)**, 23-26.

Christensen, C.M. (1997) *The innovator's dilemma: When new technologies cause great firms to fail*, Boston, MA: Harvard Business School Press.

Christensen, C.M., Johnson, M.W. and Rigby, D.K. (2002) Foundations for growth: How to identify and build disruptive new business. *Sloan Management Review* **43(Spring)**, 22-31.

Christensen, C.M. and Overdorf, M. (2000) Meeting the challenge of disruptive change. *Harvard Business Review* **(March-April)**.

Christensen, C.M. and Raynor, M.E. (2003) *The innovator's solution: Creating and sustaining successful growth*, Boston, MA: Harvard Business School Press.

Christensen, C.M. and Rosenbloom, R. (1995) Explaining the attacker's advantage: Technological paradigms, organizational dynamics, and the value network. *Research Policy* **24(2)**, 233-257.

Cilliers, P. (1998) *Complexity and post-modernism: Understanding complex systems*, London: Routledge.

Cohen, M.D., March, J.G. and Olsen, J.P. (1972) A garbage can model of organizational choice. *Administrative Science Quarterly* **17(1)**, 1-25.

Congdon, P. (2001) *Bayesian statistical modelling*, Chichester: John Wiley & Sons.

Cooper, R.G. (1990) Stage-gate systems: A new tool for managing new products. *Business Horizons* **33**, 44-54.

Cooper, R.G. (1994) Third-generation new product processes. *Journal of Product Innovation Management* **11**, 3-14.

Coty, C. (2002) Computational tools that identify drug likeness. *Drug Discovery & Development* **(September)**, 67-72.

Covaliu, Z. and Oliver, R. (1995) Representation and solution of decision problems using sequential decision diagrams. *Management Science* **41(12)**, 1860-1881.

Cox, B. (1989) Strategies for drug discovery: Structuring serendipity. *Pharmaceutical Journal* , 329-337.

Critchfield, G.C. and Willard, K.E. (1986) Probabilistic analysis of decision trees using Monte Carlo simulation. *Medical Decision Making* **6(2)**, 85-92.

Cyert, R.M. and March, J.G. (1992) *A behavioral theory of the firm*, Second edn. Oxford: Blackwell Publishers Ltd.

Dahan, E. and Mendelson, H. (2001) An extreme-value model of concept testing. *Management Science* **47(1)**, 102-116.

De Meyer, A., Loch, C.H. and Pich, M.T. (2001) A framework for project management under uncertainty. *INSEAD Working Paper*.

De Meyer, A., Loch, C.H. and Pich, M.T. (2002) Managing project uncertainty: From variation to chaos. *Sloan Management Review* **43(Winter)**, 60-67.

Devaney, R.L. (1989) *An introduction to chaotic dynamical systems*, Reading, MA.: Addison-Wesley.

DeWitte, R.S. (2002) On experimental design in drug discovery. *Current Drug Discovery* **(February)**, 19-22.

Doty, H.D. and Glick, W.H. (1994) Typologies as a unique form of theory building: Toward improved understanding and modeling . *Academy of Management Review* **19(2)**, 230-251.

Drews, J. (1998) Innovation deficit revisited: Reflections on the productivity of pharmaceutical R&D. *Drug Discovery Today* **3**, 491-494.

Drews, J. (2000) Drug Discovery: A historical perspective. *Science 287* **(March)**, 1960-1964.

Drews, J. (2003) Strategic trends in the drug industry. *Drug Discovery Today* **8(9)**, 411-420.

Durand, T. (1992) Dual Technology Trees: Assessing the intensity and strategic significance of technological change. *Research Policy* **21**, 361-380.

Duyck, J. (2003) Attrition and translation. *Science* **302(October 24)**, 603-605.

Eisenhardt, K.M. (1989) Building theories from case study research. *Academy of Management Review* **14(4)**, 532-550.

Eisenhardt, K.M. and Tabrizi, B.N. (1995) Accelerating adaptive processes: Product innovation in the global computer industry. *Administrative Science Quarterly* **40**, 84-110.

Ekvall, G. (1997) Organizational conditions and levels of creativity. *Creativity and Innovation Management* **6(4)**, 195-205.

Flood, R.L. (1987) Complexity: A definition by construction of a conceptual framework. *Systems Research* **4(3)**, 177-185.

Fonseca, J. (2002) *Complexity and innovation in organizations*, London: Routledge.

Food and Drug Administration (2004) Innovation stagnation: Challenge and opportunity on the critical path to new medical products. *FDA Report* **(March)**.

Garud, R. and Van de Ven, A.H. (1992) An empirical evaluation of the internal corporate venturing process. *Strategic Management Journal* **13**, 93-109.

Gassmann, O. and von Zedtwiz, M. (2003) Trends and determinants of managing virtual R&D teams. *R&D Management* **33(3)**, 243-262.

Gell-Mann, M. (1995) *The quark and the jaguar: Adventures in the simple and the complex*, New York: Abacus.

Gerwin, D. and Barrowman, N.J. (2002) An evaluation of research on integrated product development. *Management Science* **48(7)**, 938-953.

Getz, K.A. and De Bruin, A. (2000) Breaking the development speed barrier: Assessing successful practices of the fastest drug development companies. *Drug Information Journal* **34**, 725-736.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P. and Trow, M. (1994) *The new production of knowledge: The dynamics of science and research in contemporary societies*, London: Sage.

Gilbert, N. and Doran, J. (1994) *Simulating societies: The computer simulation of social phenomena*, London: UCL Press Ltd.

Glaser, B.G. and Strauss, A. (1967) *The discovery of grounded theory: Strategies for qualitative research* , Chicago: Aldine.

Goldberg, D.E. (1989) *Genetic Algorithms in search, optimization and machine learning*, Boston: Addison-Wesley.

Goldberg, D.E. (2000) The design of innovation: Lessons from genetic algorithms, lessons for the real world. *Technological Forecasting and Social Change* **64(1)**, 7-12.

Grandori, A. (1997) An organizational assessment of interfirm coordination modes. *Organization Studies* **18(6)**, 897-925.

Hann, M.M., Leach, A.R. and Harper, G. (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. *Journal of Chemical Information and Computer Sciences* **41**, 856-864.

Hatch, M.J. and Tsoukas, H. (1997) Complex thinking about organizational complexity: The appeal of a narrative approach to complexity theory. *Warwick Business School Research Bureau* **(No. 251)**.

Henderson, R.M. (1994a) The evolution of integrative capability: Innovation in cardiovascular drug discovery. *Industrial and Corporate Change* **3**, 607-630.

Henderson, R.M. (1994b) Managing innovation in the information age. *Harvard Business Review* **(January-February)**, 100-105.

Henderson, R.M. and Clark, K.B. (1990) Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* **35**, 9-30.

Henderson, R.M. and Cockburn, I. (1994) Measuring competence? Exploring firm effects in pharmaceutical research. *Strategic Management Journal* **15**, 63-84.

Henderson, R.M. and Cockburn, I. (1996) Scale, scope and spillovers: The determinants of research productivity in Drug Discovery. *Rand Journal of Economics* **27(1)**, 32-59.

Hesse, H. (1943) *Das Glasperlenspiel*, Zürich: Fretz und Wasmuth Verlag.

Higgs, R.E., Bemis, K.G. and Wikel, J.H. (1997) Experimental designs for selecting molecules from large chemical databases. *Journal of Chemical Information and Computer Sciences* **37**.

Holland, J. (1992) *Adaptation in natural and artificial systems*, Cambridge, Massachusetts: MIT Press.

Holland, J.H. (1998) *Emergence: From chaos to order*, Oxford: Oxford University Press.

Hopkins, A. (2003) Differentiating Inherent Risks Between Novel Targets. In: *Intelligent Drug Discovery & Development Conference 2003*, Philadelphia, PA.

Horrobin, D.F. (2003) Modern biomedical research: An internally self-consistent universe with little contact with medical reality? *Nature Reviews Drug Discovery* **2(February)**, 151-154.

Howard, R.A. (1988) Decision analysis: Practice and promise. *Management Science* **34(6)**, 679-695.

Iansiti, M. (1998) *Technology integration: Making critical choices in a dynamic world* , Boston, MA: Harvard Business School Press.

Iansiti, M. and Clark, K.B. (1994) Integration and dynamic capability: Evidence from product development in automobiles and mainframe computers. *Industrial and Corporate Change* **3(3)**, 557-605.

Jacobson, M.J. (2001) Problem solving, cognition and complex systems: Differences between experts and novices. *Complexity* **6(3)** , 41-49.

Jensen, F.V. (2001) *Bayesian networks and decision graphs*, New York: Springer-Verlag.

Johnson, J. (2000) The "Can you trust it?" problem of simulation science in the design of sociotechnical systems. *Complexity* **6(2)** , 34-40.

Kaufmann, S.A. (1993) *The origins of order: Self-organization and selection in evolution*, Oxford: Oxford University Press.

Kennedy, T. (1997) Managing the drug discovery/development interface. *Drug Discovery Today* **2(10)**, 436-444.

Khurana, A. (1999) Managing complex production processes. *Sloan Management Review* **(Winter)**, 85-97.

Khurana, A. and Rosenthal, S.R. (1997) Integrating the fuzzy front end of new product development. *Sloan Management Review* **(Winter)**, 103-120.

Khurana, A. and Rosenthal, S.R. (1998) Towards holistic "front-ends" in new product development. *Journal of Product Innovation Management* **15**, 57-74.

Kiel, L.D. (1991) Lessons from the non-linear paradigm: Applications of the theory of dissipative structures in the social sciences. *Social Science Quarterly* **72(3)**, 431-442.

Kim, J. and Wilemon, D. (2002) Focusing the fuzzy front-end in new product development. *R&D Management* **32(4)**, 269-279.

Kim, J. and Wilemon, D. (2003) Sources and assessment of complexity in NPD projects. *R&D Management* **33(1)**, 15-30.

Klaes, M. (1997) Sociotechnical constituencies, game theory, and the diffusion of compact discs: An interdisciplinary investigation into the market for recorded music. *Research Policy* **25(8)**, 1221-1234.

Kleijnen, J.P.C., Cheng, R.C.H. and Bettonvil, B. (2001) Validation of trace-driven simulation models: Bootstrap tests. *Management Science* **47(11)**, 1533-1538.

Koput, K.W. (1992) Dynamics of innovation idea generation in organizations: Randomness and chaos in the development of a new medical device. *Unpublished PhD dissertation,* School of Business, University of California at Berkeley.

Krishnan, V., Eppinger, S.D. and Whitney, D. (1997) A model based framework to overlap product development activities. *Management Science* **43**.

Kubinyi, H. (2003) Drug research: Myths, hype and reality. *Nature Reviews Drug Discovery* **2**, 665-668.

Kuhn, T.S. (1996) *The structure of scientific revolutions*, 3rd edn. Chicago: The University of Chicago Press.

Langley, A. (1999) Strategies for theorizing from process data. *Academy of Management Review* **24(4)**, 691-710.

Lee, J., Boatwright, P. and Kamakura, W.A. (2003) A Bayesian model for prelaunch sales forecasting of recorded music. *Management Science* **49(2)**, 179-196.

Lee, M. and Na, D. (1994) Determinants of technical success in product development when innovation radicalness is considered. *Journal of Product Innovation Management* **11**, 62-68.

Leifer, R., McDermott, C.M., O'Connor, G.C., Peters, L.S., Rice, M. and Veryzer, R.W. (2000) *Radical innovation: How mature companies can outsmart upstarts*, Boston, MA: Harvard Business School Press.

Leonard-Barton, D. (1992) Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal* **13**, 111-125.

Leonard-Barton, D. (1995) *Wellsprings of knowledge: Building and sustaining the sources of innovation*, Boston, MA: Harvard Business School Press.

Lesko, L.J., Rowland, M., Peck, C.C. and Blaschke, T.F. (2000) Optimizing the science of drug development: Opportunities for better candidate selection and accelerated evaluation in humans. *Pharmaceutical Research* **17(11)**, 1335-1344.

Loch, C.H., Pich, M.T. and De Meyer, A. (2000) Project uncertainty and management styles. *INSEAD Working Paper 2000/31/TM/CIMSO 10.*.

Loch, C.H. and Terwiesch, C. (1998) Communication and uncertainty in concurrent engineering. *Management Science* **44(8)**, 1032-1048.

Loch, C.H., Terwiesch, C. and Thomke, S. (2001) Parallel and sequential testing of design alternatives. *Management Science* **47(5)**, 663-678.

Lynn, G.S. and Akgun, A.E. (2001) Project visioning: It's components and impact on new product success. *Journal of Product Innovation Management* **18(ER6)**, 374-387.

Lynn, G.S., Morone, J.G. and Paulson, A.S. (1996) Marketing and discontinuous innovation: The Probe and Learn process. *California Management Review* **38(3)**, 8-37.

MacCormack, A. and Verganti, R. (2003) Managing the sources of uncertainty: Matching process and context in software development. *Journal of Product Innovation Management* **20**, 217-232.

March, J.G. (1978) Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics* **9**, 587-608.

March, J.G. (1991) Exploration and exploitation in organizational learning. *Organization Science* **2(1)**, 71-87.

Masuch, M. and Lapotin, P. (1989) Beyond garbage cans: An AI model of organizational choice. *Administrative Science Quarterly* **34(1)**, 38-67.

McDonough, E.F.I. and Barczak, G. (1992) The effects of cognitive problem-solving orientation and technological familiarity on faster new product development. *Journal of Product Innovation Management* **9(1)**, 44-52.

McKee, D. (1992) An organizational learning approach to product innovation. *Journal of Product Innovation Management* **9**, 232-245.

Mena, C., Thomson, A. and Jeffrey, P. (2001) HS Marston report: An evolutionary model of the product design process. *Innovative Manufacturing Initiative: The complexities of product definition*.

Mihm, J., Loch, C.H. and Huchzermeier, A. (2003) Problem-solving oscillations in complex engineering projects. *Management Science* **46(6)**, 733-750.

Mitchell, M. (2001) *An introduction to genetic algorithms*, Cambridge, Masachusetts: MIT Press.

Molina, A. (1990) Transputers and transputer-based parallel computers: Sociotechnical constituencies and the build-up of British-European capabilities in information technologies. *Research Policy* **19**, 309-333.

Montgomery, D.C. (1997) *Design and analysis of experiments*, John Wiley & Sons.

Müller, P. (1999) Simulation-based optimal design. In: Bernardo, J.M., Berger, J.O., David, A.P. and Smith, A.F.M., (Eds.) *Bayesian Statistics 6*, pp. 459-474. Oxford: Oxford University Press.

Newell, A. and Simon, H.A. (1976) Computer science as empirical enquiry: Symbols and search. *Communications of the ACM* **19(3)**, 113-126.

O' Connor, G.C. (1998) Market learning and radical innovation: A cross-case comparison of eight radical innovation projects. *Journal of Product Innovation Management* **15(2)**, 151-166.

Oprea, T.I. (2002) Virtual screening in lead discovery: A viewpoint. *Molecules* **7**, 51-62.

Oprea, T.I., Davis, A.M., Teague, S.J. and Leeson, P.D. (2001) Is there a difference between leads and drugs? A historical perspective. *Journal of Chemical Information and Computer Sciences* **41**, 1308-1315.

Parmigiani, G. (2002) *Modeling in medical decision making: A Bayesian approach*, Chichester: John Wiley & Sons.

Partington, D. (2000) Building grounded theories of management action. *British Journal of Management* **11**, 91-102.

Pearl, J. (2000) *Causality: Models, reasoning, and inference*, Cambridge: Cambridge University Press.

Pelikan, M. and Goldberg, D.E. (2003) A hierarchy machine: learning to optimize from nature and humans. *Complexity* **8(5)**, 36-45.

Pich, M.T., Loch, C.H. and De Meyer, A. (2002) On uncertainty, ambiguity and complexity in project management. *Management Science* **48(8)**, 1008-1023.

Pickering, L. (2001) ADME/Tox models can speed develoment. *Drug Discovery & Development* **(January)**, 34-38.

Pidd, M. (1992) *Computer simulation in management science*, 3rd edn. New York: John Wiley & Sons.

Pisano, G.P. (1994) Knowledge, integration, and the locus of learning: An empirical analysis of process development. *Strategic Management Journal* **15**, 85-100.

Pisano, G.P. (1996) Learning before doing in the development of new process technology. *Research Policy* **25(7)**, 1097-1119.

Pisano, G.P. (1997) *The development factory: Unlocking the potential of process innovation*, Boston, MA: Harvard Business School Press.

Poole, M.S., Van de Ven, A.H., Dooley, K. and Holmes, M.E. (2000) *Organizational change and innovation processes*, Oxford: Oxford University Press.

Popper, K.R. (1959) *The logic of scientific discovery*, London: Routledge.

Reiss, T. and Hinze, S. (2000) Innovation process and techno-scientific dynamics. In: Jungmittag, A., Reger, G. and Reiss, T., (Eds.) *Changing innovation in the pharmaceutical industry: Globalization and new ways of drug development*, Berlin: Springer –Verlag.

Roemer, T.A., Ahmadi, R.H. and Wang, R.H. (2000) Time-cost trade-offs in overlapped product development. *Operations Research* **48(6)**, 858-865.

Romme, A.G.L. (2003) Making a difference: Organization as design. *Organization Science* **14(5)**, 558-573.

Rumelhart, D.E., McClelland, J.L. and the PDP Research Group (1986) *Parallel distributed processes: Explorations in the microstructure of cognition. Volume 1: Foundations*, Cambridge, MA: A Bradford Book.

Sandberg, J. (2000) Understanding human competence at work: An interpretative approach. *Academy of Management Journal* **43(1)**, 9-25.

Schrader, S., Riggs, W.M. and Smith, R.P. (1992) Choice over uncertainty and ambiguity in technical problem solving. *Journal of Engineering and Technology Management* **10(1)**, 73-99.

Schroder, H., Driver, M. and Streufert, S. (1967) *Human information processing*, New York: Holt, Rinehart and Winston.

Schwenk, C. (1985) The use of participant recollection in the modeling of organizational processes. *Academy of Management Review* **10(3)** , 496-503.

Shenhar, A.J. and Dvir, D. (1996) Toward a typological theory of project management. *Research Policy* **25(4)**, 607-632.

Shenoy, P.P. (1992) Valuation-based systems for Bayesian decision analysis. *Operations Research* **40(3)**, 463-484.

Singh, K. (1997) The impact of technological complexity and interfirm cooperation on business survival. *Academy of Management Journal* **40(2)**, 339-367.

Smart, P., Gupta, A., Bessant, J., Peakman, T. and Van Dyck, W. (2004) Networking for innovation: Exploring a dynamic capabilities perspective. *R&D Management* **Special Issue on Innovation and Intangible Assets**, Extended abstract submitted.

Smith, G.F. (1988) Towards a heuristic theory of problem structuring. *Management Science* **34**, 1489-1506.

Smith, J.Q. (1989) Influence diagrams for Bayesian decision analysis. *European Journal of Operational Research* **40**, 363-376.

Smith, R. and Eppinger, S.D. (1997a) A prediction model of sequential iteration in engineering design. *Management Science* **43(8)**, 1104-1120.

Smith, R.P. and Eppinger, S.D. (1997b) Identifying controlling features of engineering design iterations. *Management Science* **43(3)**, 276-293.

Sobek II, D.K., Ward, A.C. and Liker, J.K. (1999) Toyota's principles of set-based concurrent engineering. *Sloan Management Review* **(Winter)**, 67-83.

Spirtes, P., Glymour, C. and Scheines, R. (2000) *Causation, prediction, and search*, Second edn. Cambridge, MA: The MIT Press.

Stacey, R. (1995) The science of complexity: An alternative perspective for strategic change processes. *Strategic Management Journal* **16**, 477-495.

Stacey, R.D. (2001) *Complex responsive processes in organizations: Learning and knowledge creation*, London: Routledge.

Stacey, R.D., Griffin, D. and Shaw, P. (2000) *Complexity and management*, London: Routledge Series on Complexity and Emergence in organisations.

Tatikonda, M.V. and Rosenthal, S.R. (2000) Technology novelty, project complexity, and product development execution success. *IEEE Transactions on Engineering Management* **47(1)**, 74-87.

Terwiesch, C., Loch, C.H. and De Meyer, A. (2002) Exchanging preliminary information in concurrent engineering: Alternative coordination strategies. *Organization Science* **13(4)**, 402-419.

Thomke, S. (1998a) Managing experimentation in the design of new products. *Management Science* **44(6)**, 743-762.

Thomke, S. (1998b) Simulation, learning and R&D performance: Evidence from automotive development. *Research Policy* **27**, 55-74.

Thomke, S. (2001) Enlightened experimentation: The new imperative for innovation. *Harvard Business Review* **(February)**, 67-75.

Thomke, S. (2003) *Experimentation matters: Unlocking the potential of new technologies for innovation*, Boston, MA: Harvard Business School Press.

Thomke, S. and Fujimoto, T. (2000) The effect of 'front-loading' problem-solving on product development performance. *Journal of Product Innovation Management* **17(2)**, 128-142.

Thomke, S., von Hippel, E. and Franke, R. (1998) Modes of experimentation: An innovation process - and competitive variable. *Research Policy* **27**, 315-332.

Tidd, J., Bessant, J. and Pavitt, K. (2001) *Managing innovation: Integrating technologicl, market and organizational change*, Second edn. Chichester: John Wiley & Sons.

Tollman P. (2001) A revolution in R&D. How genomics and genetics are transforming the biopharmaceutical industry. *The Boston Consulting Group*.

van Aken, J.E. (2004a) The field-tested and grounded technological rule as product of Mode 2 management research. *British Journal of Management,* Revised and resubmitted March 31, 2004.

van Aken, J.E. (2004b) Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules . *Journal of Management Studies* **41(2)**, 219-246.

Van de Ven, A.H. and Johnson, P.E. (2004) Knowledge for science and practice. *Academy of Management Review,* Revised and resubmitted March 10, 2004**.**

Van de Ven, A.H. and Polley, D. (1992) Learning while innovating. *Organization Science* **3(1)**, 92-116.

Van de Ven, A.H. and Poole, M.S. (1995) Explaining development and change in organizations. *Academy of Management Review* **20**, 510-540.

Van Dyck, W. (2004a) Introducing advanced technology in pharmaceutical R&D: The need for a new paradigm? In: Pacl, H., (Ed.) *The future of pharmaceutical R&D: Challenges and trends*, Frankfurt: Festel Capital, Forthcoming.

Van Dyck, W. (2004b) Managing complexity in fuzzy front-end innovation: The need for a paradigm shift? *Cranfield School of Management Working Paper Series*, paper submitted.

Van Dyck, W. (2004c) Predictive performance of front-loaded experimentation strategies in pharmaceutical discovery: A Bayesian perspective. *Cranfield School of Management Working Paper Series*, paper submitted.

van Fraassen, B.C. (1980) *The scientific image*, Oxford: Clarendon Press.

Verganti, R. (1997) Leveraging on systemic learning to manage the early phases of product innovation projects. *R&D Management* **27**, 377-392.

Verganti, R. (1999) Planned flexibility: Linking anticipation and reaction in product development projects. *Journal of Product Innovation Management* **16(4)**, 363-376.

Veryzer, R.W. (1998) Discontinuous innovation and the new product development process. *Journal of Product Innovation Management* **15(July)**, 304-321.

von Foerster, H. (1973) Cybernetics of cybernetics (physiology of revolution). *The Cybernetician* **3**, 30-32.

Weick, K.E. (1995) *Sensemaking in organizations*, London: Sage Publications.

West, J. and Iansiti, M. (2003) Experience, experimentation, and the accumulation of knowledge: The evolution of R&D in the semiconductor industry. *Research Policy* **32**, 809-825.

Wheelwright, S.C. and Clark, K.B. (1992) *Revolutionizing product development: Quantum leaps in speed, efficiency, and quality* , New York: The Free Press.

Whitney, D.E. (1990) Designing the design process. *Research in Engineering Design* **2**, 3-13.

Yang, K. and El-Haik, B. (2003) *Design for Six Sigma: A roadmap for product development*, New York: Mc Graw Hill.

Yin, R.K. (1989) *Case study research. Design and methods*, London: Sage Publications.

Young, S.S., Lam, R.L.H. and Welch, W.J. (2002) Initial compound selection for sequential screening. *Current Opinion in Drug Discovery & Development* **5(3)**, 422-427.

Young, S.S., Sheffield, C.F. and Farmen, M. (1997) Optimum utilization of a compound collection or chemical library for Drug Discovery. *Journal of Chemical Information and Computer Sciences* **37(5)**, 892-899.

Yu, H. and Adedayo, A. (2003) ADME-Tox in drug discovery: Integration of experimental and computational technologies. *Drug Discovery Today* **8(18)**, 852-861.

APPENDIX A: Exploratory case studies visual mapping and temporal bracketing results

Case statistics

Case study	*# empirical indicators*	*Respondents*	*# interviews*
DNA-Zyme	31	Marcus Brewster	2
Micro-emulsions	91	Tina Arien	3
Supercritical Fluids	38	Tina Arien Christina	2 1
Nanosuspension	23	Marc François	2
Gamma Controlled release	45	Marc Deweer	2
Alfa Immediate release	24	Marc Deweer	2
Totals		**5**	**14**

Case written sources

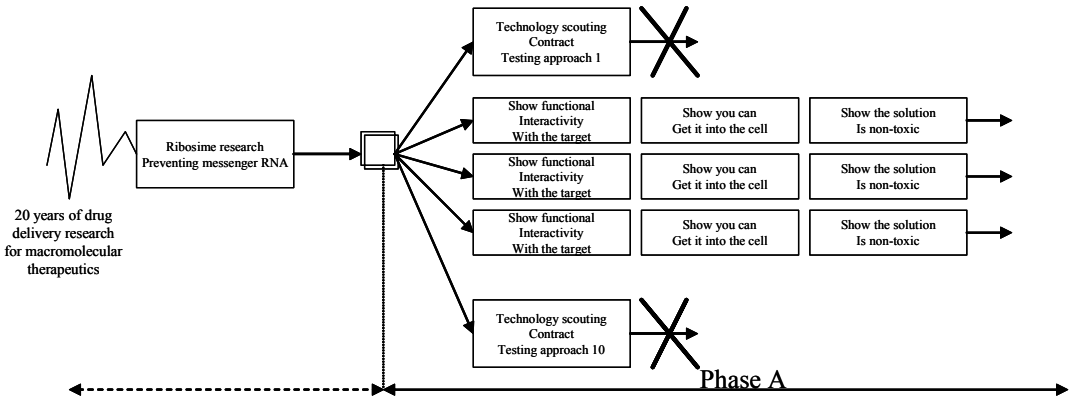
Sources used to document PharmaCo research context and case visual maps:

(1) Brewster, Peeters, Noppe, Arien (1997) 'The Drug Delivery Research Group: A component of pharmaceutical development, Group implementation, integration and research planning', internal confidential PharmaCo publication.
(2) PharmaCo (2001) 'Pharmaceutical Development facts & figures', internal publication.
(3) PharmaCo (2001) 'Global best practice: Direct compression formulation development flow chart', internal confidential publication.
(4) PharmaCo (2001) 'Project [Gamma] formulation development rationale', internal confidential publication.
(5) PharmaCo (1998) 'Use of supercritical fluid technology to broaden applicability of melt extrusion: Proposal for a research grant for funding from the Excellence in Science Award Program sponsored by the Corporate Office of Science and Technology', internal publication.
(6) PharmaCo (2002) 'SCF technology in combination with meltextrusion: Timeline of SCF activities', internal confidential presentation.
(7) PharmaCo (1998), 'Use of novel microemulsions for the formulation of hydrophobic drugs: Proposal for a research grant for funding from the Excellence in Science Award Program sponsored by the Corporate Office of Science and Technology', internal publication.
(8) PharmaCo (2000) 'Project Omega: DNAzymes – How do they work?', internal confidential publication.

(9) PharmaCo Molecular Oncology Group (2000) 'Penetration-mediated delivery of anti c-myc DNAzyme molecules in cell-based assays, proof of principle study proposal, internal confidential publication.

In the following tables empirical indicators will be indicated with a case name and number (Case name, x) at the end referring to the interview notes in the exploratory case studies database.

Case DNAzyme basic data

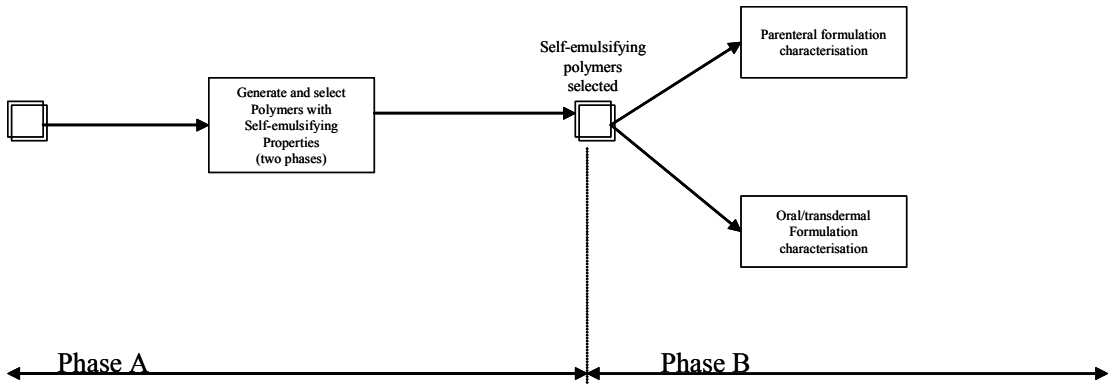


Approach dimensions	Applies to Phase	Relevant empirical indicators
Target setting	A	"I don't know if everybody does, but we took a stepped approach, because first you have to show it interacts with the target, than you have to show you can get it into the cells, then you have to show you have a non-toxic material , then you have to show you get one way or another some activity in an in-vivo system , then you have to have a number of curtails because this area is fought with problems, coming from that very first experiment, where you show some functional activity. (DNAZyme, 7)
	A	"What Michel [project manager] tried to do is to set up, even before we started looking at these people, what we like to see for the very first criteria; that was functional DNA-Zyme activity in the cells" (DNAZyme, 22)
	A	"…I think the idea was really; let's identify these criteria [solution critical requirements] We have the opportunity to look at our range of technologies, but having said that, the very first step is to get rid of as many as possible. What would happen, if all of these passed the first test, I don't know. But we knew that the bar was so high, that it wasn't likely to happen." (DNAZyme, 31)
Experimentation	A	"We also went to the scientific literature and made a special point to specific scientific media. In the beginning there was a relatively long list, we looked at anything that might be useful; so at one point we must have had 10 or 12 different approaches" (DNAZyme, 4)
	A	"You knew that you had to look at some pretty avant-garde therapies for approaches" (DNAZyme, 19)
	A	"…It is so high risk you could never put all your eggs in the same basket. You never approach it by just assuming one unproven technology is going to be the best. You have to spread in as many reasons as possible…so Michel went to a number of companies in the area with the idea establishing a number of parallel feasibility contracts…we also tried to diversify risk by looking at all known

		possibilities of delivery" (DNAZyme, 21, 3, 5)
	A	"The trouble was that, since we felt that it was a difficult project, it had to be very highly parallelized in terms of how to move forward, […]you also have to have good criteria (DNAZyme, 23)
	A	"The more complicated the design space is, the more you tend to tailor your approach to answer a specific question, rather than learn as much as you can about that area that probably is not the case in a university, where off-course the goal of a university is something different of the goal of a pharmaceutical company in terms of generating space. It's just a realism that a pharmaceutical company is moving as quickly as possible to a space you can provide a product for" (DNAZyme, 25)
		That were the three, we thought were the best, that we could move foreword with and that we are currently interacting with (DNAZyme, 12)
Learning	A	"…It is always stepwise, but the difference was the extent to which we thought we could have some sort of input into the different processes.….[x] had more development time and there was also more technology we could help them with, [y] was brand new, they had a polymer, very little was known, and because of that we thought we had a more open playfield….we interacted most with them (DNAZyme, 9, 11, 10)
	A	"…There was so much published on so many failures, that we knew we could not do with the traditional approaches so we started already at the beginning with things that were pretty off the wall..." (DNAZyme, 17)
	A	" …We knew, that when we had any success, it would be from novel approaches and not from things that have been tried fifty times before". (DNAZyme, 18)
Coordination	A	"…Since we felt that it was a difficult project it had to be very highly paralleled in terms of how to move forward…and you have to have good criteria to focus as quickly as possible on those companies and institutions that were able to demonstrate that…" (DNAZyme, 23, 24)
	A	"So, there are two big pharma-ideas about this; one is that you sit back, and hope for the best to wait for somebody to give you something off the shelf or you can try to have some impact on their development, by going in early and trying to steer, manipulate the direction which is going to be most interesting and beneficial for you" (DNAZyme, 6).
	A	"…Everybody got the same message what was going to be the first milestone, what we wanted to see at that point. For those companies, that were either at the first milestone or that we had a good feeling about we broadened the topic to say, the next area will be this, and than after that it will be…always raising the hurdle…" (DNAZyme, 26)
		"..I don't know if everybody does, but we took a stepped approach with each of the candidate solutions…first they had to show you

	A	can get it into the cell, then you have to have some in-vivo activity, then…" (DNAZyme, 7) " We went through basically screening with all of these" (DNAZyme, 13).
	A	"…From the very beginning we always had in mind the stepped-approach, that if the first sets of experiments were positive, we would make commitment during the next step, if it would not have been positive we would have dropped it " (DNAZyme, 8).
		"…Here is something the project leader was convinced we had to do in a relatively systematic way…either we moved forward, either we stopped…in the end …that were the three we thought were the best, that we could move forward with and currently interacting with…" (DNAZyme, 12, 14) … "either we moved forward, or we stopped… but the data that came back as the most positive, were these three" (DNAZyme, 15, 16).

Case Micro-emulsions basic data

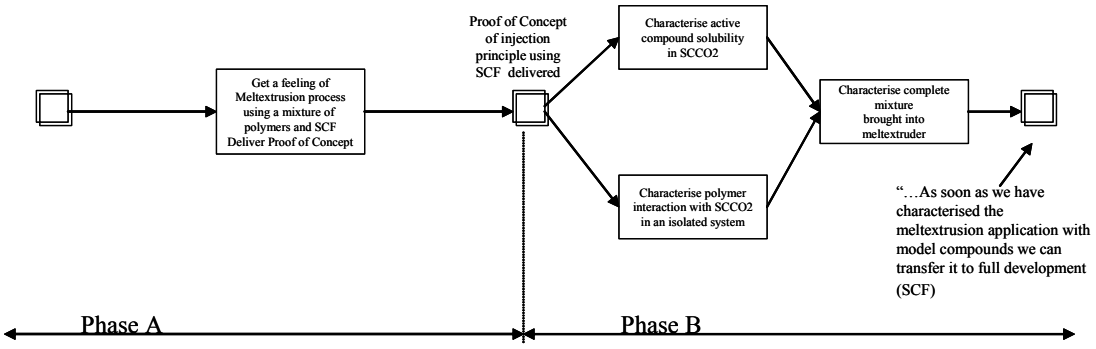


Approach dimensions	Applies to Phase	Relevant empirical indicators
Target setting	A	"…The objective of these first two phases [refers to project process chart] is really to check for entire families of polymers whether it is possible to form micro-emulsions and to encapsulate an active ingredient, or not". (Micro, 45)
	A	"First you check the functioning of the system as a whole or not, only then you take the next step. This is why we split up the project in different steps (Micro, 35)
Experimentation	A	"…[refers to project process chart] the process is really to check for each family of polymers if micro-emulsions can be formed, then if active substance can be encapsulated. If it doesn't work, go back, if it does, take another family and start all over again…. After this initial phase we will be able to fully characterise the formulation for parenteral and oral

		applications" [refers to project process chart] (Micro, 12)
	B	"…[refers to project process chart] then we would like to characterise parenteral and oral applications [in parallel] exploring how much active substance can be solved, whether a capsule can be filled, does the capsule break when it gets into contact with the acids in the gut? How about the toxicity and permeability for each of these applications…"(Micro, 11, 12)
	B	"…this project is actually a platform that can be used for different applications,…parenteral, oral, and transdermal" (Micro, 23, 25)
	B	"In the first phase the question to be answered is; is it stable or not? In this phase, the experimental plan will be more directed towards exploration of ranges of parameters –is this range a good range to function at production scale?- (Micro, 54)
	A,B	"In Phase A you select, in Phase B you model and optimize the polymer structure…" (Micro, 72)
	B	"Design of Experiments will also be easier to be used in this phase than in the previous because you now have a good idea which parameters are going to influence more or less strongly the process (Micro, 55)
	B	"The end point of this Phase B is really to have a technology available, sufficiently characterised to have a compound project use it" (Micro, 79)
Learning	A	"[Before starting]You check the literature from different perspectives and try to combine approaches without doing many experiments (Micro, 57)
Coordination	A,B	"…This is definitely a reciprocal collaboration contract we have with UCL [university] (Micro, 37)
	A	"…Based upon some first experiments we estimated the time needed to do synthesis and characterisation of the compounds, then, we estimated it would take so many months to synthesize so many polymers…." (Micro, 26) … a professor from UCL has also made an estimation, then we compared and saw we came to about the same results (Micro, 27)
	A,B	"The transition to the next phase [from A to B] is really dependent upon the results you get. And if you don't get a result you have to stop unless the literature provides you with further clues" (Micro, 50)
	B	"You develop an application on model compounds, not for a final R-number [specific drug project] We will only transfer into Full Development when we will be sure it will be bio-available, enhances solubility and is non-toxic. This needs to be shown first on model compounds (Micro, 61)
	B	"I think the transition into Full Development is made as soon as there's a compound in need for this technology in its relevant application domain (Micro, 63)

	A,B	"Now the technology is already sufficiently characterized for one type of polymer to allow us to test its solubility in micro-emulsions for one specific compound falling within this application domain" (Micro, 64) "Depending on the results it is possible to go back from Phase B to Phase A because we get results that do not allow us to go through" (Micro, 71)

Case Supercritical fluids basic data

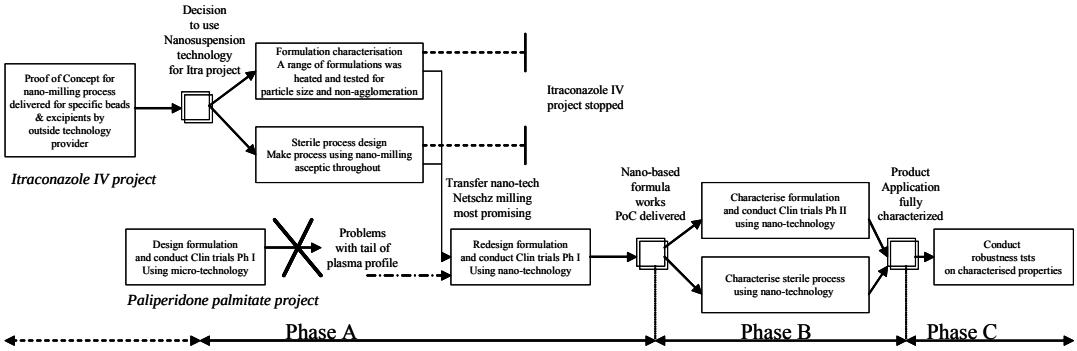


Approach dimensions	Applies to Phase	Relevant empirical indicators
Target setting	A	"We checked solubility of polymers in SCF and decided whether we could proceed or not. If it would not have been possible to solubilize polymers we could not have proceeded". (SCF, 24)
	B	"As soon as we find out if and how we can get the active substance into the SCF we will have to check how good the active can be solubilized, or which volume you can usefully bring into the extruder, what is the stability, bio-availability, of the mixture?"(SCF, 26)
	A	Comparing SCF to microemulsions' first screening phase: "For SCF there will also be a screening of classes of active substances and polymers for which SCF can be a solution. Which ones are feasible? Then a certain system will be selected and the most useful extruder parameters will be characterised". (SCF, 30)
	A	"For the meltextrusion application we stated as an objective that the temperature with SCF had to be lower as the temperature we would have had without SCF. And this for a pressure that would not be so high it would endanger the functioning of the meltextruder. (SCF, 38)

Experimentation	A	"…At first instance Christina [project manager] checked the solubility of the active substance in the supercritical fluid, in parallel she checked whether the polymer could be extruded using supercritical fluids…at the first experiments she saw it was not really possible to bring liquid CO2 into the extruder, so it had to be a gas form" (SCF, 7)
	A	"…Mid 2000 Proof of Concept was delivered in the Supercritical fluids project when polymers in interaction with $SCCO_2$ showed that they can decrease the meltextrusion temperature and that active substance can be made soluble in $SCCO_2$ under certain critical conditions" (SCF, 4)
	A	"…In [Phase A] … we tested both concepts on one polymer. Both single and twin screw worked but the latter gave the best blend properties so we chose the last one to show proof of concept (SCF, 32)… "In [Phase A] you check; is it feasible or not?" (SCF, 37)
	B	"…As soon as we have a bit of a view of the process [Phase A] we will use Design of Experiments to characterise and optimise the process (SCF, 9).
	A	"In this mode the question is much more; is it stable or not? Or, [for this parameter] is this a good working range to have the application work at full production scale? DoE will be easier in this phase [Phase B] while you have an idea which parameters will influence the process more or less. So you know the factors to do a proper DoE". (SCF, 27, 28)
	B	"At the beginning [in Phase A] it will be very difficult to select the right variables to do a DoE" (SCF, 29)
	A,B	"…[After parallel characterisation of both areas] then we could bring the characterised supercritical fluid with active substance into the characterised extrusion process, bringing the two together…if it works we can start modelling the nozzle, if it doesn't work we'll have to check whether we could use carbon solvents" (SCF, 14)
	B	"…First, we looked whether products of our pipeline could be brought into the extruder using supercritical fluids, then [in Phase B] we do a full physico-chemical characterisation of the whole extrusion system, later [Full Development] it will be scaled-up to full production level (SCF, 6)
		"…You start with determining under which conditions you can add SCCO2 to the polymer, then you look for the lowest possible temperature at which this can happen…you characterise parameters like speed of addition, pressure, single versus twin-screw geometry of the extrusion screw…" (SCF, 8)
Learning	A	"…The first year an intensive literature study was needed complemented with a lot of contacts outside JJPRD to get a grasp of the technology" (SCF, 22)
	A,B	"…The characterisation of the supercritical fluid-active substance interaction is done in house with problem-solving help of the University of [x], for the extrusion characterisation part we do this

		in collaboration with the University of [y]…. It is not a real collaboration contract we have with them, it's more consultancy, they give advise to us…no milestone agreements exist with them" (SCF, 17, 18, 19)
Coordination	B	"You split up the concept proven in [Phase A] into subparts that you will characterise" (SCF, 35)
	B	" In [Full Development] you check for a specific compound application or dosage form like a capsule, whether stability in the log run, in-vivo bio-availability, the whole gamut of parameters are ok (SCF, 36).

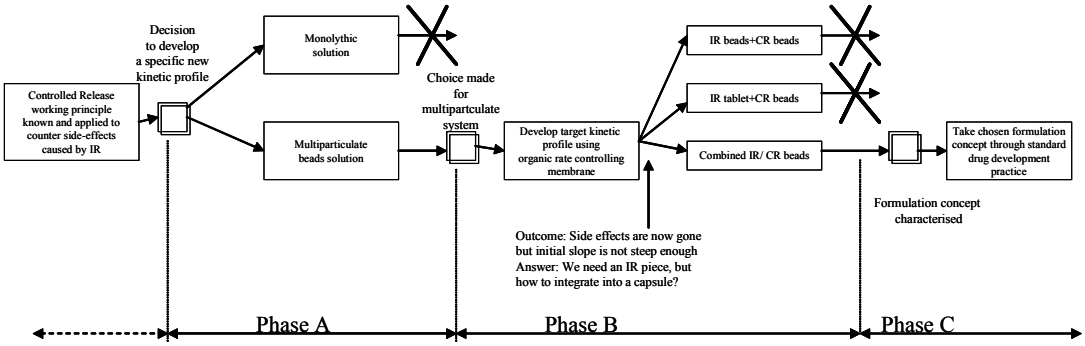
Case Nanosuspension basic data



Approach dimensions	Applies to Phase	Relevant empirical indicators
Target setting	A	"…At project start the basic Nanosuspension process technology was known. Our task was to see whether we could reproduce this process with our compounds at the level of clinical studies, and whether we could make it sterile…" (Nano, 1)
	A	"We tried to have a working product as quickly as possible" (Nano, 12)…then we had to change several excipients… then we solved the suspension problems…finally we had a working formula…this was the end of Concept Selection (Nano, 13)
Experimentation	B	"…We split up the work in two coherent design packages [Formulation design and Sterile process design] each exploring and characterizing part of the solution…" (Nano, 2)
	A	"…finally we had a working formula…this was the end of Concept Selection" (Nano, 13).
	B	"Then we started characterising the product…every parameter; particle size as a function of time, sedimentation as a function of time, etc. we checked for robustness" (Nano, 14)

	B	"Characterisation was done immediately on the project compound [not on model compound first] (Nano, 15)
Learning	A	"…Knowledge transfer happened mainly at the beginning, (Nano, 16)
	B	"… later we used the relationship only for problem solving, and it was not a joint development effort…" (Nano, 16)
	A,B	"… The basic technology we bought was the milling process and excipients/beads that gave this type of product this stability. We brought the sterile knowledge to the project" (Nano, 7)
Coordination	A,B	"Formulation and process were very much interwoven in this project" (Nano, 5)
	A,B	"During Phase A you select a concept, in the second step you optimize your concept" (Nano, 19, 20)

Case Gamma Controlled Release basic data

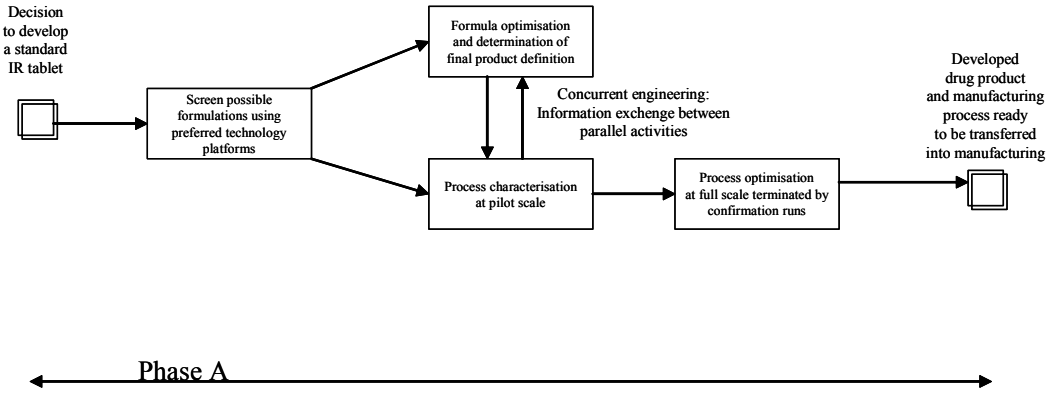


Approach dimensions	Applies to Phase	Relevant empirical indicators
Target setting	A	"The challenge was to develop a CR profile that would rise very quickly, stay up the whole day, and fall down in the evening so patients can sleep, all this without peaks to prevent side-effects"…(CR, 7) "The overall proof of concept for this was known. However, what was not known was how to get this significant quick rise at the beginning" (CR, 8)
	A	"After Phase A we had to have a working concept. It was actually a combination of selecting and characterizing a concept" (CR, 36)
	A	"At the end of Phase A you know you will have a working concept

		with a good kinetic profile that will also be manufacturable" (CR, 37) We selected a multiparticulate concept with an IR and CR piece (CR, 39) Then we said, we will be working using an aqueous solution and fine-tuned it to get a workable profile. That was Phase B or concept characterisation" (CR, 40)
	B	"In Phase B you are discovering ranges. The driving factor is kinetics. You know after Phase A which kinetic profile you need. How you will get it you discover in Phase B" (CR, 38)
	A,B	"Target setting in concept characterisation [Phase A and B] is done in a domain you don't know. You don't know what the value ranges of the variables are" (CR, 42)
	A,B,C	"The target in Phase A and B is evolving in the sense that you're discovering domain limits. While in Phase C you know the limits but now your efforts are directed to developing a product" (CR, 43)
	B,C	"The endpoint of Phase B is a working concept that still needs to be developed. In Phase C you know it works" (CR, 44)
Experimentation	A	"We chose a multiparticulate system instead of a monolytic system. So we chose in fact for a capsule containing a population of coated beads" (CR, 3)
	B	"Technically there are two possible ways you can coat a filler sphere; using a suspension, or using a solution. Both were tested in parallel and the best was chosen taking into consideration technical parameters and cost" (CR, 9, 12)
	B	"So this is concept characterisation [Phase B]; identifying the uncertainty, what are the parameters, and then screening them…" (CR, 29)
	C	"This is why I call concept application [Phase C] process-related. If you know this parameter is critical, what should I do with the production process to always fall within the parameter ranges (CR, 29)…During Phase B you may have determined particle size to be of critical importance, and you may have determined its critical ranges to be applied. Now, during Phase C you will have to determine, using DoE which knobs to turn on the machine in order to stay within these ranges" (CR, 32)
	B,C	"Another example; during Phase B you check which thickness of the coating layer gives which level of kinetics. PK people will tell you it goes too slow or too fast. So, during Phase B you develop a profile. Now, during Phase C you will determine how you have to design your production process to get these specific levels of thickness. Then we will have chosen and optimized for the most cost-effective process (CR, 33, 34)
	B,C	"You can apply DoE in Phase B to design your blend for example. But you will most certainly apply it in Phase C" (CR, 35)
Learning	A,B	"…There was no standard formula for CR or technical solution, so we started with a literature review and based upon some experience around IR coating complemented with some initial

		experiments we designed a formula" (CR, 11, 12)
	A,B	"…We knew theoretically how it should work, techniques were described in the literature how to do it at lab scale. However nobody had ever done clinical studies with it…so we tried a number of solutions in parallel without being able to fall back to standard technology platforms" (CR, 15, 16)
Coordination	B	"…Finally we were able to fill CR pallets in capsules but noticed the release profile was too slow at the beginning so we had to add an IR component, but how to do this? We brainstormed a number of options and checked for cost, compliance and logistic issues, finally we chose for… (CR, 22)
	C	"Formula screening and process optimisation is then done using a standard six sigma procedure. Eventually, you have to have a working commercially available product" (CR, 25)

Case Alfa Immediate Release basic data



Approach dimensions	Applies to Phase	Relevant empirical indicators
Target setting	A	"…The product definition is documented in the TCDS [Target Core Data Sheet] and can evolve up to a certain point in the stage-gate process after which no more modifications are allowed" (IR, 23)
Experimentation	A	"A standard platform formulation is like a rough formula. We took three platforms; lactose, calciumphosphate, and cellulose. …" (IR, 6)
	A	"…In parallel we used three standard platform formulations and tested each of them for stability…then we picked one out, the best. …Then you can start refining the formula" (IR, 7,8)

	A A A A	"…For these type of projects [standard solid tablets and oral dosage forms] we use an integrated set of development techniques under the banner of 'Design Excellence', which consists of techniques like pre-designed DoE, FMEA, VOC, QTC etc." (IR, 22) "…At each gate of the stage-gate process we have to be able to show which Design Excellence techniques have been used with what results (IR, 24) "…Development was done using the 'Direct compression formulation development flow chart' [shows this documented procedure (see ref (3))]. We use this procedure to run all our standard tablet formulations" (IR, 19) "…[refers to ref (3)]After parallel formula concept screening the process gets developed concurrently with formula optimisation because it is possible that you should adapt the formula because of the process" " (IR, 19)
Learning	A A A	"…The difficulty in this project was time, normally it takes 12 months to develop the formulation, now we had to do it in 8 months. The only solution was to work with standard platform formulations" (IR, 1) "…We use the standard platform formula. Of course, results of formulation work are never black or white…experience, skills are still important…if results are on the low side, which of the parallel-developed solutions should we pick out? (IR, 8) "…Experience built up over the years, by working with different products, and also experience in interpreting results are crucial in determining which type of filler we will use" (IR, 9)
Coordination	A A	"…Development was done using the 'Direct compression formulation development flow chart' [shows this documented procedure (see ref (3))]. We use this procedure to run all our standard tablet formulations" (IR, 19) "…GANTT charts [shows project chart] are used to track progress…" (IR, 22)

APPENDIX B: Complexity-handling modes empirical indicators leading to descriptor sets

Concept Selection		
Dimension	***Empirical indicators***	***Descriptor set***
Target setting approach	[1] "What Michel [project manager] tried to do is to set up, even before we started looking at these people, what we like to see for the very first criteria; that was functional DNA-Zyme activity in the cells" (DNAZyme, 22) "…I think the idea was really; let's identify these criteria [solution critical requirements] We have the opportunity to look at our range of technologies, but having said that, the very first step is to get rid of as many as possible. What would happen, if all of these passed the first test, I don't know. But we knew that the bar was so high, that it wasn't likely to happen." (DNAZyme, 31) "…The objective of these first two phases [refers to project process chart] is really to check for entire families of polymers whether it is possible to form micro-emulsions and to encapsulate an active ingredient, or not". (Micro, 45) "For the meltextrusion application we stated as an objective that the temperature with SCF had to be lower as the temperature we would have had without SCF. And this for a pressure that would not be so high it would endanger the functioning of the meltextruder. (SCF, 38) "We checked solubility of polymers in SCF and decided whether we could proceed or not. If it would not have been possible to solubilize polymers we could not have proceeded". (SCF, 24) "…At project start the basic Nanosuspension process technology was known. Our task was to see whether we could reproduce this process with our compounds at the level of clinical studies, and whether we could make it sterile…" (Nano, 1)	[1] Target defined up-front as minimal system critical requirements to pass
Experimentation approach	[2] "…It is so high risk you could never put all your eggs in the same basket. You never approach it by just assuming one unproven	[2] Run parallel experiments to characterise critical variables affecting

	technology is going to be the best. You have to spread in as many reasons as possible…so Michel went to a number of companies in the area with the idea establishing a number of parallel feasibility contracts…we also tried to diversify risk by looking at all known possibilities of delivery" (DNAZyme, 21, 3, 5) "The trouble was that, since we felt that it was a difficult project, it had to be very highly parallelized in terms of how to move forward, […]you also have to have good criteria (DNAZyme, 23) "…[refers to project process chart] the process is really to check for each family of polymers if micro-emulsions can be formed, then if active substance can be encapsulated. If it doesn't work, go back, if it does, take another family and start all over again…. After this initial phase we will be able to fully characterise the formulation for parenteral and oral applications" [refers to project process chart] (Micro, 12) "…At first instance Christina [project manager] checked the solubility of the active substance in the supercritical fluid, in parallel she checked whether the polymer could be extruded using supercritical fluids…at the first experiments she saw it was not really possible to bring liquid CO2 into the extruder, so it had to be a gas form" (SCF, 7) "We tried to have a working product as quickly as possible" (Nano, 12)…then we had to change several excipients… then we solved the suspension problems…finally we had a working formula…this was the end of Concept Selection (Nano, 13) [3] That were the three, we thought were the best, that we could move foreword with and that we are currently interacting with (DNAZyme, 12). "In Phase A you select, in Phase B you model and optimize the polymer structure…" (Micro, 72) "…In [Phase A] … we tested both concepts on one polymer. Both single and twin screw worked but the latter gave the best blend properties so we chose the last one to show proof of concept (SCF, 32)… "In [Phase A] you check; is it feasible or not?" (SCF, 37)	response for different candidate system solutions [3] System solution selection

	Comparing SCF to microemulsions' first screening phase: "For SCF there will also be a screening of classes of active substances and polymers for which SCF can be a solution. Which ones are feasible? Then a certain system will be selected and the most useful extruder parameters will be characterised". (SCF, 30) "…finally we had a working formula…this was the end of Concept Selection" (Nano, 13). "During Phase A you select a concept, in the second step you optimize your concept" (Nano, 19, 20) [4] "…The objective of these first two phases [refers to project process chart] is really to check for entire families of polymers whether it is possible to form micro-emulsions and to encapsulate an active ingredient, or not". (Micro, 45) "First you check the functioning of the system as a whole or not, only then you take the next step. This is why we split up the project in different steps (Micro, 35) "…Mid 2000 Proof of Concept was delivered in the Supercritical fluids project when polymers in interaction with $SCCO_2$ showed that they can decrease the meltextrusion temperature and that active substance can be made soluble in $SCCO_2$ under certain critical conditions" (SCF, 4) "…In [Phase A] … we tested both concepts on one polymer. Both single and twin screw worked but the latter gave the best blend properties so we chose the last one to show proof of concept (SCF, 32)… "In [Phase A] you check; is it feasible or not?" (SCF, 37) "We tried to have a working product as quickly as possible" (Nano, 12)…then we had to change several excipients… then we solved the suspension problems…finally we had a working formula…this was the end of Concept Selection (Nano, 13)	[4] Show Proof of Concept by spelling out assumptions about the set of relevant variables and their functional relationships
Learning approach	[5] "…There was so much published on so many failures, that we knew we could not do with the traditional approaches so we started already at the beginning with things that were pretty off the wall..." (DNAZyme, 17)	[5] External explicit: Mainly at project start learning from published science. Later ad-hoc for problem solving.

	" …We knew, that when we had any success, it would be from novel approaches and not from things that have been tried fifty times before". (DNAZyme, 18) "[Before starting]You check the literature from different perspectives and try to combine approaches without doing many experiments (Micro, 57) "…The first year an intensive literature study was needed complemented with a lot of contacts outside JJPRD to get a grasp of the technology" (SCF, 22) "…Knowledge transfer happened mainly at the beginning, (Nano, 16) "… later we used the relationship only for problem solving, and it was not a joint development effort…" (Nano, 16) [6] "…It is always stepwise, but the difference was the extent to which we thought we could have some sort of input into the different processes.….[x] had more development time and there was also more technology we could help them with, [y] was brand new, they had a polymer, very little was known, and because of that we thought we had a more open playfield….we interacted most with them (DNAZyme, 9, 11, 10) "…This is definitely a reciprocal collaboration contract we have with UCL [university] (Micro, 37) "…The characterisation of the supercritical fluid-active substance interaction is done in house with problem-solving help of the University of [x], for the extrusion characterisation part we do this in collaboration with the University of [y]…. It is not a real collaboration contract we have with them, it's more consultancy, they give advise to us…no milestone agreements exist with them" (SCF, 17, 18, 19)	[6] External tacit: On-going knowledge transfer by interaction between teams and external technology suppliers.
Coordination approach	[7] "…Everybody got the same message what was going to be the first milestone, what we wanted to see at that point. For those companies, that were either at the first milestone or that we had a good feeling about we broadened the topic to say, the next area will be this, and than after that it will	[7] Define milestone targets that are reached if results can be shown

	be…always raising the hurdle…" (DNAZyme, 26) "..I don't know if everybody does, but we took a stepped approach with each of the candidate solutions…first they had to show you can get it into the cell, then you have to have some in-vivo activity, then…" (DNAZyme, 7) " We went through basically screening with all of these" (DNAZyme, 13). "…First, we looked whether products of our pipeline could be brought into the extruder using supercritical fluids, then [in Phase B] we do a full physico-chemical characterisation of the whole extrusion system, later [Full Development] it will be scaled-up to full production level (SCF, 6) [8] "…Based upon some first experiments we estimated the time needed to do synthesis and characterisation of the compounds, then, we estimated it would take so many months to synthesize so many polymers…." (Micro, 26) … a professor from UCL has also made an estimation, then we compared and saw we came to about the same results (Micro, 27) [9] "…Since we felt that it was a difficult project it had to be very highly paralleled in terms of how to move forward…and you have to have good criteria to focus as quickly as possible on those companies and institutions that were able to demonstrate that…" (DNAZyme, 23, 24) "The more complicated the design space is, the more you tend to tailor your approach to answer a specific question, rather than learn as much as you can about that area that probably is not the case in a university, where off-course the goal of a university is something different of the goal of a pharmaceutical company in terms of generating space. It's just a realism that a pharmaceutical company is moving as quickly as possible to a space you can provide a product for" (DNAZyme, 25) "…Everybody got the same message what was going to be the first milestone, what we wanted to see at that point. For those companies, that were either at the first milestone or that we had a good feeling about we broadened the topic to say, the next area	[8] Estimate work package effort/ timeline based on first experiments or expert knowledge [9] Focus on experiments capable of selecting as quickly as possible solutions that meet all system critical requirements [10] Through close monitoring of progress:

	will be this, and than after that it will be…always raising the hurdle…" (DNAZyme, 26) [10] "…From the very beginning we always had in mind the stepped-approach, that if the first sets of experiments were positive, we would make commitment during the next step, if it would not have been positive we would have dropped it " (DNAZyme, 8). "…At first instance Christina [project manager] checked the solubility of the active substance in the supercritical fluid, in parallel she checked whether the polymer could be extruded using supercritical fluids…at the first experiments she saw it was not really possible to bring liquid CO2 into the extruder, so it had to be a gas form" (SCF, 7)	Eliminate as quickly as possible candidate system solutions not meeting one of the system critical requirements

Concept Characterisation		
Dimension	***Empirical indicators***	***Descriptor set***
Target setting approach	[1] "In the first phase the question to be answered is; is it stable or not? In this phase, the experimental plan will be more directed towards exploration of ranges of parameters –is this range a good range to function at production scale?- (Micro, 54) "You develop an application on model compounds, not for a final R-number [specific drug project] We will only transfer into Full Development when we will be sure it will be bio-available, enhances solubility and is non-toxic. This needs to be shown first on model compounds (Micro, 61) "Now the technology is already sufficiently characterized for one type of polymer to allow us to test its solubility in micro-emulsions for one specific compound falling within this application domain" (Micro, 64) "As soon as we find out if and how we can get the active substance into the SCF we will have to check how good the active can be solubilized, or which volume you can usefully bring into the extruder, what is the stability, bio-availability, of the mixture?"(SCF, 26) "Target setting in concept characterisation [Phase A and B] is done in a domain you don't know. You don't know what the value ranges of the variables are" (CR, 42) "The challenge was to develop a CR profile that would rise very quickly, stay up the whole day, and fall down in the evening so patients can sleep, all this without peaks to prevent side-effects"…(CR, 7) "The overall proof of concept for this was known. However, what was not known was how to get this significant quick rise at the beginning" (CR, 8) "The target in Phase A and B is evolving in the sense that you're discovering domain limits. While in Phase C you know the limits but now your efforts are directed to developing a product" (CR, 43)	[1]Target moving toward feasible application domain requirements to pass
Experimentation approach	[2] "…[refers to project process chart] then we would like to characterise parenteral and oral	[2]Run parallel/concurrent experiments to

	applications [in parallel] exploring how much active substance can be solved, whether a capsule can be filled, does the capsule break when it gets into contact with the acids in the gut? How about the toxicity and permeability for each of these applications…"(Micro, 11, 12) "In Phase A you select, in Phase B you model and optimize the polymer structure…" (Micro, 72) "…[refers to project process chart] then we would like to characterise parenteral and oral applications [in parallel] exploring how much active substance can be solved, whether a capsule can be filled, does the capsule break when it gets into contact with the acids in the gut? How about the toxicity and permeability for each of these applications…"(Micro, 11, 12) "…this project is actually a platform that can be used for different applications,…parenteral, oral, and transdermal" (Micro, 23, 25) "In the first phase the question to be answered is; is it stable or not? In this phase, the experimental plan will be more directed towards exploration of ranges of parameters –is this range a good range to function at production scale?- (Micro, 54) "…You start with determining under which conditions you can add SCCO2 to the polymer, then you look for the lowest possible temperature at which this can happen…you characterise parameters like speed of addition, pressure, single versus twin-screw geometry of the extrusion screw…" (SCF, 8) "You split up the concept proven in [Phase A] into subparts that you will characterise" (SCF, 35) "Then we started characterising the product…every parameter; particle size as a function of time, sedimentation as a function of time, etc. we checked for robustness" (Nano, 14) "Technically there are two possible ways you can coat a filler sphere; using a suspension, or using a solution. Both were tested in parallel and the best was chosen taking into consideration technical parameters and cost" (CR, 9, 12) "So this is concept characterisation [Phase B]; identifying the uncertainty, what are the parameters, and then screening them…" (CR, 29)	characterise all variables affecting response for different uncertainty areas within proof of concept delivered solution [3] Integrate uncertainty areas into limited / characterised application domain

	[3] "The end point of this Phase B is really to have a technology available, sufficiently characterised to have a compound project use it" (Micro, 79) "Now the technology is already sufficiently characterized for one type of polymer to allow us to test its solubility in micro-emulsions for one specific compound falling within this application domain" (Micro, 64) "…[After parallel characterisation of both areas] then we could bring the characterised supercritical fluid with active substance into the characterised extrusion process, bringing the two together…if it works we can start modelling the nozzle, if it doesn't work we'll have to check whether we could use carbon solvents" (SCF, 14) "At the end of Phase A you know you will have a working concept with a good kinetic profile that will also be manufacturable" (CR, 37) We selected a multiparticulate concept with an IR and CR piece (CR, 39) Then we said, we will be working using an aqueous solution and fine-tuned it to get a workable profile. That was Phase B or concept characterisation" (CR, 40)	
Learning approach	[4] [5] "…The characterisation of the supercritical fluid-active substance interaction is done in house with problem-solving help of the University of [x], for the extrusion characterisation part we do this in collaboration with the University of [y]…. It is not a real collaboration contract we have with them, it's more consultancy, they give advise to us…no milestone agreements exist with them" (SCF, 17, 18, 19) "…Knowledge transfer happened mainly at the beginning, (Nano, 16)"… later we used the relationship only for problem solving, and it was not a joint development effort…" (Nano, 16) "…We knew theoretically how it should work, techniques were described in the literature how to do it at lab scale. However nobody had ever done clinical studies with it…so we tried a number of solutions in parallel without being able to fall back to standard technology platforms" (CR, 15, 16) [6] "…There was no standard formula for CR or technical solution, so we started with a literature	[4]External explicit: Mainly at project start learning from published science. Later ad-hoc for problem solving. [5] External tacit: Knowledge-transfer mainly at project start between external technology supplier and team. Later for problem solving. [6] Internal tacit: Use of pockets of previous knowledge

	review and based upon some experience around IR coating complemented with some initial experiments we designed a formula" (CR, 11, 12) "… The basic technology we bought was the milling process and excipients/beads that gave this type of product this stability. We brought the sterile knowledge to the project" (Nano, 7)	
Coordination approach	[7] "…We split up the work in two coherent design packages [Formulation design and Sterile process design] each exploring and characterizing part of the solution…" (Nano, 2) " In [Full Development] you check for a specific compound application or dosage form like a capsule, whether stability in the log run, in-vivo bio-availability, the whole gamut of parameters are ok (SCF, 36). "…You start with determining under which conditions you can add SCCO2 to the polymer, then you look for the lowest possible temperature at which this can happen…you characterise parameters like speed of addition, pressure, single versus twin-screw geometry of the extrusion screw…" (SCF, 8) [8] "The end point of this Phase B is really to have a technology available, sufficiently characterised to have a compound project use it" (Micro, 79) "You develop an application on model compounds, not for a final R-number [specific drug project] We will only transfer into Full Development when we will be sure it will be bio-available, enhances solubility and is non-toxic. This needs to be shown first on model compounds (Micro, 61) "Technically there are two possible ways you can coat a filler sphere; using a suspension, or using a solution. Both were tested in parallel and the best was chosen taking into consideration technical parameters and cost" (CR, 9, 12) "In Phase B you are discovering ranges. The driving factor is kinetics. You know after Phase A which kinetic profile you need. How you will get it you discover in Phase B" (CR, 38) [9] "Design of Experiments will also be easier to be used in this phase than in the previous because	[7] Define uncertainty areas and assumptions to be tested per area [8] Define milestone targets for uncertainty areas that are reached if results can be shown. [9] Use DoE experimental guides to systematize the

	you now have a good idea which parameters are going to influence more or less strongly the process (Micro, 55) "In this mode the question is much more; is it stable or not? Or, [for this parameter] is this a good working range to have the application work at full production scale? DoE will be easier in this phase [Phase B] while you have an idea which parameters will influence the process more or less. So you know the factors to do a proper DoE". (SCF, 27, 28) "…As soon as we have a bit of a view of the process [Phase A] we will use Design of Experiments to characterise and optimise the process (SCF, 9). "You can apply DoE in Phase B to design your blend for example. But you will most certainly apply it in Phase C" (CR, 35) [10] [11] "I think the transition into Full Development is made as soon as there's a compound in need for this technology in its relevant application domain (Micro, 63) "…[After parallel characterisation of both areas] then we could bring the characterised supercritical fluid with active substance into the characterised extrusion process, bringing the two together…if it works we can start modelling the nozzle, if it doesn't work we'll have to check whether we could use carbon solvents" (SCF, 14) "…You start with determining under which conditions you can add SCCO2 to the polymer, then you look for the lowest possible temperature at which this can happen…you characterise parameters like speed of addition, pressure, single versus twin-screw geometry of the extrusion screw…" (SCF, 8) "…First, we looked whether products of our pipeline could be brought into the extruder using supercritical fluids, then [in Phase B] we do a full physico-chemical characterisation of the whole extrusion system, later [Full Development] it will be scaled-up to full production level (SCF, 6) "Another example; during Phase B you check which thickness of the coating layer gives which level of kinetics. PK people will tell you it goes too slow or too fast. So, during Phase B you	testing process [10] Guide progress through real-time coordination of concurrent results of different uncertainty areas [11] Bring focus through adaptive learning i.e.; assumptions testing, learning, continue/redirect efforts to characterise the feasible application domain

	develop a profile. Now, during Phase C you will determine how you have to design your production process to get these specific levels of thickness. Then we will have chosen and optimized for the most cost-effective process (CR, 33, 34) "The challenge was to develop a CR profile that would rise very quickly, stay up the whole day, and fall down in the evening so patients can sleep, all this without peaks to prevent side-effects"…(CR, 7) "The overall proof of concept for this was known. However, what was not known was how to get this significant quick rise at the beginning" (CR, 8) "After Phase A we had to have a working concept. It was actually a combination of selecting and characterizing a concept" (CR, 36) "At the end of Phase A you know you will have a working concept with a good kinetic profile that will also be manufacturable" (CR, 37) We selected a multiparticulate concept with an IR and CR piece (CR, 39) Then we said, we will be working using an aqueous solution and fine-tuned it to get a workable profile. That was Phase B or concept characterisation" (CR, 40) [12] "Depending on the results it is possible to go back from Phase B to Phase A because we get results that do not allow us to go through" (Micro, 71)	[12] Possible go back to previous mode if application domain cannot be delivered or if new application domain emerges

Concept Application		
Dimensions	***Empirical indicators***	***Concept Development***
Target setting approach	[1] "…The product definition is documented in the TCDS [Target Core Data Sheet] and can evolve up to a certain point in the stage-gate process after which no more modifications are allowed" (IR, 23)	[1] Target defined up-front as product/process requirements to pass. Only modifiable at formal stage-gates
Experimentation approach	[2] "This is why I call concept application [Phase C] process-related. If you know this parameter is critical, what should I do with the production process to always fall within the parameter ranges (CR, 29)…During Phase B you may have determined particle size to be of critical importance, and you may have determined its critical ranges to be applied. Now, during Phase C you will have to determine, using DoE which knobs to turn on the machine in order to stay within these ranges" (CR, 32) "Another example; during Phase B you check which thickness of the coating layer gives which level of kinetics. PK people will tell you it goes too slow or too fast. So, during Phase B you develop a profile. Now, during Phase C you will determine how you have to design your production process to get these specific levels of thickness. Then we will have chosen and optimized for the most cost-effective process (CR, 33, 34) "A standard platform formulation is like a rough formula. We took three platforms; lactose, calciumphosphate, and cellulose. …" (IR, 6) "…In parallel we used three standard platform formulations and tested each of them for stability…then we picked one out, the best. …Then you can start refining the formula" (IR, 7,8) "…For these type of projects [standard solid tablets and oral dosage forms] we use an integrated set of development techniques under the banner of 'Design Excellence', which consists of techniques like pre-designed DoE, FMEA, VOC, QTC etc." (IR, 22)	[2] Run parallel/concurrent experiments to optimise variables values within a solution [3] Followed by concurrent engineering driven by QFD derived product definition [4] FMEA based process design
Learning approach	[5] "…The difficulty in this project was time, normally it takes 12 months to develop the	[5] Internal explicit: Learning based on formal procedure-

	formulation, now we had to do it in 8 months. The only solution was to work with standard platform formulations" (IR, 1) [6] "…We use the standard platform formula. Of course, results of formulation work are never black or white…experience, skills are still important…if results are on the low side, which of the parallel-developed solutions should we pick out? (IR, 8) "…Experience built up over the years, by working with different products, and also experience in interpreting results are crucial in determining which type of filler we will use" (IR, 9)	based inquiry [6] Internal tacit: Learning by doing, based on previous internal experience
Coordination approach	[7] [8] [9] [10] [11] "You can apply DoE in Phase B to design your blend for example. But you will most certainly apply it in Phase C" (CR, 35) "Formula screening and process optimisation is then done using a standard six sigma procedure. Eventually, you have to have a working commercially available product" (CR, 25) "…At each gate of the stage-gate process we have to be able to show which Design Excellence techniques have been used with what results (IR, 24) "…Development was done using the 'Direct compression formulation development flow chart' [shows this documented procedure (see ref (3))]. We use this procedure to run all our standard tablet formulations" (IR, 19) "…[refers to ref (3)]After parallel formula concept screening the process gets developed concurrently with formula optimisation because it is possible that you should adapt the formula because of the process" " (IR, 19) "…GANTT charts [shows project chart] are used to track progress…" (IR, 22)	[7]Define product/process requirements to be met [8] Use pre-designed DoE experimental guides to systematize the experimentation process [9] Use GANTT type plans & schedules for cross-activity programming and tracking task completion [10] Monitor plan variation and act accordingly by executing contingency plans [11] Use of standard approaches and documented best practices to problem-solving

APPENDIX C: Confirmatory case study replication results

Case statistics

Case study	*# empirical indicators*	*Respondents*	*# interviews*
PharmaCo EMEA Discovery organisation			
	23	Jan Hoflack & Claire Macky	1
	8	Claire Macky	4
	7	Michael Engels	2
	4	Claire Macky & Michael Engels	2
Simulation model verification meetings		Claire Macky & Michael Engels	6
Totals	**42**	**3**	**15**

Case written sources

Sources used to document PharmaCo research context and case visual maps:

(10) PharmaCo 'EMEA Discovery Research process flowchart', Internal confidential publication.
(11) PharmaCo 'Assessment for transition from HTS to H2L', Internal confidential publication.
(12) PharmaCo 'Bioavailability = complex in-vivo disposition processes: Can they be reduced to discrete mechanisms to be modelled?', Internal confidential publication.
(13) Hoflack, J. (2002) 'The drug discovery challenge: Better targets, better compounds, better processes', internal publication for presentation.
(14) Mackie, C. (2002) 'Pre-clinical expertise in drug discovery: Changing the paradigm', internal publication for presentation.
(15) Maes, V. (2004) 'Discovery operational cost, timelines and resource breakdown', internal confidential spreadsheet.
(16) Toxicology reference assay data, internal PharmaCo confidential publication

In the following tables empirical indicators will be indicated with a number [x] at the end referring to the interview notes in the confirmatory case database.

Empirical descriptors for complexity-handling modes used in Discovery Research confirmatory case study

Complexity-handling modes	Target setting approach		
	Description	**Phase A empirical indicators**	**Phase B empirical indicators**
Concept Selection	• Target defined up-front as minimal system critical requirements to pass	• "Very often you take the high-risk approach for a new target, and then you have only about 30% chance of getting some efficacy meaning your compound has no effect on the disease. Many companies who have gone through this genome-based DD have experienced this and therefore we stress target validation so much [before starting problem-solving]"[16]. • It is really exploring the potential of the series. In LO you say I have explored the potential of a series, now I think I can find within that series the molecule I'm looking for .The focus now is on finding within a series a compound that fits my criteria. [22].	• No supporting indicators
Concept Characterisation	• Target moving toward feasible application domain requirements to pass	• No supporting indiscators	• No supporting indicators
Concept Development	• Target defined up-front as product/process requirements to pass. Only modifiable at formal stage-gates	• No supporting indicators	• No supporting indicators

Complexity-handling modes	Experimentation approach		
	Description	**Phase A empirical indicators**	**Phase B empirical indicators**
Concept Selection	• Define uncertainty to characterise critical variables and to anticipate problems to be solved (added due to [10] • Run parallel experiments to characterise critical variables affecting response for different candidate system solutions • System solution selection • Show Proof of Concept by spelling out assumptions about the set of relevant variables and their functional relationships	• So, for me at the beginning, these are just a set of choices, options, whatever they are, they need to be optimized in a certain direction, and selected.[2] • At the back of the HTS they do things like prior art, which is about patentability, ease of chemistry, preliminary in silico models to anticipate problems as hypotheses to be tested in-vitro.[10] • We take a sample of each family and they are very big so we don't want to optimize each and every compound within that family. Now in H2L we want to get a picture across the family of compounds to see where the strengths/weaknesses lie.[12] • So, in H2L we take a snapshot picture of the chemical families/clusters and see/prioritize which of them is worth further pursuing in LO. So, in essence the task of H2L is to check which of the clusters are worth prioritizing and taking them further to LO.[13] • Therefore, we run in parallel with pharmacology our screens [of different families], so that at the end you could for example say; I blocked my potency in-vitro, but I increased my solubility and permeability, which could mean a better overall in-vivo compound.[14]	• No supporting indicators
Concept Characterisation	• Run parallel/concurrent experiments to characterise all variables affecting response for different uncertainty areas within proof of concept delivered solution • Integrate uncertainty areas into limited / characterised application domain	• No supporting indicators (while never on THE PoC delivered solution).	• No supporting indicators
Concept Development	• Run parallel/concurrent experiments to optimise variables values within a solution • Followed by concurrent engineering driven by QFD derived product definition • FMEA based process design	• No supporting indicators	• No supporting indicators

Complexity-handling modes	Learning approach		
	Description	**Phase A empirical indicators**	**Phase B empirical indicators**
Concept Selection	• External explicit: Mainly at project start learning from published science. Later ad-hoc for problem solving. • External tacit: On-going knowledge transfer by interaction between teams and external technology suppliers. • Internal tacit: use of pockets of previous knowledge (added due to [17]	• "Dr Paul would say; don't work on something you have no experience with, do that 10 years from now. We would probably say; everything which is intra-cellular forget about it and focus on extra-cellular. We would probably say; the compound collection is attuned to this, we know the chemistry involved, this is a choice we as discovery management need to make" [17]	• No supporting indicators
Concept Characterisation	• External explicit: Mainly at project start learning from published science. Later ad-hoc for problem solving. • External tacit: Knowledge-transfer mainly at project start between external technology supplier and team. Later for problem solving. • Internal tacit: Use of pockets of previous knowledge	• No supporting indicators	• No supporting indicators
Concept Development	• Internal explicit: Learning based on formal procedure-based inquiry • Internal tacit: Learning by doing, based on previous internal experience	• No supporting indicators	• No supporting indicators

Complexity-handling modes	Coordination approach		
	Description	**Phase A empirical indicators**	**Phase B empirical indicators**
Concept Selection	• Define milestone targets that are reached if results can be shown • Estimate work package effort/ timeline based on first experiments or expert knowledge • Focus on experiments capable of selecting as quickly as possible solutions that meet all system critical requirements • Through close monitoring of progress: Eliminate as quickly as possible candidate system solutions not meeting one of the system critical requirements	• In summary, in H2L you score chemical classes from HTS for overall developability, and the one that gets the highest score overall, gets developed first. For each chemical class you say; if we have to take this class forward, what do we have to optimize in and out? Knowing then what lies ahead of you can prioritize, because you know that fe your task will be larger if you do this class.[15] • A flowchart is splitting it up into different components,… Criteria get harder as you move along, but also the methods become more resourceful, more difficult to perform. In H2L it is more a generic type of flowchart. The one in LO is built upon the perception of the series in H2L and you get specific biology questions, which are disease related.[20] • We look for an overall trade-of between properties. You look at your criteria and say; can I make this into a drug? If so, then you go on tox study, if it's clean you go DE.[26] • However, if you don't have any room or scope for manoeuvre it shouldn't go, because then you're optimising rubbish before you've even started.[29] • Normally, I wouldn't like to take a complete series through into LO if I cannot be convinced that I can get oral exposure, because most of the medication we're looking for is oral medicine. If it's not orally absorbed I cannot do any tox, kinetics, PK/PD, anything.[30]	• No supporting indicators
Concept Characterisation	• Define uncertainty areas and assumptions to be tested per area • Define milestone targets for uncertainty areas that are reached if results can be shown. • Use DoE experimental guides to systematize the testing process • Guide progress through real-time coordination of concurrent results of different uncertainty areas • Bring focus through adaptive learning i.e.; assumptions testing, learning, continue/redirect efforts to	• No supporting indicators	• First glance is then the first time you introduce the project to DE, these are the basics of the compound and this is what we aim to do in the next phase. But First Glance is not necessarily the compound you'll be handing over to the next phase. Then at the end of this phase (handover to DE) you have what I just indicated but now on the proper compound

	characterise the feasible application domain • Possible go back to previous mode if application domain cannot be delivered or if new application domain emerges		
Concept Development	• Define product/process requirements to be met • Use pre-designed DoE experimental guides to systematize the experimentation process • Use GANTT type plans & schedules for cross-activity programming and tracking task completion • Monitor plan variation and act accordingly by executing contingency plans • Use of standard approaches and documented best practices to problem-solving	• No supporting indicators	• No supporting indicators

APPENDIX E: Simulation system Visual Basic program code

This procedure starts the simulation (# of runs entered in the “result” sheet)

```
Sub CommandButton1_Click()
Start_at_Line = 4
Number_of_runs = Worksheets("Results").Cells(1, 4).Value
Worksheets("Results").Range("A" & Start_at_Line + 1 & ":AW65536").ClearContents
Start_time = Time
For i = 1 To Number_of_runs
Application.ScreenUpdating = False
Worksheets("Model").Range("D7:AW65000").ClearContents
On Error GoTo Start_here
Start_here:
Model.Start
Application.Calculation = xlCalculationAutomatic
Worksheets("Results").Cells(i + Start_at_Line, 1).Value = i
Worksheets("Results").Cells(i + Start_at_Line, 2).Value = Worksheets("Model").Cells(2, 49).Value
Worksheets("Results").Cells(i + Start_at_Line, 3).Value = Worksheets("Model").Cells(3, 49).Value
Worksheets("Results").Cells(i + Start_at_Line, 4).Value = Worksheets("Model").Cells(4, 47).Value
Worksheets("Results").Cells(i + Start_at_Line, 5).Value = Worksheets("Model").Cells(4, 48).Value

Worksheets("Results").Cells(i + Start_at_Line, 6).Value = Worksheets("Model").Cells(2, 47).Value
Worksheets("Results").Cells(i + Start_at_Line, 7).Value = Worksheets("Model").Cells(2, 48).Value
Worksheets("Results").Cells(i + Start_at_Line, 8).Value = Worksheets("Model").Cells(3, 47).Value
Worksheets("Results").Cells(i + Start_at_Line, 9).Value = Worksheets("Model").Cells(3, 48).Value
Worksheets("Results").Cells(1, 9).Value = Time - Start_time
Worksheets("Results").Cells(2, 4).Value = i
Worksheets("Results").Cells(i + Start_at_Line, 12).Value = Worksheets("Model").Cells(2, 51).Value
Application.ScreenUpdating = True
Next i
For j = Start_at_Line + 1 To Start_at_Line + Number_of_runs
Worksheets("Results").Cells(j, 13).Value = "=IF(D" & j & "<>0,F" & j & "/D" & j & ","""---""")"
Worksheets("Results").Cells(j, 14).Value = "=IF(E" & j & "<>0,I" & j & "/E" & j & ","""---""")"
Worksheets("Results").Cells(j, 15).Value = "=IF(B" & j & "<>0,F" & j & "/B" & j & ","""---""")"
Worksheets("Results").Cells(j, 16).Value = "=IF(C" & j & "<>0,H" & j & "/C" & j & ","""---""")"
Worksheets("Results").Cells(j, 17).Value = "=D" & j & "/Model!C6"
Next j
Worksheets("Results").Cells(1, 9).Value = Time - Start_time
End Sub
```

The code below constitutes the main part of the simulation tool. After the generation and classification of compounds in classes, the model applies the selected discovery scenario (Old paradigm, Front-loaded paradigm or Front Loading). Most parameters are inputted directly in the worksheets to allow an easier control by the user. Each step of the discovery process (HTS, H2L, LO) are modelled by a procedure.
The first module (“Start”) is the central procedure from where other procedures are called from.

```
Const Start_at_Line = 6
Sub Start()
   Real
   Application.Calculation = xlCalculationAutomatic
   HTS
   Application.Calculation = xlCalculationAutomatic
   H2L
   Application.Calculation = xlCalculationAutomatic
   LO
```

```
   Application.Calculation = xlCalculationAutomatic
End Sub
Private Sub CommandButton1_Click()
Application.Calculation = xlCalculationManual
Number_of_Classes = Worksheets("Model").Cells(4, 2).Value
Number_of_Ref = Worksheets("Model").Cells(5, 2).Value
Number_of_Comp = Worksheets("Model").Cells(6, 2).Value
Worksheets("Model").Range("A7:BB65000").ClearContents
Worksheets("Model").Range("1:65000").Ungroup
Worksheets("Model").Range("1:65000").Ungroup
Line = Start_at_Line
For i = 1 To Number_of_Classes
      Line = Line + 1
      Worksheets("Model").Cells(Line + 1, 1) = "Class H(" & i & ")"
   For j = 1 To Number_of_Ref
         Line = Line + 2
         Worksheets("Model").Cells(Line + 1, 2) = "Ref Compound H(" & i & "," & j & ")"
         Line = Line + 1
      For h = 1 To Number_of_Comp
            Line = Line + 1
            Worksheets("Model").Cells(Line, 3) = "Compound H(" & i & "," & j & "," & h & ")"
      Next h
      h = 1
   Next j
   j = 1
Next i

Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
      Line = Line + 1
      Worksheets("Model").Range(Line + 1 & ":" & Line + (Number_of_Comp + 3) * Number_of_Ref).Group
   For j = 1 To Number_of_Ref
         Line = Line + 3
            Worksheets("Model").Range(Line & ":" & Line + Number_of_Comp - 1).Group
      For h = 1 To Number_of_Comp
            Line = Line + 1
      Next h
      h = 1
   Next j
   j = 1
Next i

Application.Calculation = xlCalculationAutomatic
End Sub

Private Sub Real()
Application.Calculation = xlCalculationManual
a = Worksheets("Distributions").Cells(48, 7).Value
c = Worksheets("Distributions").Cells(6, 7).Value
b = Worksheets("Distributions").Cells(27, 7).Value
Sigma_pot_ref = Worksheets("Model").Cells(2, 4).Value
Sigma_pot_comp = Worksheets("Model").Cells(3, 4).Value
Sigma_bio_ref = Worksheets("Model").Cells(2, 5).Value
Sigma_bio_comp = Worksheets("Model").Cells(3, 5).Value
Number_of_Classes = Worksheets("Model").Cells(4, 2).Value
Number_of_Ref = Worksheets("Model").Cells(5, 2).Value
Number_of_Comp = Worksheets("Model").Cells(6, 2).Value
Line = Start_at_Line

' ------ Genration of potency values --------

For i = 1 To Number_of_Classes
      Line = Line + 1
      Randomize
      Potency = 9.5 - c * (Log(Rnd() * (Exp(9 / c) - 1) + 1))
      Select Case Potency
         Case Is > 9.5
            Worksheets("Model").Cells(Line, 4).Value = 9.5
         Case Is < 0.5
            Worksheets("Model").Cells(Line, 4).Value = 0.5
         Case Else
            Worksheets("Model").Cells(Line, 4).Value = Potency
```

```
      End Select
   For j = 1 To Number_of_Ref
         Randomize
         Line = Line + 2
         Potency_Ref = Application.WorksheetFunction.NormInv(Rnd(), Potency, Sigma_pot_ref)
         Select Case Potency_Ref
            Case Is > 9.5
               Worksheets("Model").Cells(Line, 4).Value = 9.5
            Case Is < 0.5
               Worksheets("Model").Cells(Line, 4).Value = 0.5
            Case Else
               Worksheets("Model").Cells(Line, 4).Value = Potency_Ref
         End Select
         Line = Line + 1
      For h = 1 To Number_of_Comp
            Randomize
            Line = Line + 1
            Potency_comp = Application.WorksheetFunction.NormInv(Rnd(), Potency_Ref, Sigma_pot_comp)
            Select Case Potency_comp
               Case Is > 9.5
                  Worksheets("Model").Cells(Line, 4).Value = 9.5
               Case Is < 0.5
                  Worksheets("Model").Cells(Line, 4).Value = 0.5
               Case Else
                  Worksheets("Model").Cells(Line, 4).Value = Potency_comp
         End Select
      Next h
      h = 1
   Next j
   j = 1
Next i

' ------ Genration of bio-availability values --------

Line = Start_at_Line
For i = 1 To Number_of_Classes
      Line = Line + 1
      Randomize
      Bioavailability = 9.5 - b * (Log(Rnd() * (Exp(9 / b) - 1) + 1))
      Select Case Bioavailability
         Case Is > 9.5
            Worksheets("Model").Cells(Line, 5).Value = 9.5
         Case Is < 0.5
            Worksheets("Model").Cells(Line, 5).Value = 0.5
         Case Else
            Worksheets("Model").Cells(Line, 5).Value = Bioavailability
      End Select
   For j = 1 To Number_of_Ref
         Randomize
         Line = Line + 2
         Bioavailability_Ref = Application.WorksheetFunction.NormInv(Rnd(), Bioavailability, Sigma_bio_ref)
         Select Case Bioavailability_Ref
            Case Is > 9.5
               Worksheets("Model").Cells(Line, 5).Value = 9.5
            Case Is < 0.5
               Worksheets("Model").Cells(Line, 5).Value = 0.5
            Case Else
               Worksheets("Model").Cells(Line, 5).Value = Bioavailability_Ref
         End Select
         Line = Line + 1
      For h = 1 To Number_of_Comp
            Randomize
            Line = Line + 1
            Bioavailability_comp = Application.WorksheetFunction.NormInv(Rnd(), Bioavailability_Ref, Sigma_bio_comp)
            Select Case Bioavailability_comp
               Case Is > 9.5
                  Worksheets("Model").Cells(Line, 5).Value = 9.5
               Case Is < 0.5
                  Worksheets("Model").Cells(Line, 5).Value = 0.5
               Case Else
                  Worksheets("Model").Cells(Line, 5).Value = Bioavailability_comp
         End Select
      Next h
```

```
        h = 1
    Next j
    j = 1
Next i

' ------ Genration of Toxicity values --------
Line = Start_at_Line
For i = 1 To Number_of_Classes
        Line = Line + 1
    For j = 1 To Number_of_Ref
            Line = Line + 3
        For h = 1 To Number_of_Comp
                Randomize
                Line = Line + 1
                '----- Old tox = a * Log(Rnd() * (Exp(9 / a) - 1) + 1) + 0.5
                Toxicity_comp = 10 - (0.25 * Rnd() + 0.75) * Worksheets("Model").Cells(Line, 4).Value
                Select Case Toxicity_comp
                    Case Is > 9.5
                        Worksheets("Model").Cells(Line, 6).Value = 9.5
                    Case Is < 0.5
                        Worksheets("Model").Cells(Line, 6).Value = 0.5
                    Case Else
                        Worksheets("Model").Cells(Line, 6).Value = Toxicity_comp
            End Select
        Next h
        h = 1
    Next j
    j = 1
Next i


Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
    Line = Line + 1

'------ Averages ------

    Average_Col_D = "=Average(D" & Line + 2
    Average_Col_E = "=Average(E" & Line + 2
    Average_Col_F = "=Average(F" & Line + 2
    For x = 2 To Number_of_Ref
        Average_Col_D = Average_Col_D & ",D" & Line + (x - 1) * (Number_of_Comp + 3) + 2
        Average_Col_E = Average_Col_E & ",E" & Line + (x - 1) * (Number_of_Comp + 3) + 2
        Average_Col_F = Average_Col_F & ",F" & Line + (x - 1) * (Number_of_Comp + 3) + 2
    Next x
        Average_Col_D = Average_Col_D & ")"
        Average_Col_E = Average_Col_E & ")"
        Average_Col_F = Average_Col_F & ")"

    Worksheets("Model").Cells(Line, 4).Value = Average_Col_D
    Worksheets("Model").Cells(Line, 5).Value = Average_Col_E
    Worksheets("Model").Cells(Line, 6).Value = Average_Col_F

'--------------------------------

    For j = 1 To Number_of_Ref
            Line = Line + 2
            Worksheets("Model").Cells(Line, 4).Value = "=Average(D" & Line + 1 & ":D" & Line + Number_of_Comp & ")"
            Worksheets("Model").Cells(Line, 5).Value = "=Average(E" & Line + 1 & ":E" & Line + Number_of_Comp & ")"
            Worksheets("Model").Cells(Line, 6).Value = "=Average(F" & Line + 1 & ":F" & Line + Number_of_Comp & ")"
            Line = Line + 1
        For h = 1 To Number_of_Comp
                Line = Line + 1
        Next h
        h = 1
    Next j
    j = 1
Next i
'--------- Selection on Real Values  ------------------
Line = Start_at_Line
For i = 1 To Number_of_Classes
        Line = Line + 1
```

```
   For j = 1 To Number_of_Ref
         Line = Line + 3
      For h = 1 To Number_of_Comp
         Line = Line + 1
            Worksheets("Model").Cells(Line, 44).Value =
"=IF(AND(RC4>=R2C37,AND(RC5>=R3C37,RC6<=R4C37)),""Pre-selected"","""")"
            Worksheets("Model").Cells(Line, 45).Value = "=If(RC42 = ""X"",""Pre-selected"","""") & "" - "" & RC44"
            If Worksheets("Model").Cells(Line, 4).Value < Worksheets("Model").Cells(Line, 5).Value Then
               Worksheets("Model").Cells(Line, 46).Value = Worksheets("Model").Cells(Line, 4).Value
            Else
               Worksheets("Model").Cells(Line, 46).Value = Worksheets("Model").Cells(Line, 5).Value
            End If
      Next h
      h = 1
   Next j
   j = 1
Next i
'-----------------------------------------------------------
Application.Calculation = xlCalculationAutomatic
End Sub

Private Sub HTS()
Sigma_pot_HTS = Worksheets("Model").Cells(2, 9).Value
Sigma_bio_HTS = Worksheets("Model").Cells(2, 10).Value
Sigma_tox_HTS = Worksheets("Model").Cells(2, 11).Value
Number_of_Classes = Worksheets("Model").Cells(4, 2).Value
Number_of_Ref = Worksheets("Model").Cells(5, 2).Value
Number_of_Comp = Worksheets("Model").Cells(6, 2).Value
Application.Calculation = xlCalculationManual
Worksheets("Model").Range("I7:R65000").ClearContents
Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
      Line = Line + 1

'------ Averages ------

   Average_Col_I = "=Average(I" & Line + 2
   Average_Col_J = "=Average(J" & Line + 2
   Average_Col_K = "=Average(K" & Line + 2
   For x = 2 To Number_of_Ref
      Average_Col_I = Average_Col_I & ",I" & Line + (x - 1) * (Number_of_Comp + 3) + 2
      Average_Col_J = Average_Col_J & ",J" & Line + (x - 1) * (Number_of_Comp + 3) + 2
      Average_Col_K = Average_Col_K & ",K" & Line + (x - 1) * (Number_of_Comp + 3) + 2
   Next x
      Average_Col_I = Average_Col_I & ")"
      Average_Col_J = Average_Col_J & ")"
      Average_Col_K = Average_Col_K & ")"

   Worksheets("Model").Cells(Line, 9).Value = Average_Col_I
   Worksheets("Model").Cells(Line, 10).Value = Average_Col_J
   Worksheets("Model").Cells(Line, 11).Value = Average_Col_K

'---------------------------------


         Worksheets("Model").Cells(Line, 17).Value = "=if(Iserror(RANK(O" & Line & ",$O$8:$O$65536)) =
False,RANK(O" & Line & ",$O$8:$O$65536),"""")"

         Worksheets("Model").Cells(Line, 18).Value = "=If(Q" & Line & "< $R$3,""X"","""")"

   For j = 1 To Number_of_Ref
               Line = Line + 2
               Worksheets("Model").Cells(Line, 9).Value = "=Average(I" & Line + 1 & ":I" & Line + Number_of_Comp & ")"
               Worksheets("Model").Cells(Line, 10).Value = "=Average(J" & Line + 1 & ":J" & Line + Number_of_Comp &
")"
               Worksheets("Model").Cells(Line, 11).Value = "=Average(K" & Line + 1 & ":K" & Line + Number_of_Comp &
")"
               Randomize
               Potency_HTS = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line, 4).Value,
Sigma_pot_HTS)
               Select Case Potency_HTS
                     Case Is > 9.5
```

```
                    Worksheets("Model").Cells(Line, 9).Value = 9.5
                Case Is < 0.5
                    Worksheets("Model").Cells(Line, 9).Value = 0.5
                Case Else
                    Worksheets("Model").Cells(Line, 9).Value = Potency_HTS
            End Select
            Randomize
            Bioavailability_HTS = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line,
5).Value, Sigma_bio_HTS)
            Select Case Bioavailability_HTS
                Case Is > 9.5
                    Worksheets("Model").Cells(Line, 10).Value = 9.5
                Case Is < 0.5
                    Worksheets("Model").Cells(Line, 10).Value = 0.5
                Case Else
                    Worksheets("Model").Cells(Line, 10).Value = Bioavailability_HTS
            End Select
            Randomize
            Toxicity_HTS = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line, 6).Value,
Sigma_tox_HTS)
            Select Case Toxicity_HTS
                Case Is > 9.5
                    Worksheets("Model").Cells(Line, 11).Value = 9.5
                Case Is < 0.5
                    Worksheets("Model").Cells(Line, 11).Value = 0.5
                Case Else
                    Worksheets("Model").Cells(Line, 11).Value = Toxicity_HTS
            End Select

        Line = Line + 1
    For h = 1 To Number_of_Comp
        Line = Line + 1
    Next h
    h = 1
  Next j
  j = 1
Next i

Application.Calculation = xlCalculationAutomatic
Application.Calculation = xlCalculationManual

Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
    Line = Line + 1
    Worksheets("Model").Cells(Line, 13).Value = "=IF(AND(I" & Line & ">=$M$2,AND(J" & Line & ">=$M$3,K" & Line
& "<=$M$4)),""Pre-selected"","""")"

        Select Case Worksheets("Model").Cells(2, 15).Value
            Case Is = "Potency"
                If Worksheets("Model").Cells(Line, 13).Value = "Pre-selected" Then
                    Worksheets("Model").Cells(Line, 15).Value = Worksheets("Model").Cells(Line, 9).Value
                End If
            Case Is = "Bio-availability"
                If Worksheets("Model").Cells(Line, 13).Value = "Pre-selected" Then
                    Worksheets("Model").Cells(Line, 15).Value = Worksheets("Model").Cells(Line, 10).Value
                End If
            Case Is = "Min"
                If Worksheets("Model").Cells(Line, 13).Value = "Pre-selected" Then
                    Worksheets("Model").Cells(Line, 15).Value = "=Min(I" & Line & ":J" & Line & ")"
                End If
            Case Is = "Weighted Average"
                If Worksheets("Model").Cells(Line, 13).Value = "Pre-selected" Then
                    Worksheets("Model").Cells(Line, 15).Value = "=($P$3*I" & Line & "+$P$4*J" & Line & ")/($P$3+$P$4)"
                End If
        End Select

  For j = 1 To Number_of_Ref
        Line = Line + 3
    For h = 1 To Number_of_Comp
        Line = Line + 1
    Next h
    h = 1
```

```
    Next j
    j = 1
Next i


Application.Calculation = xlCalculationAutomatic
End Sub


Private Sub CommandButton2_Click()
Start
End Sub

Private Sub CommandButton4_Click()
Application.Calculation = xlCalculationManual
Number_of_Classes = Worksheets("Model").Cells(4, 2).Value
Number_of_Ref = Worksheets("Model").Cells(5, 2).Value
Number_of_Comp = Worksheets("Model").Cells(6, 2).Value
Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
        Line = Line + 1
            Select Case Worksheets("Model").Cells(2, 15).Value
                Case Is = "Potency"
                    If Worksheets("Model").Cells(Line, 13).Value = "Pre-selected" Then
                        Worksheets("Model").Cells(Line, 15).Value = Worksheets("Model").Cells(Line, 9).Value
                    End If
                Case Is = "Bio-availability"
                    If Worksheets("Model").Cells(Line, 13).Value = "Pre-selected" Then
                        Worksheets("Model").Cells(Line, 15).Value = Worksheets("Model").Cells(Line, 10).Value
                    End If
                Case Is = "Min"
                    If Worksheets("Model").Cells(Line, 13).Value = "Pre-selected" Then
                        Worksheets("Model").Cells(Line, 15).Value = "=Min(I" & Line & ":J" & Line & ")"
                    End If
                Case Is = "Weighted Average"
                    If Worksheets("Model").Cells(Line, 13).Value = "Pre-selected" Then
                        Worksheets("Model").Cells(Line, 15).Value = "=($P$3*I" & Line & "+$P$4*J" & Line & ")/($P$3+$P$4)"
                    End If
            End Select

    For j = 1 To Number_of_Ref
            Line = Line + 3
        For h = 1 To Number_of_Comp
                Line = Line + 1
        Next h
        h = 1
    Next j
    j = 1
Next i


Application.Calculation = xlCalculationAutomatic
End Sub

Private Sub H2L()
Sigma_pot_H2L = Worksheets("Model").Cells(2, 21).Value
Sigma_bio_H2L = Worksheets("Model").Cells(2, 22).Value
Sigma_tox_H2L = Worksheets("Model").Cells(2, 23).Value
Number_of_Classes = Worksheets("Model").Cells(4, 2).Value
Number_of_Ref = Worksheets("Model").Cells(5, 2).Value
Number_of_Comp = Worksheets("Model").Cells(6, 2).Value
Application.Calculation = xlCalculationManual
Worksheets("Model").Range("U7:AD65000").ClearContents
Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
        Line = Line + 1
            Checked_HTS = Worksheets("Model").Cells(Line, 18).Value
            If Checked_HTS = "X" Then

'------ Averages ------
```

```
   Average_Col_U = "=Average(U" & Line + 2
   Average_Col_V = "=Average(V" & Line + 2
   Average_Col_W = "=Average(W" & Line + 2
   For x = 2 To Number_of_Ref
      Average_Col_U = Average_Col_U & ",U" & Line + (x - 1) * (Number_of_Comp + 3) + 2
      Average_Col_V = Average_Col_V & ",V" & Line + (x - 1) * (Number_of_Comp + 3) + 2
      Average_Col_W = Average_Col_W & ",W" & Line + (x - 1) * (Number_of_Comp + 3) + 2
   Next x
      Average_Col_U = Average_Col_U & ")"
      Average_Col_V = Average_Col_V & ")"
      Average_Col_W = Average_Col_W & ")"

   Worksheets("Model").Cells(Line, 21).Value = Average_Col_U
   Worksheets("Model").Cells(Line, 22).Value = Average_Col_V
   Worksheets("Model").Cells(Line, 23).Value = Average_Col_W

'---------------------------------

         Worksheets("Model").Cells(Line, 29).Value = "=if(Iserror(RANK(AA" & Line & ",$AA$8:$AA$65536)) =
False,RANK(AA" & Line & ",$AA$8:$AA$65536),"""")"

         Worksheets("Model").Cells(Line, 30).Value = "=If(AC" & Line & "< $AD$3,""X"","""")"
         End If
   For j = 1 To Number_of_Ref
               Line = Line + 2
               If Checked_HTS = "X" Then
               Worksheets("Model").Cells(Line, 21).Value = "=Average(U" & Line + 1 & ":U" & Line + Number_of_Comp &
")"
               Worksheets("Model").Cells(Line, 22).Value = "=Average(V" & Line + 1 & ":V" & Line + Number_of_Comp &
")"
               Worksheets("Model").Cells(Line, 23).Value = "=Average(W" & Line + 1 & ":W" & Line + Number_of_Comp
& ")"
               Randomize
               Potency_H2L = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line, 4).Value,
Sigma_pot_H2L)
               Select Case Potency_H2L
                     Case Is > 9.5
                        Worksheets("Model").Cells(Line, 21).Value = 9.5
                     Case Is < 0.5
                        Worksheets("Model").Cells(Line, 21).Value = 0.5
                     Case Else
                        Worksheets("Model").Cells(Line, 21).Value = Potency_H2L
               End Select
               Randomize
               Bioavailability_H2L = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line,
5).Value, Sigma_bio_H2L)
               Select Case Bioavailability_H2L
                     Case Is > 9.5
                        Worksheets("Model").Cells(Line, 22).Value = 9.5
                     Case Is < 0.5
                        Worksheets("Model").Cells(Line, 22).Value = 0.5
                     Case Else
                        Worksheets("Model").Cells(Line, 22).Value = Bioavailability_H2L
               End Select
               Randomize
               Toxicity_H2L = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line, 6).Value,
Sigma_tox_H2L)
               Select Case Toxicity_H2L
                     Case Is > 9.5
                        Worksheets("Model").Cells(Line, 23).Value = 9.5
                     Case Is < 0.5
                        Worksheets("Model").Cells(Line, 23).Value = 0.5
                     Case Else
                        Worksheets("Model").Cells(Line, 23).Value = Toxicity_H2L
               End Select
               End If
         Line = Line + 1
      For h = 1 To Number_of_Comp
            Line = Line + 1
      Next h
      h = 1
   Next j
   j = 1
```

```
Next i

Application.Calculation = xlCalculationAutomatic
Application.Calculation = xlCalculationManual

Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
      Line = Line + 1
         Checked_HTS = Worksheets("Model").Cells(Line, 18).Value
         If Checked_HTS = "X" Then
         Worksheets("Model").Cells(Line, 25).Value = "=IF(AND(U" & Line & ">=$Y$2,AND(V" & Line & ">=$Y$3,W" &
Line & "<=$Y$4)),""Pre-selected"","""")"

         Select Case Worksheets("Model").Cells(2, 27).Value
            Case Is = "Potency"
               If Worksheets("Model").Cells(Line, 25).Value = "Pre-selected" Then
                  Worksheets("Model").Cells(Line, 27).Value = Worksheets("Model").Cells(Line, 21).Value
               End If
            Case Is = "Bio-availability"
               If Worksheets("Model").Cells(Line, 25).Value = "Pre-selected" Then
                  Worksheets("Model").Cells(Line, 27).Value = Worksheets("Model").Cells(Line, 22).Value
               End If
            Case Is = "Min"
               If Worksheets("Model").Cells(Line, 25).Value = "Pre-selected" Then
                  Worksheets("Model").Cells(Line, 27).Value = "=Min(U" & Line & ":V" & Line & ")"
               End If
            Case Is = "Weighted Average"
               If Worksheets("Model").Cells(Line, 25).Value = "Pre-selected" Then
                  Worksheets("Model").Cells(Line, 27).Value = "=($AB$3*U" & Line & "+$AB$4*V" & Line &
")/($AB$3+$AB$4)"
               End If
         End Select
         End If
   For j = 1 To Number_of_Ref
         Line = Line + 3
      For h = 1 To Number_of_Comp
            Line = Line + 1
      Next h
      h = 1
   Next j
   j = 1
Next i


Application.Calculation = xlCalculationAutomatic

End Sub

Private Sub LO()
Sigma_pot_LO = Worksheets("Model").Cells(2, 21).Value
Sigma_bio_LO = Worksheets("Model").Cells(2, 22).Value
Sigma_tox_LO = Worksheets("Model").Cells(2, 23).Value
Number_of_Classes = Worksheets("Model").Cells(4, 2).Value
Number_of_Ref = Worksheets("Model").Cells(5, 2).Value
Number_of_Comp = Worksheets("Model").Cells(6, 2).Value
Application.Calculation = xlCalculationManual
Worksheets("Model").Range("AG7:AP65000").ClearContents
Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
      Line = Line + 1
         Checked_H2L = Worksheets("Model").Cells(Line, 30).Value
         For j = 1 To Number_of_Ref
               Line = Line + 2
                     Line = Line + 1
         For h = 1 To Number_of_Comp
            If Checked_H2L = "X" Then
               Randomize
               Potency_LO = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line, 4).Value,
Sigma_pot_LO)
               Select Case Potency_LO
                     Case Is > 9.5
```

```
                    Worksheets("Model").Cells(Line, 33).Value = 9.5
                Case Is < 0.5
                    Worksheets("Model").Cells(Line, 33).Value = 0.5
                Case Else
                    Worksheets("Model").Cells(Line, 33).Value = Potency_LO
            End Select
            Randomize
            Bioavailability_LO = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line,
5).Value, Sigma_bio_LO)
            Select Case Bioavailability_LO
                Case Is > 9.5
                    Worksheets("Model").Cells(Line, 34).Value = 9.5
                Case Is < 0.5
                    Worksheets("Model").Cells(Line, 34).Value = 0.5
                Case Else
                    Worksheets("Model").Cells(Line, 34).Value = Bioavailability_LO
            End Select
            Randomize
            Toxicity_LO = Application.WorksheetFunction.NormInv(Rnd(), Worksheets("Model").Cells(Line, 6).Value,
Sigma_tox_LO)
            Select Case Toxicity_LO
                Case Is > 9.5
                    Worksheets("Model").Cells(Line, 35).Value = 9.5
                Case Is < 0.5
                    Worksheets("Model").Cells(Line, 35).Value = 0.5
                Case Else
                    Worksheets("Model").Cells(Line, 35).Value = Toxicity_LO
            End Select

        Worksheets("Model").Cells(Line, 41).Value = "=if(Iserror(RANK(AM" & Line & ",$AM$8:$AM$65536)) =
False,RANK(AM" & Line & ",$AM$8:$AM$65536),"""")"
        Worksheets("Model").Cells(Line, 42).Value = "=If(AO" & Line & "< $AP$3,""X"","""")"
        End If
        Line = Line + 1
    Next h
    h = 1
  Next j
  j = 1
Next i

Application.Calculation = xlCalculationAutomatic
Application.Calculation = xlCalculationManual

Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
    Line = Line + 1
      Checked_H2L = Worksheets("Model").Cells(Line, 30).Value
    For j = 1 To Number_of_Ref
      Line = Line + 3
    For h = 1 To Number_of_Comp
        If Checked_H2L = "X" Then
      Worksheets("Model").Cells(Line, 37).Value = "=IF(AND(AG" & Line & ">=$AK$2,AND(AH" & Line &
">=$AK$3,AI" & Line & "<=$AK$4)),""Pre-selected"","""")"

      Select Case Worksheets("Model").Cells(2, 39).Value
        Case Is = "Potency"
          If Worksheets("Model").Cells(Line, 37).Value = "Pre-selected" Then
            Worksheets("Model").Cells(Line, 39).Value = Worksheets("Model").Cells(Line, 33).Value
          End If
        Case Is = "Bio-availability"
          If Worksheets("Model").Cells(Line, 37).Value = "Pre-selected" Then
            Worksheets("Model").Cells(Line, 39).Value = Worksheets("Model").Cells(Line, 34).Value
          End If
        Case Is = "Min"
          If Worksheets("Model").Cells(Line, 37).Value = "Pre-selected" Then
            Worksheets("Model").Cells(Line, 39).Value = "=Min(AG" & Line & ":AH" & Line & ")"
          End If
        Case Is = "Weighted Average"
          If Worksheets("Model").Cells(Line, 37).Value = "Pre-selected" Then
            Worksheets("Model").Cells(Line, 39).Value = "=($AN$3*AG" & Line & "+$AN$4*AH" & Line &
")/($AN$3+$AN$4)"
          End If
```

```
            End Select
            End If
                Line = Line + 1
        Next h
        h = 1
    Next j
    j = 1
Next i

Application.Calculation = xlCalculationAutomatic
Application.Calculation = xlCalculationManual
Line = Start_at_Line
Line = Line + 1
For i = 1 To Number_of_Classes
        Line = Line + 1
        For j = 1 To Number_of_Ref
            Line = Line + 3
        For h = 1 To Number_of_Comp
            If Worksheets("Model").Cells(Line, 42).Value = "X" Then
                Worksheets("Model").Cells(Line, 47).Value = "=if(Iserror(RANK(AT" & Line & ",$AT$8:$AT$65536)) =
False,RANK(AT" & Line & ",$AT$8:$AT$65536),"""")"
            End If
            Line = Line + 1
        Next h
        h = 1
    Next j
    j = 1
Next i

Application.Calculation = xlCalculationAutomatic
End Sub
```