MAINTAINING SENSING INFLUENCE

AN AUTOPOIETIC BEHAVIOUR MODEL FOR
REMEDYING PERSISTENT HUMAN ERROR

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr ir J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op vrijdag 8 september 2006 om 10.00 uur

door

Machtelt VAN DER VLUGT
werktuigbouwkundig ingenieur
geboren te Heemskerk

Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr ir P.A. Wieringa
Prof. dr J. Dankelman

Samenstelling promotiecommissie:
Rector Magnificus, voorzitter
Prof. dr ir P.A. Wieringa, Technische Universiteit Delft, promotor
Prof. dr J. Dankelman, Technische Universiteit Delft, promotor
Prof. dr G. Grote, Eidgenössische Technische Hochschule Zürich
Prof. dr ir J.A. Mulder, Technische Universiteit Delft
Prof. dr A.R. Hale, Technische Universiteit Delft
Prof. dr M. Neerincx, Technische Universiteit Delft
Dr J.O. Kroesen, Technische Universiteit Delft

Title:	Maintaining Sensing Influence, an autopoietic behaviour model for remedying persistent human error. PhD thesis, Delft University of Technology, Delft, the Netherlands, September 2006.
Author:	Machtelt van der Vlugt
Cover design:	Machtelt van der Vlugt
Copyright:	Machtelt van der Vlugt, Delft, the Netherlands, 2006
Print:	Ponsen en Looijen BV
ISBN-10:	90-6464-018-1
ISBN-13:	978-90-6464-018-6

CONTENTS

1 INTRODUCTION 1

1.1 BACKGROUND: DIFFICULTIES WITH HUMAN ERRORS THAT PERSIST 2

1.2 OBJECTIVE. 6

1.3 MOTIVATION OF CHOICE FOR AVIATION 6

1.4 METHODOLOGY APPLIED. 7

1.5 OUTLINE OF THIS THESIS 8

REFERENCES 9

2 REMEDYING PERSISTENT HUMAN ERROR: THE NEED FOR A NEW BEHAVIOUR MODEL 11

2.1 INTRODUCTION. 12

2.2 MODELLING THE BEHAVIOUR OF A FEEDBACK SYSTEM WITH REGARD TO REMEDYING ERROR 13

2.3 DECISION-MAKING FEEDBACK SYSTEMS. 16

Behaviour of decision-making feedback systems 16

Restoring predictability of decision-making feedback systems 18

2.4 THE HUMAN AS A DECISION-MAKER: A SURVEY OF CURRENT BEHAVIOUR MODELS 19
2.5 DETERMINISTIC FEEDBACK SYSTEMS 22
Modelling the behaviour of deterministic feedback systems 23
Predictability of deterministic feedback systems 24
2.6 CAN HUMAN BEHAVIOUR BE MODELLED BY A DETERMINISTIC FEEDBACK SYSTEM? 27
2.7 DISCUSSION: REMEDYING PERSISTENT HUMAN ERRORS 29
Common denominator of current behaviour models: induction 30
Approach proposed: deduction 30
2.8 CONCLUSION: STARTING POINTS FOR A NEW BEHAVIOUR MODEL 31
REFERENCES 32

3 MAINTAINING SENSING INFLUENCE: A MODEL THAT EXPLAINS HUMAN ERROR 37
3.1 INTRODUCTION 38
3.2 STARTING POINTS FOR THE BEHAVIOUR MODEL 39
3.3 A SHORT INTRODUCTION TO AUTOPOIESIS 41
The general pattern of living systems is autopoiesis 41
Structural change of living systems 43
Structural coupling of living systems 44
3.4 DISCUSSION ON AUTOPOIESIS AND HUMAN SOCIAL BEHAVIOUR 46
Human's existence is independent of resigning from social institutions 48
Social coupling is essential for a human's existence 48
3.5 HUMAN SOCIAL BEHAVIOUR REVISITED WITH AUTOPOIESIS 49
3.6 HUMAN BEHAVIOUR REPHRASED AS MAINTAINING SENSING INFLUENCE (MSI) CONDITIONED BY SOCIAL REPRESENTING 51
Human social coupling: social representing 51
Structural change: Sensing Influence and Noting Influence 53
Human behaviour explained with the critical process MSI 54
3.7 ERRONEOUS BEHAVIOUR EXPLAINED WITH MSI 56
Context change is MSI 56
Observer context and (in)consistent operator behaviour 58
The evident, situational and reference context 58

Erroneous behaviour reveals a change to an evident context 61
3.8 DISCUSSION 62
Non-objectivity or the involvement of the observer 62
No Predictions 63
No free choice 64
3.9 CONCLUSIONS 65
REFERENCES 65

4 RE-EXAMINATION OF TWO CASES OF PERSISTENT HUMAN ERROR 69
4.1 INTRODUCTION 70
4.2 UNDERSTANDING HUMAN FAILURES VIA MAINTAINING SENSING INFLUENCE . . 71
Outline of the Maintaining Sensing Influence (MSI) model 71
Human failures reveal a context discrepancy 72
4.3 CASE 1: FLIGHT 801 OF KOREAN AIR 73
Short outline of the case 73
Explanation for the crew's contribution given by the NTSB. 76
Explanation for the crew's contribution based on the MSI-model. 76
4.4 CASE 2: FLIGHT 1420 OF AMERICAN AIRLINES 78
A short outline of the case 78
Explanation for the crew's contribution given by the NTSB 80
Explanation for the crew's contribution with the MSI-model. 81
4.5 EXTENDING THE MSI-EXPLANATION OF BOTH ACCIDENTS TO AVIATION IN GENERAL 82
4.6 THE ADDED VALUE OF SI WITH REGARD TO SA 84
4.7 CONCLUSIONS 85
REFERENCES 86

5 MSI, A NEW CONCEPT TO RECOVER FROM UNDESIRED DECISION BEHAVIOUR 89
5.1 INTRODUCTION 91
5.2 THE DIFFICULTY WITH DECISION-MAKING 92
5.3 CONCEPT FOR EXPLAINING (UNDESIRED) DECISION-MAKING 93
Concept of Maintaining Organization 93
Concept of Maintaining Sensing Influence (MSI) 94

Maintaining Sensing Inlfluence through context change 95
Hypotheses 96
5.4 IMPLEMENTATION OF THE CONCEPT IN THE STUDENT COUNSELING DOMAIN. . 97
Justification for student counselling 97
MSI-based analysis of the lack of study progress that persists 98
MSI to explain undesired behaviour. 99
Regaining NI, MSI based interventions 100
Results using the MSI-concept 100
5.5 STUDENT COUNSELLING VERSUS AIRCREW INTERVENTIONS 102
5.6 SIGNIFICANCE OF THE MSI-CONCEPT 103
5.7 CONCLUSIONS 104
REFERENCES 105

6 TRAINING AIRCREWS TO RECOVER FROM HUMAN ERROR, A PROPOSAL FOR A TRAINING OUTLINE 107
6.1 INTRODUCTION 109
6.2 PRESENT AIRCREW TRAINING 110
6.3 MODEL USED FOR DEVELOPING THE TRAINING OUTLINE 112
Concept of Maintaining Sensing Influence 112
Maintaining SI through context change 113
6.4 HYPOTHESES FOR MSI-CONCEPT 114
6.5 IDENTIFYING NEEDS TO TRAIN RECOVERY FROM HUMAN ERROR 115
Maintaining NI is essential for positive training 116
An example of negative training 116
6.6 OUTLINE FOR TRAINING FOR RECOVERY 117
6.7 FUTURE IMPLEMENTATION 119
6.8 CONCLUSION 120
REFERENCES 121

7 DISCUSSION 123
7.1 HUMAN ERROR MANAGEMENT 124
Existing approaches to human error management. 124

A new approach to human error management: MSI-approach 125
Connecting the MSI-approach with the existing approaches 126
The contribution of the MSI-approach to Human Error Management 128
7.2 PROCEDURAL SYSTEM DEFENCES: RULE MANAGEMENT 129
7.3 ENGINEERED SYSTEM DEFENCES: MAN-MACHINE INTERACTION 131
7.4 RECOMMENDATIONS FOR FURTHER RESEARCH 131
Training for recovery from human error in aviation 132
Added value of the MSI-model in medicine 133
7.5 CONCLUSIONS 133
REFERENCES 134

SUMMARY 137

SAMENVATTING 141

DANKWOORD 147

CURRICULUM VITAE 151

Chapter 1

INTRODUCTION

The main objective of this thesis is to develop a conceptual behaviour model that gives starting points for designing additional measures to remedy human errors that persist, despite the measures that should have prevented them or at least have made their effects harmless.

"The task is not so much to see what no one has yet seen, but to think what nobody has yet thought, about that which everybody sees."

Erwin Schrödinger (1887-1991)

1.1 Background: difficulties with human errors that persist

On December 18, 2003 Federal Express Corporation flight 647 crashed while landing at Memphis International Airport in Memphis, Tenessee. The National Transportation and Safety Board (NTSB) that examined the accident, found that the first officer failed to properly apply crosswind landing techniques to align the aeroplane and to properly arrest the plane's descent rate before the touch down. As a result the aeroplane touched down extremely hard with its nose still pointed into the wind and veered off the right side of the runway. As for the captain of the flight they found that he failed to adequately monitor the first officer's performance and failed to take any corrective action during the final approach and landing, such as commanding a go-around or taking over the flight controls, which could have prevented the accident. The NTSB concluded that the crew performance was causal to the accident. Other factors such as weather and plane condition did not play a significant role (NTSB 2005).

As pointed out by Bainbridge (1997), Hollnagel (1996), and Maurino *et al.* (1995), human operators remain essential for establishing high performance and safety of high reliability domains, such as aviation, nuclear power industry, air traffic control and medicine. This applies especially for situations that cannot be defined clearly. However, as the example above points out, human operators can jeopardize the system safety as well. Unfortunately, several studies into aviation accidents revealed that, although the accident rate has decreased to a substantially low rate in the last two decades (see Figure 1.1 and Figure 1.2), the share of crew involved accidents remained relatively high compared with the share of accidents attributed to technical errors, as is shown in Figure 1.3 (Boeing 2005). Recently, the concern has increased of not being able to reduce the share of crew-involved accidents further, not only because of the high impact of a possible accident caused by a sequence of human errors, but also because of the expected extra growth of the number of flight movements. If the accident rate is to be reduced further, more emphasis has to be placed on remedying human error, considering the relatively high share of human errors (Shappell and Wiegmann 2001).

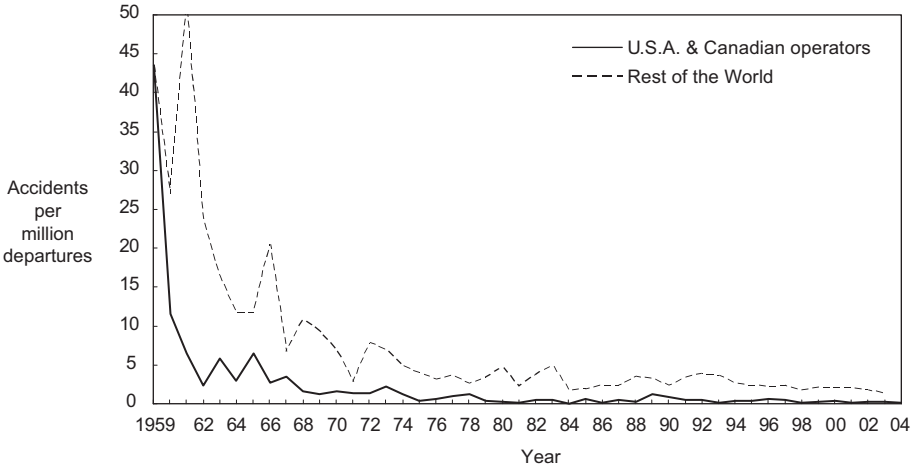


Figure 1.1 *Accident rate (number of accidents per million departures) of USA and Canadian commercial airlines (solid line) and world wide commercial airlines (dashed line) from 1959 through 2004 (adapted from Boeing statistical summary May 2005)*

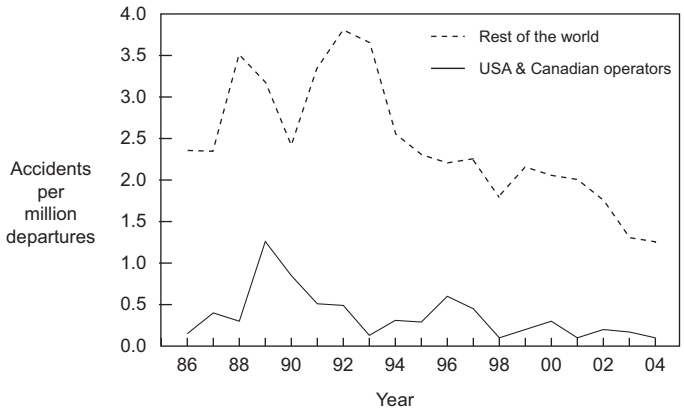


Figure 1.2 *Accident rate of USA and Canadian commercial airlines (solid line) and world wide commercial airlines (dashed line) from 1986 through 2004 (adapted from Boeing statistical summary May 2005).*

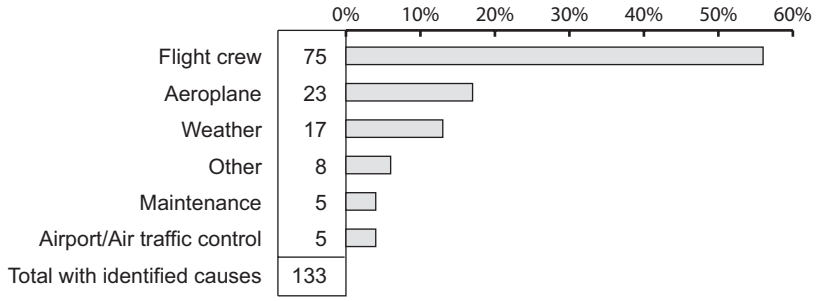


Figure 1.3 *Accidents by primary cause of world wide commercial jet aeroplanes from 1995 through 2004 (adapted from Boeing statistical summary May 2005)*

Historically, much effort has been put into minimising the share of human error in accidents and near misses. Inherent to the effect of lowering the accident rate in high reliability domains is that the most difficult and unpredictable forms of human error have remained. Researchers (NTSB 2001, Orasanu *et al.* 1998) found that a common decision-making error that remained is that operators stuck to their initial decision and, inherently, ignored evidence that did not support their initial decision, especially under high stress. This form of persistent human (decision) error is known as fixation, tunnel vision, cognitive lockup and preoccupation. One speaks of fixation when human operators maintain their course of actions in the face of opportunities to revise. Fixation is a behaviour pattern that typically becomes noticeable in high reliability domains, because accidents in these domains occur through a succession of small errors that only jointly can lead to an accident. Human operator teams involved in several major accidents in high reliability domains, such as Chernobyl and Three Mile Island have exhibited this behaviour pattern (Reason 1990; Perrow 1984).

Fixation and similar forms of persistent human error have been considered as intangible and impossible to rule out (Woods *et al.* 1994), not only because they occur unexpectedly, but also because they occur with experienced human operators. One of the main difficulties with persistent human errors is that they defeat the available remedial measures incorporated in the system that have been developed to prevent human error or make their effects harmless. These measures are the so-called system defences. High reliability systems incorporate several layers of system defences that each have to be defeated before an accident can occur. As is illustrated in Figure 1.4, fixation only becomes apparent when the human operator persists in committing errors, defeating the

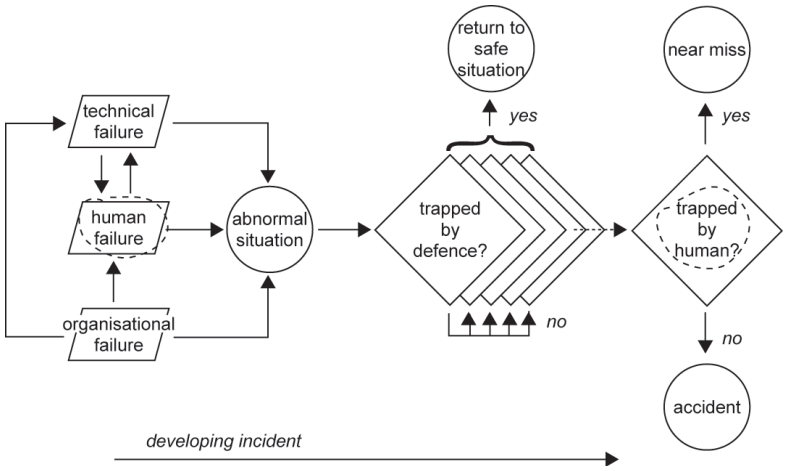


Figure 1.4 *A schematic representation of an incident/accident trajectory adapted from Van der Schaaf (1992). The development initiates with a minor failure originating from the technical system, the organisation or the human operator. The defences, being technical systems, procedures and trained human operators, have been integrated in the system design to trap the initial failures before they lead to a near-miss or accident. Note the double role of the human operator who is both a system defence and a threat to the system defences.*

successive defence layers. When fixation develops the last measure to avert a possible accident is to remedy it instantaneously. This remedial measure has to differ essentially from the available remedial measures incorporated in the system defences, because apparently, these do not satisfy in case of fixation.

Many measures have been developed that prevent human error from occurring altogether by urging the operator(s) to take action when the situation requires it, or that at least make the effects of human error harmless when they do occur. However, these measures are based primarily on the state of the environment (i.e. the system they are supervising) and only secondarily on the state of the human operator. Therefore, the focus of this research is on human operators that persist to err, because there appears to be some room left for improvement. Note, that although human operators who persist in erring are the subject of investigation, it is not the intention to blame them.

1.2 OBJECTIVE

To be able to develop measures that can remedy instantaneously persistent human error, an understanding of human operator behaviour is necessary, appropriate for this goal. Therefore, the main objective of this thesis is:

> *to develop a conceptual behaviour model that gives starting points for developing remedial measures that are able to establish a recovery from persistent human errors for which the available remedies do not satisfy any more.*

To be able to fulfil this objective, the first step is to explore the existing difficulties with remedying instantaneously persistent human error like fixation. Therefore, a more sophisticated understanding of remedying errors is required. Although remedying human error has been explored exhaustively and with success, the practice of successfully remedying persistent human errors is still poorly known.

The next step is to develop a new conceptual behaviour model that formulates the essential process that is involved in both human successful behaviour and human erroneous behaviour. This model has to overcome the difficulties of the existing behaviour models with regard to persistent human behaviour. However, it may not contradict with the existing behaviour models because they have shown their effectiveness with regard to lowering the accident rate in high reliability domains.

1.3 MOTIVATION OF CHOICE FOR AVIATION

The domain chosen to conduct the empirical part of the research is aviation, specifically commercial airline pilots flying an aeroplane. Of all high reliability domains, commercial aviation is considered to be the most reliable and safest domain, having far less than its fair share of accidents (Reason 2000). Although a significant reduction of the accident rate is the goal of the FAA (Wiegmann and Shappell 2001), our choice for aviation has primarily been influenced by reasons of research opportunities.

The outstanding performance and safety levels achieved in aviation are used as means for a further research into the problem of persistent human error. Moreover, extensive accident reports are available that offer detailed information for our research in the decision-making domain. Other domains do not offer such an easily accessible public da-

tabase of detailed accident reports, which makes it more difficult to explore the problem of fixation.

For the success of this research, human error should emerge spontaneously for the observer. For when an observer deliberately succeeds in forcing a human operator to commit an error, the action of the operator, seen from the observer's viewpoint, is not a human operator error any more. Strictly viewed, when the operator fails to commit the error expected by the observer, he errs in the context of the experiment. In many experiments this is practically the case: the solution is embedded in the experiment and the observer who forces the error implicitly knows the solution even before the problem emerges. To circumvent this paradox, the observer ideally has to wait until a human operator error occurs spontaneously during the natural course of events. Aviation fulfils this requirement of spontaneity and is therefore an optimal domain for conducting the empirical part of our research. Every year, commercial airline pilots have to participate in both a so-called "prof-check" training session and a recurrent training session in a full flight simulator. These sessions are fully incorporated in their daily routine and matter for their career. Errors committed in such a session are not forced by the instructor who supervises the sessions.

In addition to the requirement of human error that emerges spontaneously, the error should be easily recognized based on a norm. In aviation pilot behaviour is highly protocolized in the Standard Operation Procedures. Erroneous behaviour therefore is assumed to be recognized easier for a non-pilot than in comparable high reliability domains such as medicine, which has many unwritten rules unknown to a non-medical observer.

1.4 METHODOLOGY APPLIED

The methodology applied in this research is taken from the hypothetical-deductive method applied by Maturana and Varela (1992). They model behaviour of living systems by proposing an essential process that is universal to all specific forms of behaviour of living systems. Consequently, it is not necessary to incorporate all possible situations that a living system such as a human operator may encounter, to be able to represent all specific characteristic forms of behaviour. Therefore our objective is to reformulate the phenomenon of human error by reconsidering the universality of human error in human behaviour.

This methodology is opposite to the approach proposed by Rasmussen (1983), who abandoned the idea of a general behaviour model and adopted the idea of a framework consisting of several specific behaviour models for defined categories of work/task situations. Because these specific models have proved to be successful, our newly developed behaviour model is then acceptable when it does not contradict nor excludes the existing behaviour models or the existing approaches to human error.

1.5 Outline of this thesis

The main part of this thesis concerns the theoretical development of a conceptual behaviour model, preceded by a theoretical re-examination of the existing behaviour models for remedying human error. The other chapters are more practically oriented. Each practical chapter offers a short outline of the newly developed model, and therefore, can be read independently from the theoretical chapters.

Chapter 2 shows the need for a new behaviour model for remedying persistent human errors. Therefore the essential requirement for successfully remedying errors is identified for exploring the existing behaviour models. This requirement has been used to define the criteria for the behaviour model to be developed.

In Chapter 3, a behaviour model is developed based on the necessities identified in Chapter 2. Maturana and Varela's conceptual theory of autopoiesis is introduced in short. This theory is considered to be sufficient for describing human behaviour but insufficient for human social behaviour. Based on logical reasoning, it is made plausible that the concept of autopoiesis applies as well to human social behaviour. Then, in addition to the concept of Maturana and Varela, a new conceptual human social behaviour model is presented, the MSI-model. This model describes human social behaviour as continually **M**aintaining **S**ensing **I**nfluence.

Chapter 4 carries out a re-examination of two crew-involved accidents in aviation based on the newly developed MSI-model. The National Transportation and Safety Board already examined these two cases and concluded that the accidents were likely due to the crew's poor situation awareness. Our re-examination of these cases using the MSI-approach has resulted in new insights with regard to persistent human errror.

In Chapter 5 the added value of the MSI-model is investigated by applying it in a student counselling setting for students who, after they passed their Bachelor examination,

showed persistent undesired study behaviour regarding their intent to become a Master of Science. A comparison is drawn between the student counselling setting and the cockpit setting.

In Chapter 6 a proposal for a training outline is discussed that should aim specifically at remedying persistent human error. A number of type-qualification training sessions for aspirant airline pilots and a number of compulsory recurrent training sessions for airline pilots were attended and evaluated based on the model.

Chapter 7 discusses the added value of the newly developed model to human error management, rule management and man-machine interaction with regard to remedying persistent human error.

REFERENCES

Bainbridge L (1983) Ironies of Automation. *Automatica*, **19:** 775-779.

Boeing (2005) Statistical Summary Of Commercial Jet Airplane Accidents, World Wide Operations 1959-2004. Seattle Washington, Aviation Safety Boeing Commercial Airplanes.

Maturana HR, Varela FJ (1992) *The Tree of Knowledge; The Biological Roots of Human Understanding*. Shambhala Publications.

Maurino DE, Reason J, Johnston N, Lee RB (1995) Widening the search for accident causes: a theoretical framework. In: *Beyond Aviation Human Factors; Safety in High Technology Systems*. Aldershot, UK: Avebury Aviation, Ahgate Publishing Ltd., 1-29.

NTSB (2001) Runway Overrun During Landing American Airlines Flight 1420, McDonnell Douglas MD-82, N215AA Little Rock, Arkansas. NTSB/AAR-01/02. Washington DC, National Transportation Safety Board.

NTSB (2005) Hard Landing, Gear Collapse of Federal Express Flight 647, Boeing MD-10-10F, Memphis Tennesse. NTSB/AAR-05/01. Washington DC, National Transportation Safety Board.

Orasanu J, Martin L, Davison J. Errors in Aviation Decision Making: Bad Decisions or Bad Luck? (1998) Fourth Conference on Naturalistic Decision Making, Warrenton Virginia. NASA.

Perrow C (1984) *Normal Accidents: Living with High-Risk Technologies*. New York:

Basic Books.

Rasmussen J (1983) Skills, Rules, and Knowledge: Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-13:** 257-266.

Reason J (1990) *Human Error*. New York: Cambridge University Press.

Reason J (2000) Human error: models and management. *BMJ*, **320:** 768-770.

Shappell SA, Wiegmann D. (2000) The Human Factors Analysis and Classification System - HFACS. DOT/FAA/AM-00/7. Washington, FAA.

Van der Schaaf TW (1992) *Near miss reporting in the chemical process industry*. PhD Thesis: Technische Universiteit Eindhoven.

Wiegmann DA, Shappell SA. (2001) A Human Error Analysis of Commercial Aviation Accidents Using the Human Factors Analysis and Classification System (HFACS). DOT/FAA/AM-01/3. Washington DC, FAA.

Woods DD, Johannesen LJ, Cook RI, Sartert NB (1994) *Behind Human Error: Cognitive Systems, Computers, and Hindsight*. Columbus: CSERIAC.

Chapter 2

REMEDYING PERSISTENT HUMAN ERROR: THE NEED FOR A NEW BEHAVIOUR MODEL

The aim of this chapter is to show the need for a new behaviour model for remedying human errors that continue to exist despite the measures that should have prevented or remedied them. A review of the current human behaviour models is the basis for our conclusion that a different view on human behaviour, and therefore on human error, is required for remedying persistent human error successfully.

"The hard sciences are successful because they deal with soft problems, the soft sciences are struggling because they deal with hard problems."

Heinz von Foerster, 1972

2.1 INTRODUCTION

In system theory and control theory it has been shown that effectively remedying system error can only be achieved when the dynamics of the system to be controlled are known (Flach 2004; Stassen *et al.* 1990; Conant and Ashby 1970). In addition, the Law of Requisite Variety postulated by Ashby (1956, 1960) states that for an effective remedy of system errors, the richness in variety of the controller's behaviour should match the richness in variety of the system to be controlled. Because the system to be "controlled" is the human operator himself, a behaviour model representing the dynamics of human behaviour is required for remedying human error effectively. Persistent human error implies that the richness in variety of the operator's behaviour exceeds the richness in variety of the model-based remedies. Hence, the human behaviour model applied contains an insufficient richness in variety of human behaviour. Hollnagel has repeatedly argued for the need of a theory of behaviour that accounts for the variety of human performance (Hollnagel *et al.* 2006).

Ashby (1960) and Capra (1996) argued that humans can be considered as *deterministic* feedback systems, although not simple ones. Whereas Rasmussen (1983) argued that humans cannot be regarded as simple *deterministic input-output* systems, but that they have to be regarded as *decision-making feedback* systems. Decision-making, as an alternative of determinism, is considered to be a component of feedback behaviour (Heylighen 1998). Feedback means that the system is oriented towards a goal. A feedback system that moves towards its goal in a unique way is considered to be deterministic. A feedback system that can move towards its goal in several possible ways is considered to be non-deterministic. Rasmussen considered humans as *decision-making* systems because the behaviour actually shown by humans could not be predicted by the existing deterministic feedback models. To explore the issues of modelling human operator behaviour with regard to remedying persistent human error, both deterministic feedback systems and its alternative, decision-making feedback systems are discussed.

The next section addresses the aspects of modelling the behaviour of feedback systems with regard to remedying error. Section 2.3 discusses the aspects of modelling behaviour when the feedback system is considered to be non-deterministic. Then a short survey

is given of the existing human behaviour models. For details about the mentioned behaviour models the reader is referred to Hollnagel (1993), Reason (1990), Rasmussen (1986) and Stassen *et al.* (1990). Section 2.5 addresses the aspects of modelling behaviour when the feedback system is considered to be deterministic. It will be shown how deterministic feedback systems can generate behaviour that is unpredictable to the observer. In Section 2.6, the characteristics of human behaviour are translated into a deterministic feedback model for human systems. Section 2.7 discusses what the current behaviour models have in common and what they lack with respect to remedying persistent human error. Finally, the necessities are defined for a behaviour model to be developed that can give starting points for remedying persistent human error.

2.2 MODELLING THE BEHAVIOUR OF A FEEDBACK SYSTEM WITH REGARD TO REMEDYING ERROR

The aim is to explore the issues of developing a human behaviour model that could remedy persistent human error successfully. To obtain a reference point this section firstly discusses these issues with regard to a system. Ashby (1956) defined the behaviour of a system as a change from one system state to another system state, in which the system's state at a particular time is represented by the numerical values of its variables at that time. The system's behaviour is thus formed by a series of output states. The system's behaviour *model* is constituted by a series of output states related to a series of input states. The variables taken to model the system's behaviour depend on the goal the observer wants to achieve with the system. The observer chooses the input states such that the output states should realize this goal. Based on the behaviour model, the observer assess the system's behaviour. Behaviour that is unexpected to the observer regarding the model is viewed as a system error. The behaviour is considered to be erroneous when it does not lead to the desired goal.

For remedying the system error, the observer has to restore the predictability of the system's behaviour, otherwise all efforts will be like shots in the dark. The observer's assumption about the system determines the way the predictability is restored, and consequently the way the error can be remedied. The unpredicted/erroneous behaviour can be clarified in two ways: an input state has not been taken into account, or the set of variables is insufficient. Assuming that an input state was omitted, the observer can restore predictability either by taking this state into account or by assuring that the system does not encounter this input state once more. Assuming that the series of input states

regarding the model is complete, the observer may ascribe erroneous behaviour to (see Figure 2.1):

- incomplete observability of the system, i.e. it contains unobservable states;
- incomplete accessibility of the system, i.e. the system is adapting
- non-determinism of the system.

If the observer ascribes the unpredictability of the output to an incomplete observability of the system, he assumes that the system contains unobservable internal variables that take internal states unknown to him. That is, the system's output state depends not only on its input state but also on its internal states. Consequently, the relation between input state and output state is not single-valued. The observer may take into account the system's previous states (its past) to restore predictability. In other words, by assuming the existence of memory or experience. By taking the past into account the observer is able to fill in the internal states unknown to him and therefore to transform the indefinite input-output relationship to a definite one. The fewer the observable variables the more the observer is forced to regard past events as influencing the system's present behaviour (Ashby 1956).

If the observer ascribes the output's unpredictability to its incomplete accessibility, its present relationship between its variables depends on its previous output. In other words, its output state once observed for a given input state will most likely be different for the same input state given later. A system whose input-output relationship changes due to its output is called an adapting system. The output of an adapting system can be unpredictable even when the system is completely observable. The observer may restore the predictability by replacing the old system for a new system. By replacing the old system for a new system the observer "resets" the input state. In this case, it is important to bear in mind that the predictability belongs to the group of systems and not to the individual system. Note, that the dependence of the present input-output relationship on its previous output differs from a dependence of the present input on its previous output. For each feedback system applies that the present input depends on its previous output, however, only for adapting feedback systems applies that its present input-output relationship depends on its previous output.

If the observer assumes the system to be non-adapting and completely observable, for example a dice, he may ascribe the unpredictability of the system to non-determinism of the system. The observer can only restore the predictability of the system's output

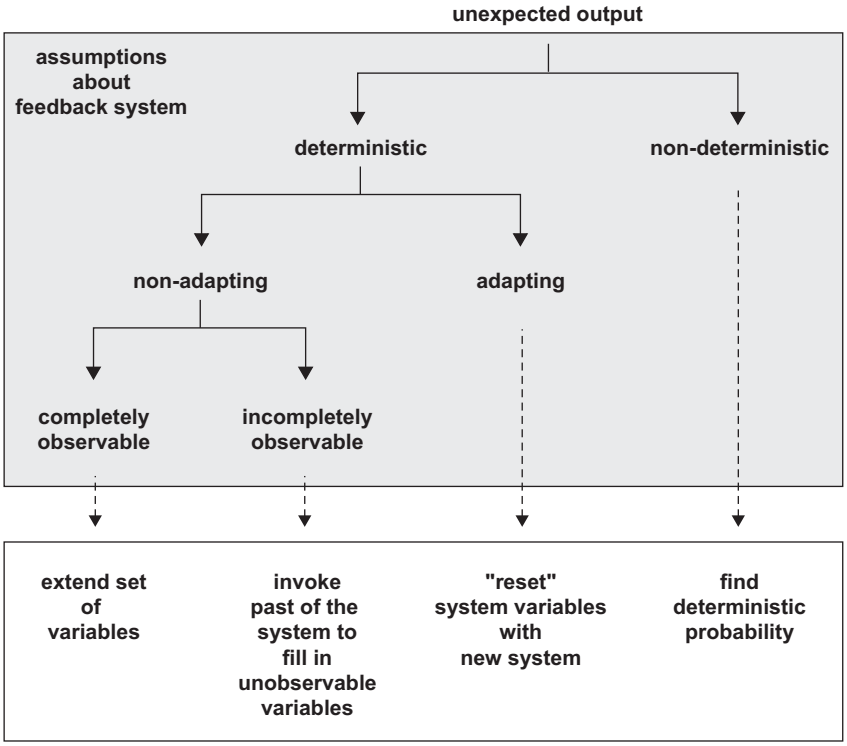


Figure 2.1 *Depending on the assumptions the observer makes about the system, the possibilities to remedy predictability differ. If the system is assumed to be completely observable, non-adapting and deterministic then the possibilities are constrained to extending the set of observable variables. If instead it is assumed to be incompletely observable then the observer may invoke the past of the system. If it is adapting the observer has to reset the input by replacing the old system for a new one, independent of its observability; unobservability only complicates the situation for the observer. If it is impossible for the observer to restore predictability completely, then he may look for a frequency of occurrence of the various output states.*

by looking for a frequency of occurrence of the various possible output states. Then probability becomes the variable of the model and not the system variables that provide the probability. In practice this method is restricted to behaviour that is not too rich in variety (Capra 1996; Ashby 1956).

When it is impossible for the observer to restore the predictability of the system, it does not necessarily mean that the system is non-deterministic. Determinism refers to a system and is an assumption of the observer, whereas predictability refers to the system's output and is an observer assessment. Predictability is thus not the same as determinism. Although predictability implies determinism, determinism does not imply predictability. Non-determinism means unpredictability, but unpredictability does not necessarily mean non-determinism (see Figure 2.1).

Restoring predictability may become a never-ending pursuit for the observer when the system is both adapting and incompletely observable. Because for each unobservable variable the observer must invoke a past state of the system to restore its predictability. When, in addition, the numerical value of these unobservable internal variables (the internal state) also depends on the previous output, each time the internal state changes unknown to the observer, he must invoke a past state to repair the system's predictability. This implies that the system's present behaviour is dependent on the particular sequence of input states the system encountered reaching into the remote past (Von Foerster 2003).

2.3 DECISION-MAKING FEEDBACK SYSTEMS

Modelling humans as *decision-making* feedback systems emerged because in practice the variety of behaviour generated by the existing *deterministic* feedback models was too limited with regard to the variety of the behaviour generated in reality. Since determinism implies that given the input state the human's behaviour is considered to be determined, non-determinism assumes humans as to have the ability to decide based on rationality, which action (output) is the best to obtain a preferred goal. Because the observer considers the decision-maker to have feedback, they necessarily share the same goal. When they do not share the goal to be achieved, the observer does not regard the decision-maker to have feedback.This viewpoint has led to behaviour models in which the human is modelled as a decision-maker (or problem-solver).

BEHAVIOUR OF DECISION-MAKING FEEDBACK SYSTEMS

Decision-making considers two situations, the present situation (or the initial situation) and the desired situation, the goal (or the end situation). The present situation in relation to the goal constitutes the decision-situation. The decision-maker's task is to transform

the initial situation into the goal. In each decision-situation the decision-maker has to choose from a set of available actions for bringing the present situation closer to the goal.

Considering that the behaviour of the decision-maker is predictable, the decision-making approach distinguishes two extreme decision-situations, the very strictly defined or trivial decision-situation and the very ill-defined or random decision-situation (see Figure 2.2). The predictability of the decision-maker's behaviour (the way he reaches the goal) depends on how the decision-maker experiences the decision-situation. In a trivial decision-situation the behaviour shown by the decision-maker is predictable to the observer. In this special case of decision-making, the observer assumes that the decision-maker is facing a set of possible actions of which only one is guaranteed to bring the present situation closer to the goal. This means that in each decision-situation the decision-maker's actions with regard to the goal (the next decision-situation) are determined. An example of trivial decision-making is a thermostat regulating the temperature. Trivial decision-making can be represented by a definite relationship.

In a random decision-situation, the way the decision-maker will take to reach the goal is unpredictable for the observer. He assumes that the decision-maker is facing a set of possible actions of which none is guaranteed to bring the present state closer to the goal state. When the consequences are undetermined and, in addition unknown to the decision-maker, the only way for him to find the right action is to try and see what the next state is. In the view of the observer, the decision-maker then shows trial-and-error behaviour, searching at random for a next situation closer to the goal and retaining this situation when it is closer to the goal. Assuming there is a goal, there is no guarantee that the decision-maker will reach it in a finite time (Heylighen 1998).

In practice the middle way between these two extremes of decision-situations is common. The way the decision-maker reaches the goal is not solely determined by the decision-situation, but it is not random either. Rather, the decision-maker seems to have some idea of the consequences of the possible actions. The richness in variety of his decision-making behaviour seems not only to be constrained by the decision-situation but also by the decision-maker's past events.

Restoring predictability of decision-making feedback systems

Considering that the decision-maker shows unpredictable behaviour to the observer, the observer ascribes the unpredictability of the decision-maker's behaviour to the un-

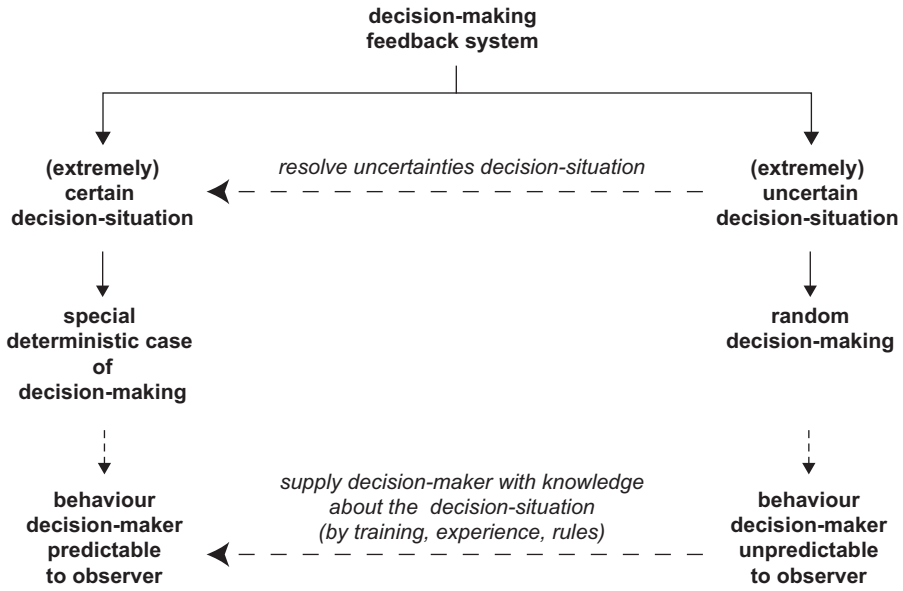


Figure 2.2 *The unpredictability of the decision-maker's behaviour is ascribed by the observer to decision-situation the decision-maker experiences given that the decision-maker is aiming at the goal as supposed by the observer. In case of a certain decision-situation the behaviour of the decision-maker is then supposed to be predictable. In case of an uncertain decision-situation the observer can restore predictability by changing the uncertain decision-situation into a certain situation, or by supplying the decision-maker with knowledge about the decision-situation either by supplying the decision-maker with rules or by generating experience with these kind of situations.*

certainty of the decision-situation to the decision-maker. To restore the predictability of the decision-maker's behaviour the observer may:

1. improve the decision-situation for the decision-maker by resolving the uncertainties
2. supplying the decision-maker with extra knowledge about the uncertain decision-situation
3. by looking for a statistical determinacy

Ad 2. The observer may restore the predictability of the decision-maker's actions by supplying him with knowledge about the decision-situation. This knowledge can be sup-

plied to the decision-maker in the form of a written rule: if the decision-situation is X then perform action Y. For the decision-maker the decision-situation then should become trivial and his behaviour should become predictable to the observer. Another way to restore the predictability of the decision-maker's actions is to introduce the experience of the decision-maker from earlier events. In case of experience, the decision-maker's actions seem to be controlled by a set of unwritten rules (= rules stored in the decision-maker's memory) which has been proven successful in decision-situations that he has encountered earlier. This turns the unpredictable behaviour of the decision-maker into predictable behaviour.

Ad 3. This approach supplies the observer with a statistical predictability but does not give direct starting points to develop remedies against errors.

2.4 The human as a decision-maker: a survey of current behaviour models

For its crucial role in high reliability domains, human operator behaviour has been attracting much interest the last decades. Therefore, a great variety of behaviour models exists ranging from the most simple to the most sophisticated. It is not the intention to go into detail about these behaviour models, neither to cover them all. Only those human behaviour models are discussed that are referred to frequently with respect to improving the performance of the human operator in high reliability domains.

The classification of the behaviour models that has been used most commonly in the man-machine Systems domain is the Skill-, Rule-, and Knowledge-based behaviour (SRK) framework introduced by Rasmussen (1983, 1986).This classification is related to the classification of decision-making. In a Skill-based task situation the task is very well defined and leaves very little freedom for the decision-maker. This is the equivalent of trivial decision-making. At the Rule-based task situation level, the human's decision-making behaviour is considered to be controlled by a set of rules: "if situation X, then do Y." This level is equivalent with non-trivial decision-making in which the decision-making has obtained or is supplied with rules about the (decision-)situation. It seems as if the action taken by the decision-maker is not only determined by the decision-situation but also by the set of rules available to the decision-maker. At the Knowledge Based task situation level, such knowledge is not available and the human's decision-making behaviour is dependent on his experience and creativity. Experience is considered to

be the reason for choosing a particular action out of the set of possible actions to reach the goal.

The theory used for modelling human behaviour at the Skill-based level stems from control theory. Well-known examples of skill-based behaviour models are the Describing Function Model of McRuer and Jex (1967) and the Optimal Control Model of Kleinman *et al.* (1971). These models are decision-situation sensitive and the set of possible actions is very limited. For an extensive treatment of this kind of models and their limitations in use is referred to Stassen (1975 in Dutch).

The main theory used for the modelling of human behaviour at the rule-based level is the Information Processing Theory. The information processing theory of cognitive behaviour became the prevalent view of cognitive science in the 1960s. Allen Newell and Herbert Simon (1972) argued that the way humans behaved could be compared to the way computers "behaved." The basic idea was that human cognition resembles that of a computer to such an extent that cognition can be represented as information processing. A computer processes information, that is, it manipulates symbols based on certain rules. The most commonly used models for this task level are the model of human information processing of Wickens (1992), and the Situation Awareness model of Endsley (1995, 1999). The model of information processing became subject of serious questioning in the 1970s due to the acknowledged deficiencies of the computational view. The first argument put forward was that information processing is based on sequential rules, whereas practical observations contradicted this assumption. The most ordinary tasks were done faster than was physically possible when simulated sequentially. Another argument put forward against human information processing models is that human behaviour is so situation specific that it is impossible to induce a general model of human behaviour (Bainbridge 1997).To resolve these shortcomings of information processing models, the neural networks theory emerged (Capra 1996).

The knowledge based level of human behaviour is the level that has become the main subject of investigation in the last two decades, especially with regard to problem solving in ill-defined uncertain decision-situations (Zsambok and Klein 1997).

The common approach to reduce the set of possible actions in each decision-situation has been to limit the study of human behaviour to specified task situations that are related to a goal the human decision-maker has to fulfil. Basically two approaches can be distinguished to study human behaviour in specified task situations: the laboratory setting

and the naturalistic or field setting. The laboratory settings are very restricted environments and have the disadvantage that they do not take into account the effects that real-life settings may have on a human decision-maker. A representative of the laboratory setting is the Behavioural Decision Theory (e.g. Doherty (1993)). In behavioural decision theory (BDT), the decision-maker's behaviour is assessed by comparing the observed actions of the decision-maker to standards obtained from Utility Theory and Bayesan Probability. Since the late 1980s several scientists among whom Hutchins and Palen (1995) and Klein and Calderwood (1997) and Hollnagel (2001) have emphasized the importance of studying behaviour in context, i.e. in natural (decision-) situations, because the richness in variety of behaviour in real-life settings was proven to be much larger. This approach adheres to the idea that one should start from the practical decision-situations arising from the field settings. Naturalistic Decision Making (NDM) researchers study the way human operators use their experience to make decisions in a specific but natural task situation (Zsambok and Klein 1997). In NDM the standards to assess operator behaviour are replaced by standards that are obtained by cognitive task analysis that identify the cognitive processes that are required to perform a task in a specific situation (Lipshitz 1997).

Decision theory describes how a decision-maker chooses an action among the possible actions (Heylighen 1998). Decision theory uses a criterion that ranks the set of possible actions based on their possible consequences with respect to the goal. The premise for this criterion is that it has to be consistent with the decision-maker's preferences. The assumption for prediction is that the decision-maker decides for the action with the highest ranking. The criterion used by the observer depends on the decision-situation. Two main decision-situations are distinguished: decision under certainty and decision under uncertainty. Decision under certainty means that each action of the set of possible actions has only one consequence. Therefore, each action is associated with a consequence. The decision-maker's preferences for these actions are represented by a value function that ranks the consequences and thus the actions. Decision under uncertainty means that each action will have one of several consequences. For the case the observer knows the probability of occurrence for each consequence, each alternative can be associated with a probability distribution. The decision-maker's preferences are estimated by a utility function. The utility is a measure of the desirability of the several consequences of each action. For the case where the observer does not know the probabilities, there are two approaches to estimate the decision-maker's preferences. The first approach uses the min-max rule (Kahneman *et al.* 1982), which means that the decision-maker chooses the action such that the worst possible consequence of the chosen action is

better than the best possible consequence of any other action. The second approach reduces the decision situation under uncertainty with unknown probabilities to a decision situation under uncertainty with expected probabilities based on expert assessments or on analysis of previous decisions made in similar situations (Heylighen 1998).

Decision theory itself is not concerned with either defining the goal, designing the possible actions or assessing the consequences. It considers them as given or previously determined (Heylighen 1998). Decision theory incorporates the following basic assumptions:

- the set of possible actions is complete and is available to the decision-maker;
- the decision-maker and the observer share the same goal;
- the criterion used, ranks the actions such as to be consistent with the given or previously determined decision maker's goal and preferences.
- the decision-maker decides rationally.

Note also, that decision theory restores the predictability of the decision-maker's behaviour, but it does not give starting points for remedying human error because the restoration is based on probabilities.

Although the decision-making approach has enlarged the richness in variety of modelled behaviour compared with the deterministic approach, the actual behaviour shown by human decision-makers is still not completely consistent with the richness in variety generated by the decision-maker. In practice, experienced decision-makers do not generate multiple courses of actions, nor weigh the pros and cons and choose the right course based on some evaluation criterion. Rather they seem to draw on their intuitive understanding (Dreyfus 1997).

2.5 Deterministic feedback systems

As discussed in Section 2.2, unpredictability of a system's output does not necessarily mean that the system is non-deterministic. In this section it is shown that a strictly deterministic feedback system that is completely observable can generate behaviour that is so rich in variety that it is unpredictable for the observer.

MODELLING THE BEHAVIOUR OF DETERMINISTIC FEEDBACK SYSTEMS

It is considered that every real deterministic feedback system embodies an infinite number of possible system variables. A model of the system's behaviour starts with choosing the essential variables out of the vast number of system variables such that the input-output relationship is definite. A pendulum, for example, embodies system variables like its position, velocity, acceleration, but also length, mass, temperature, tensile strength, electric conductivity etc. The observer selects the number of system variables that are minimally required to obtain a definite input-output relationship of the behaviour of interest, the essential variables. The number of essential variables depends on how closely the system variables are related to each other in respect to the behaviour of interest. Thus, if the observer's objective is to study the 2D-swing of a pendulum for all its possible initial states given its length, the essential variables that give a definite input-output description of the swing would be the pendulum's angle from the vertical and its angular velocity. The pendulum's angle alone is insufficient, because the swing's input-output relationship will then be indefinite. Other system variables, such as the pendulum's tensile strength, can be excluded as long as they have no or an insignificant effect on the 2D-swing (i.e. they remain constant).

A definite model of system behaviour consists of the initial state cq present state also called "input state," and a statement of transformation. The definition of input states and input signals (i.e. stimuli) are different, hence some remarks are made at this point. An input signal (stimulus) may affect the observer, however it does not necessarily have to affect the system. Therefore, an input signal cannot indiscriminately be regarded as input state for the system. On the other hand, an input state influences the system's behaviour by definition because it is a system state. The input state may be the effect of some input signal but the way the input signal influenced the input state can be left out of consideration. It is the input state that matters for the system's behaviour (Von Foerster 2003).

When $x_{n\times 1}(t_1)$, $x_{n\times 1}(t_2)$ represent the respective input state and output state of any deterministic system, its behaviour can be described with the following input-output relationship:

$$x_{n\times 1}(t_2) = T(x_{n\times 1}(t_1)) \tag{2-1}$$

When the system is known to obey certain laws that are based on a theory, such as the laws of Newton then the input-output relationship can be described in an algebraic form by a non-linear and time-variant differential equation of the n^{th}-order.

$$\dot{x}_{n\times 1}(t) = A_{n\times n}(x_{n\times 1}(t),t) \cdot x_{n\times 1}(t) \tag{2-2}$$

Equation 2-2 is a differential equation or a process equation and describes explicitly for each system variable its future value (end state) as a function of its present value (initial state). The form of equation 2-1 is a general form and can be used always. Equation 2-2 is a specific form of equation 2-1 and is applicable for systems that have feedback, and therefore it is applicable for biological systems like humans (Rosenblueth *et al.* 1943; Von Foerster 2003).

Predictability of deterministic feedback systems

Given an input state and an input-output relationship (the behaviour model) of the deterministic feedback system, the system's behaviour is definite and predicting behaviour basically consists of generating a series of successive states a system will take. The series of successive states is called the system's trajectory or its dynamic behaviour (Ashby 1956). When the input-output relationship is in the form of Equation 2-1, the dynamic behaviour can be found by taking the series of x, T(x), T(T(x)), T(T(T(x))), and so on. When the input-output relationship is a set of differential equations, then predicting the system's behaviour is equivalent with solving the differential equations given an input state. In this subsection, the essence of solving differential equations is briefly sketched, to illustrate the different levels of predictability of dynamic behaviour. For an extensive treatment of solving differential equations see, for instance, Boyce and Diprima (1992).

For the predictability of the behaviour of a deterministic feedback system, the concept of stability is essential. The basic idea involved in the concept of stability is the state of equilibrium. The state of equilibrium is the state that remains unchanged (i.e. invariant) by transformation. Two kinds of equilibrium states can be distinguished: the static and the dynamic. The static equilibrium state is the solution of the differential equation with constant coefficients:

$$\dot{x}(t) = A_{n\times n} \cdot x(t) = 0 \tag{2-3}$$

There are three types of static equilibrium states: the (asymptotically) stable, the critically stable and the unstable. If the static equilibrium state is stable (either asymptotically or critically), then specific trajectories that start close to this equilibrium state will move to either the equilibrium state or the cycle. If the static equilibrium state is unstable then any trajectory that does not start exactly at this equilibrium state moves away from the equilibrium state. The dynamic state of equilibrium is the solution of the differential equation with varying coefficients of the A-matrix:

$$\dot{x}(t) = A_{n\times n}(x(t),t) \cdot x(t) = 0 \qquad (2\text{-}4)$$

Varying of the coefficients leads to a varying equilibrium state, which leads to varying dynamic behaviour. Varying dynamic behaviour is a change from dynamic behaviour to dynamic behaviour, that is, a change from one input-output relationship to another input-output relationship. In other words, varying dynamic behaviour is a change in the way of behaving. The dynamic equilibrium state itself can also be stable or unstable. An unstable dynamic equilibrium state is called a bifurcation point. In a bifurcation point the system's behaviour is characterized by a sudden change in dynamic behaviour (Capra 1996; Boyce and Diprima 1992). Deterministic feedback systems showing this varying dynamic behaviour are creative systems. Biological systems such as humans are prime examples of creative systems.

When the set of differential equations contains only one invariant equilibrium state, predicting the system's dynamic behaviour is fairly simple and exact for all possible initial states. Such a set of differential equations is categorised as a linear set of differential equations. For a set of linear equations it is possible to obtain the general form of the dynamic behaviour containing all specific trajectories. A specific trajectory follows from the general form of dynamic behaviour by substituting an initial state. In case of linear equations, the initial state is not of crucial influence on reaching the end state (i.e. the equilibrium state), although it has some effect on how this end state is reached. The equilibrium solution is said to be globally stable. Linear equations are considered to have one global basin.The general trajectory falls within the region of stability or the basin of attraction, irrespective of its initial conditions.

When the set of differential equations contains several equilibrium states that are invariant, predicting the system's dynamic behaviour for all possible initial states is usually much more difficult and approximate. Such a set of equations is called non-linear. The

greater the number of possible equilibrium solutions, the higher the degree of non-linearity. The equilibrium solutions of a non-linear set of equations can either be unstable or locally stable. An equilibrium solution is then said to be locally stable when there exist a set of initial states close enough to the equilibrium solution, for which the specific trajectories approach it. This is called the local basin or the linear region. In a local basin, the set of non-linear equations can be linearized. For each local basin another set of linearized equations applies. If the initial state is close enough to a local stable equilibrium solution the specific trajectory can be obtained by solving the set of linearized equations. The non-linear regions correspond to the unstable equilibrium solutions. In the non-linear region the local basin "taken" first is highly sensitive to small changes in states. The specific trajectory will differ accordingly. Imagine for example a pendulum with friction in a vertically up position, an unstable equilibrium solution. It is highly sensitive to small perturbations in its initial state which will determine whether it will fall left or right, approaching its stable local equilibrium solution.

When the set of equations contains equilibrium solutions that are subject to changes, predicting the system's dynamic behaviour for all its possible (initial) states is not only extremely complex, it is practically impossible. This set of process equations is said to be relentlessly non-linear. Relentlessly non-linear systems differ essentially from "common" non-linear systems, because their input-output relationship (the A-matrix) is not definite but influenced by the system's previous output. In other words, its previous dynamic behaviour influences it present dynamic behaviour. These systems are deterministic systems, but they are unpredictable: an output observed for an input given once will most likely not be the same for the same input given later (Von Foerster 2003). Because the A-matrix, representing the system's parameters, is variant, the number of possible equilibrium solutions can be endless, and linearizing about an equilibrium solution is not practicable.

Predictability of the behaviour of a deterministic feedback system depends on the regularity in variety of its behaviour (Ashby 1956; Hollnagel 2001), and not necessarily on its determinism. The level of predictability of behaviour is related to the regularity in variety which in turn depends largely on the number of equilibrium solutions. Whereas the dynamic behaviour of a deterministic feedback system that can be represented by a third-order linear equation is fairly regular in its variety because it is running through its only global basin, the dynamic behaviour of a third-order non-linear equation is far less regular in its variety because it can be running through subsequent local basins. The model of weather conditions developed by Edward Lorenz (1963) demonstrates how a

simple set of non-linear equations with constant coefficients that is strictly deterministic can generate dynamic behaviour that is very rich in variety and that can be extremely sensitive to tiny perturbations in its (initial) states. From practically the same initial states the dynamic behaviour can develop completely different. Hence, for non-linear equations it is practically impossible to calculate exactly long-term trajectories, because a tiny perturbation of the system's states in the neighbourhood of an unstable local equilibrium solution will change the dynamic behaviour essentially. The dynamic behaviour generated by the Lorenz equations becomes enormously rich in variety when its coefficients are subject to changes as well. The equilibrium solutions become variant, introducing an extra dynamic on top of the existing dynamic behaviour. For an extensive treatment of the dynamic behaviour that can be generated by non-linear equations is referred to Boyce and Diprima (1992) and (Wiggins 1990). Figure 2.3 summarizes this subsection.

2.6 Can human behaviour be modelled by a deterministic feedback system?

The human operator is an autonomous adapting system exhibiting an enormous richness in variety of behaviour. Autonomy refers to feedback systems, or circular causal systems. The term adapting refers to parametric values of the human system that change in interaction with their environment. Adapting is called creative when the human changes behaviour encountering a new environment, whereas it is called learning when the human changes behaviour in successive encounters with the same environment (McRuer and Jex 1967). Considering the continuous development of human behaviour, the change of equilibrium states can be viewed as a steady process. It is this steady change of equilibrium states that makes human behaviour extremely rich in variety and therefore unpredictable for the observer. When these characteristics are translated to a behaviour model based on the assumption of deterministic feedback this implies that humans have to be modelled as relentlessly non-linear systems embodying a vast number of essential variables. Only then can the richness in variety of behaviour generated by the behaviour model match the richness in variety of the human's actual behaviour.

An observer dealing with a "human system" that generates behaviour that is too rich in variety to be modelled through input-output analysis, may constrain its richness by reducing the set of possible states of the system until the system generates predictable behaviour. Near a stable equilibrium state, dynamic behaviour is regular. Limiting the (initial) system states to a reduced region implies in general that the number of equilib-

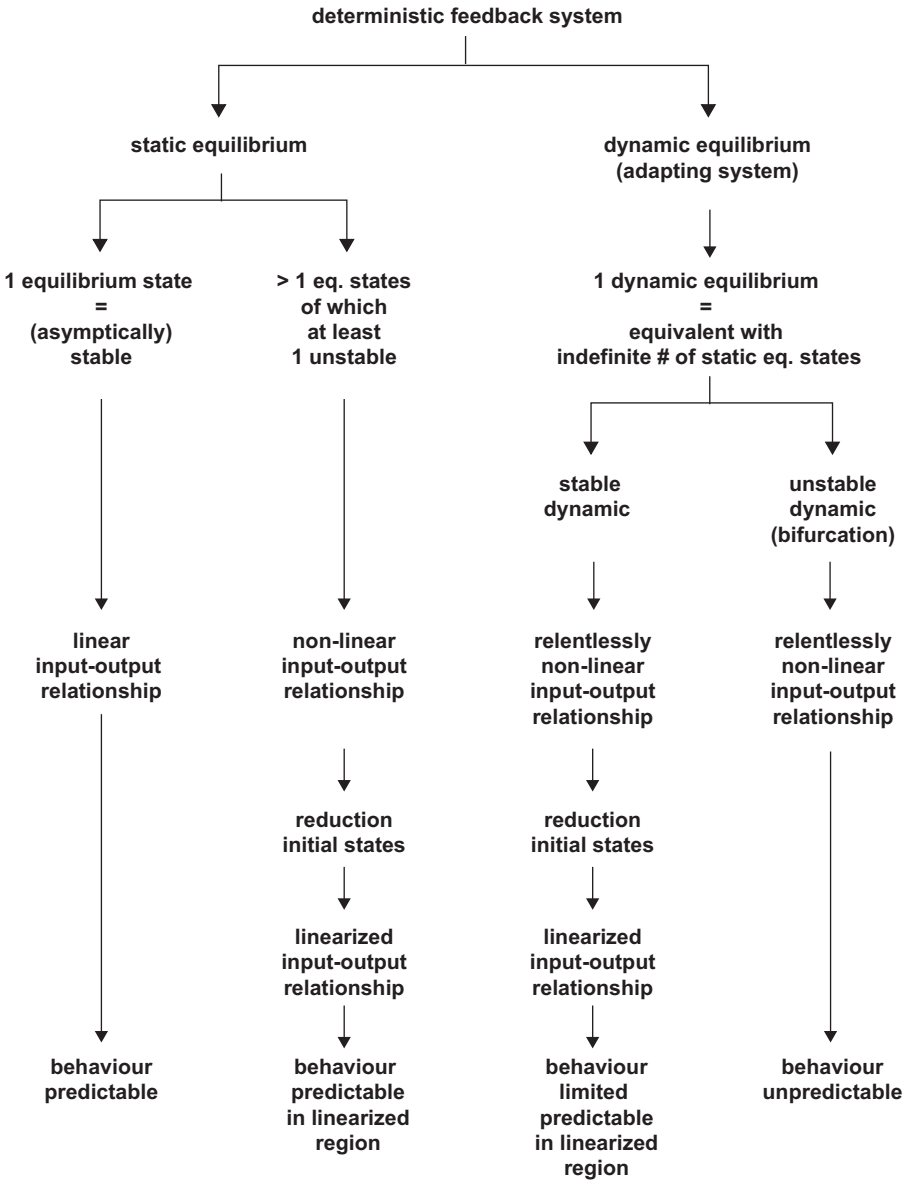


Figure 2.3 *To be able to predict behaviour the number of equilibria incorporated in the area of possible system states is important. For obtaining a define input-output relationship the number of equilibria has to be reduced to one stable equilibrium. When the system is adapting obtaining an definite input-output relationship is not possible.*

rium states the system may run through is reduced to one local equilibrium state. Depending on the stability of a local equilibrium state, the extent to which the region of initial system states has to be reduced varies. For instance, when the equilibrium state is unstable then the region of (initial) system states has to be reduced to exact one system state as for example by a pendulum in the upright position. By reducing the set of possible initial system states, the regularity in variety improves at the cost of its potential richness in variety of behaviour. By extending the set of possible system states the system may encounter, the richness in variety improves at the cost of regularity. The behaviour becomes more and more rich in variety and therefore more and more unpredictable.

A non-linear system that moves away from its equilibrium, moves from dynamic behaviour that is regular in variety towards dynamic behaviour that is rich in variety (Prigogine and Stengers 1984; Capra 1996). According to Prigogine humans are relentlessly non-linear systems that behave far from equilibrium. This generates their enormous richness in variety of behaviour. To be able to remedy persistent human error of a relentlessly non-linear system the number of possible system states have to be enlarged as much as possible so as to incorporate as many equilibrium states as possible.

2.7 Discussion: remedying persistent human errors

The essential requirement for a human behaviour model for remedying persistent human errors, is that the richness in variety of the behaviour generated by the behaviour model should correspond to the richness in variety of the human operator's actual behaviour. Thus to be able to remedy persistent human error, a human behaviour model is necessary that is general to all specific human behaviour. This means that this human behaviour model should be valid irrespective of the human's input states. For man-made systems it is fairly easy to define such a set of process equations because man-made systems are designed with an objective, such that they are almost linear (Hollnagel 2001; Van de Vegte 1990). Hence their dynamic behaviour can be represented by a linear n^{th}-order process equation with practically constant coefficients that satisfies for a large region of (initial) states. This is why technical sciences are so successful. For humans who are relentlessly non-linear systems embodying a vast number of variables, it is far less easy, if not impossible, to find such a set of global process equations based on input-output analysis. When such a set is found, it has only a limited validity, because the moment the input-output relationship has been established it most likely already has been changed.

Common denominator of current behaviour models: induction

All models discussed have in common that they have been induced through input-output analysis. To be able to model human behaviour through input-output analysis one has to weigh richness (=completeness) against regularity (=predictability). To deal with the richness in variety of human behaviour, the common approach has been to reduce the vast number of possible (decision-) situations to a specific set of situations that are considered to be essential for the dynamic behaviour of interest. Ashby calls models that are obtained this way "state determined" models (Ashby 1960). State determined models at best supply a set of local process equations. In other words, they will supply a set of process equations that are valid in a region around a local equilibrium state.

Current models satisfy until the occurrence of persistent human error. It is persistent human error that reveals the shortcomings of current behaviour models. The current approach is to repair the shortcomings by extending the possible input states the human may encounter, or to prevent as much as possible the undesired input states the human operator may encounter. Improving the human behaviour model by extending the input states a human possibly encounters is an infeasible task. Since real life situations do not neatly end at some point, the number of possible states is unlimited. Because humans behave in a continually changing environment of which it is impossible to know which change will occur at which time, and since their parametric values are continually changing, it is impossible to predict the exact values of the future dynamic behaviour of a human (Von Foerster 2003; Varela 1979).

Approach proposed: deduction

Due to the number of essential variables and the relentless non-linearity, the variety of human behaviour is too rich to be able to induce a set of global process equations through input-output analysis. However, the method of induction is not required to be able to obtain a behaviour model that gives starting points to remedy persistent human errors.

Obtaining a behaviour model other than by input-output analysis is possible but concerns the qualitative pattern of human behaviour rather than the precise quantitative values of its dynamic behaviour at a particular time in a particular situation. A qualitative analysis of humans is in terms of their equilibria. It has been observed that natural systems with feedback, like humans, develop a stable dynamic equilibrium, i.e. an Eigen behaviour (Prigogine 1989; Boyce and Diprima 1992; Von Foerster 2003). The qualitative analysis

of human behaviour then consists of identifying the human's Eigen behaviour. Then seemingly unpredictable specific human behaviour begins to reveal a general pattern. The richness in variety of this Eigen behaviour depends on the degree of non-linearity, and the number of essential variables.

One essential characteristic of Eigen behaviour is that it can be recognised without the necessity of input-output analysis. The major advantage is that the system does not have to be reduced to a manageable number of essential variables and/or initial conditions with which a set of process equations can be induced. This advantage is significant for high-dimensional systems that show relentlessly non-linear behaviour. The disadvantage of Eigen behaviour is that it is impossible to generate predictions, analytically and numerically. Eigen behaviour includes all specific dynamic behaviours implicitly, though they cannot be extracted explicitly unless the conditions and the past of the human operator are known exactly. Although Eigen behaviour can be induced from observed trajectories, the generative trajectory cannot be induced from observed Eigen behaviour (Von Foerster 2003). For example, the number "1" is the static Eigen behaviour (i.e. Eigen value) of infinitely many possible square root operations. The number "1" can be induced from any square root operation, yet one cannot induce from "1" the exact series of square root operations that generated it, unless one knows the whole past. Nevertheless this Eigen behaviour gives the opportunity to deduce possible future dynamic behaviour.

2.8 Conclusion: starting points for a new behaviour model

In this chapter current behaviour models have been reviewed, in the light of persistent human error. Human error has been defined as a deviation between the behaviour the observer expects based on the behaviour model applied and the behaviour actually shown by the operator. The starting point of control theory states that for effectively remedying errors a behaviour model is required. Based on this starting point a brief survey has been carried out into existing behaviour models and approaches that aim at remedying persistent human error. Although the current human behaviour models differ in realisation they all have a common denominator: they induce a local generic behaviour model based on a reduced set of initial states. Moreover, the goal(s) the human operator is supposed to achieve are considered to be determined and beyond discussion.

In conclusion, the essential process representing human dynamic behaviour must satisfy the following two requirements:

1. it must account for all possible initial states and possible perturbations that a human can encounter, this implies that the model should incorporate persistent human error as well as desired behaviour. In other words, the behaviour model to be developed may not contradict the existing behaviour models since these are special state determined cases of general behaviour
2. it must account for the characteristic dynamics of human behaviour, i.e. it must account for autonomy, creativity and development.

In the next chapter the shortcomings of the current behaviour models with regard to persistent human error will be resolved by assuming that an Eigen behaviour can be recognized. An Eigen behaviour is independent from initial conditions or changes in its conditions. After all, a human is a creative and autonomous system that by nature is error controlled and error controlled systems always develop Eigen behaviour.

REFERENCES

Ashby WR (1956) *An Introduction to Cybernetics*. London: Chapman & Hall.

Ashby WR (1960) *Design for a Brain, the Origin of Adaptive Behaviour*. London: Chapman & Hall.

Bainbridge L (1997) The Change in Concepts Needed to Account for Human Behavior in Complex Dynamic Tasks. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, **27:** 351-359.

Boyce WE, DiPrima RC (1992) *Elementary Differential Equations and Boundary Value Problems*. New York: John Wiley & Sons, Inc.

Capra F (1996) *The Web of Life, A New Scientific Understanding of Living Systems*. New York: Anchor Books.

Conant RC, Ashby WR (1970) Every Good Regulator of a System must be a model of that System. *International Journal of System Science*, **1:** 89-97.

Doherty ME (1993) A laboratory Scientist's View of Naturalistic Decision Making. In: Klein GA, Orasanu J, Calderwood R, and Zsambok CE, eds. *Decision Making in Actions: Models and Methods*. Norwood NJ: Ablex, 362-388.

Dreyfus HL (1997) Intuitive, Deliberative, and Calculative Models of Expert Performance. In: Zsambok CE and Klein GA, eds. *Naturalistic Decision Making*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 17-28.

Endsley MR (1995) Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, **37:** 32-64.

Endsley MR, Kaber DB (1999) Level of Automation Effects on Performance, Situation Awareness and Workload in a dynamic Control Task. *Ergonomics*, **42:** 462-496.

Flach JM (2004) The Human as a Critical Component in an Adaptive Meaning Processing System. In: *proceedings of IEEE International Conference on Systems, Man, and Cybernetics.* 2586-2591. The Hague, The Netherlands, TU Delft.

Heylighen F (1988) Formulating the problem of problem-formulation. In: Trappl R, ed. *Cybernetics and Systems*. Dordrecht: Kluwer Academic Publishers, 949-957.

Heylighen F. (1998) Problem Solving. At: *http://pespmc1.vub.ac.be/PROBSOLV.html*

Hollnagel E (1993) *Human Reliability Analysis: Context and Control*. London: Academic Press.

Hollnagel E (2001) Cognition as Control: A Pragmatic Approach to the Modelling of Joint Cognitive Systems. In: *proceedings of IEEE International Conference on Systems, Man, and Cybernetics.*

Hollnagel E, Woods DD, Leveson N (2006) *Resilience Engineering, Concepts and Precepts*. New York: Ashgate Publishers.

Hutchins E, Palen L (1995) Constructing Meaning from Space, Gesture, and Speech. In: Resnick LB and Saljo R, eds. *Discourse, tools, and reasoning: Essays on situated cognition*. Heidelberg: Springer-Verlag, 23-40.

Klein GA, Calderwood R (1991) Decision Models: Some Lessons From the Field. *IEEE Transactions on Systems, Man, and Cybernetics*, **21:** 1018-1026.

Kleinman DL, Baron S, Levison WH (1971) A Control Theoretic Approach to Man-Vehicle Systems Analysis. *IEEE Transactions on Automatic Control*, **AC:** 824-832.

Kroes PA (1996) *Ideaalbeelden van wetenschap: een inleiding tot de wetenschapsfilosofie*. Amsterdam: Boom.

Lipshitz R (1997) Naturalistic Decision Making Perspectives on Decision Errors. In: Zsambok CE and Klein GA, eds. *Naturalistic Decision Making*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 151-160.

Lorenz EN (1963) Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, **20:** 130-141.

McRuer DT, Jex HR (1967) A Review of Quasi-Linear Pilot Models. *IEEE Transactions on Human Factors in Electronics*, **HFE:** 231-249.

Newell A, Simon HA (1972) *Human Problem Solving*. Englewood Cliffs NJ: Prentice Hall.

Prigogine I (1989) The Philosophy of Instability. *Futures*, **21:** 396-400.

Prigogine I, Stengers I (1984) *Order out of Chaos*. New York: Bantam.

Rasmussen J (1983) Skills, Rules, and Knowledge: Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-13:** 257-266.

Rasmussen J (1986) *Information Processing and Human Machine Interaction; an Approach to Cognitive Engineering*. New York: Elsevier Science Publishers B.V.

Reason J (1990) *Human Error*. New York: Cambridge University Press.

Rosenblueth A, Wiener N, Bigelow J (1943) Behavior, Purpose, and Teleology. *Philosophical Science*, **10:** 18-24.

Stassen HG (1975) Mens-machine systemen: regel-, stuur- en bewakingstaken. In: Michon J.F., ed. *Handboek voor Psychonomie*. Dordrecht: Kluwer, 585-607.

Stassen HG, Johannsen G., Moray N (1990) Internal Representation, Internal Model, Human Performance Model and Mental Workload. *Automatica*, **26:** 811-820.

Van de Vegte J (1990) *Feedback Control Systems*. Englewood Cliffs NJ: Prentice Hall.

Varela FJ (1979) *Principles of Biological Autonomy*. North Holland/New York: Elsevier.

Von Foerster H (2003) *Understanding understanding*. New York: Springer.

Von Foerster H (1972) Responsibilities of Competence. *Journal of Cybernetics* **2(2):** pp. 1-6.

Wickens CD (1992) A Model of Human Information Processing. In: *Introduction to Engineering Psychology and Human Performance*. New York: HarperCollins Publ., 16-21.

Wiggins S (1990) *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. New York: Springer-Verlag.

Zsambok CE, Klein GA (1997) *Naturalistic Decision Making*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Chapter 3

MAINTAINING SENSING INFLUENCE: A MODEL THAT EXPLAINS HUMAN ERROR[1]

Humans are able to act creatively in situations that are not clearly defined and which have no ready solutions. This autonomous creativeness is the key characteristic that is general to all human behaviour. The current models describe human behaviour by listing specific characteristics of humans. Still, a human behaviour model does not exist that gives a satisfactory answer to what is the essential characteristic process that is generic to all specific human behaviour. This chapter discusses a new human behaviour model we have been developing based on Maturana and Varela's concept of autopoiesis (literally meaning self-creating), which can explain human error as well as human success based on the same essential process proper to all human behaviour. This behaviour model supplies us with the critical process Maintaining Sensing Influence (MSI) that is generic to all specific autonomous creative behaviour of humans including human error, which is viewed as a form of creative behaviour as well although an unexpected form.

1.This chapter has been submitted as: "*Maintaining Sensing Influence: an autopoietic behaviour model that explains human error,*" by M. van der Vlugt, P.A. Wieringa, and J. Dankelman

> *"The definition of a problem and the action taken to solve it largely depends on the view which the individuals or groups that discovered the problem have of the system to which it refers. A problem may thus find itself defined as a badly interpreted output, or as a faulty output of a faulty output device, or as a faulty output due to a malfunction in an otherwise faultless system, or as a correct but undesired output from a faultless and thus undesirable system."*
>
> Herbert Brün, 1971

3.1 INTRODUCTION

In high reliability domains human operators are essential for the less clearly defined situations where automation and expert systems have their drawbacks. Looking at behaviour of expert operators who deal with real life situations, researchers have found that they act autonomously and creatively in the situations that are not clearly defined or are new to them and for which there are no ready solutions (Winograd and Flores 1986, Dreyfus and Dreyfus 1986). When acting autonomously and creatively, human operators make sense of the many influences impinging on them drawing not only on acquired skills, rules and knowledge, but even more on experience and common sense (Dreyfus 1997, Varela 1991, Winograd and Flores 1986, Flach 2004, Weick 1995). On the one hand such creative autonomous behaviour is taken for granted when human operators succeed to ensure safe and efficient performance. It is so evident, as to be unnoticed. On the other hand, such creative autonomous behaviour still cannot be explained satisfactory when human operators err and jeopardise safe and efficient performance.

Creative autonomous behaviour is related to the richness in variety of human behaviour which makes it difficult to predict it exactly for all kind of situations. So far, the majority of human behaviour models is limiting this richness in variety of human behaviour by reducing the infinite number of possible situations to a defined operating domain, containing only those situations the human operator is most likely to encounter. These behaviour models are valid to the extent that their predicted behaviour resembles the real behaviour. Behaviour that is not predicted is erroneous human behaviour or human error.

The approach to understanding erroneous behaviour has been to progressively enlarge the defined operating domain so as to identify more influences and incorporate them in

the behaviour model. However, this way of progressively resolving erroneous behaviour has two drawbacks. The first is that it is impractical, if not impossible, to identify all influences that may affect the human operator. Real life situations do not end neatly at some point, there will always remain important influences not included in the model. The second drawback is that this approach models human behaviour by listing the characteristics of human behaviour specific for defined situations. However, it does not resolve what is the generic characteristic to all human behaviour, i.e. it does not give a satisfactory answer to what is the essential process that can generate all specific characteristics proper to human behaviour. The key question to be answered is: what is the essential process that makes behaviour human behaviour.

This chapter discusses a new human behaviour model we have been developing which can explain human error as well as human success based on the same essential process proper to all human behaviour. The basic concept used for this model is the concept of autopoiesis (which literally means “self-creating”) developed by Maturana and Varela (1980, 1992). This concept adheres to the idea that human behaviour is not determined by environmental influences but by their essential process: “Maintaining Organisation” (MO); external environmental influences only trigger human behaviour. However, according to Maturana and Varela the concept of autopoiesis is only valid for describing biological phenomena of human behaviour and not for describing social phenomena of human behaviour. Therefore, firstly it will be reasoned that the concept of autopoiesis is valid for describing human social behaviour as well. Next, the concept of autopoiesis is adapted to the human social behaviour by rephrasing the essential biological process of MO to the process of “Maintaining Sensing Influence” (MSI). Then the process of MSI is used to reconsider the current view on human error. Finally, the pros and cons of the MSI-model are discussed in respect to the current behaviour models.

3.2 Starting points for the behaviour model

Our aim is to develop a human behaviour model that could give starting points for remedying human error instantaneously irrespective of the specific situation. Therefore, the essential process has to be defined that is generic to all human autonomous creative behaviour. The autonomous creativity of humans that may result in desired as well as in erroneous behaviour is taken as the key to understanding human behaviour. After all, erroneous behaviour is autonomous creative behaviour as well, although unwanted by the observer.

As mentioned, the central objective of this chapter is to develop a model of behaviour, M(B). However, this model is developed by behaving, B(M). This means that this model must be developed so as to develop itself M(B(M)). Such a model of behaviour implies the one who models it. Therefore, we need a model of the modeller, i.e. a model of the observer who observes (Von Foerster 2003). This means that the observer is being included in a larger circular closure. Due to the inclusion of the observer, the model of behaviour is self-referent. Such a model differs from an observer excluded model, i.e. an objective model. An objective model requires that the observer shall not enter the model of his observations. However, if the observer is excluded from his model, there is nothing left, no observation, no model. As pointed out by several cyberneticians, like the neuro-biologists Maturana and Varela (1980), the physicist Von Foerster (2003), the anthropologist Bateson (1979) and the linguist Watzlawick (1967), human behaviour emerges as interaction between the human operator and the human observer (the community who observes). So the model representing the operator's behaviour depends on the observer. Or as Maturana and Varela put it: "every thing said is said by an observer" (1992). Our situation is the same, therefore the challenge is to develop a behaviour model that accounts for us, the observer community.

Characteristic for all biological systems is that they are feedback systems or circular closed systems. Through operational closure a dynamic stable behaviour occurs that is generic to all specific behaviour: Eigen behaviour. Eigen behaviour is common for all systems that are operational closed, i.e. that regulate themselves, that are autonomous. The major advantage of Eigen behaviours is that they can be recognised without the necessity of input/output analysis. It frees the observer from the necessity of identifying the initial states as they are automatically supplied by the end states. Eigen behaviour is characterised by its operational closure rather than determined by environmental influences.

By including the observer in the model of behaviour, a system arises that is doubly closed (Von Foerster 2003). By closing the chain the interpunction for reasoning can be arbitrarily chosen. In other words, for an effect in the present the observer can either find the cause in the past by cutting the circular chain at one spot, or the observer can find the purpose in the future by cutting the circular chain at the diametrically opposed spot. Circular causality closes the gap between cause and effect, between motive and purpose (Watzlawick 1967).

Characteristic for creative behaviour is renewal. Characteristic for autonomous creative behaviour is consolidation of known patterns of behaviour and inferring from these known patterns to renew patterns of behaviour.

Instead of listing all specific characteristic behaviour patterns by input/output analysis, we will make use of the characteristic generic to all operational closed systems: Eigen behaviour. Starting from human Eigen behaviour a generic process will be described that is able to generate all specific behaviour patterns proper to humans.

3.3 A SHORT INTRODUCTION TO AUTOPOIESIS

Since it is acknowledged that it is impossible to model all specific characteristics of human behaviour, an approach of abstraction instead of reduction is adopted, to develop a generic model of human operator behaviour. Maturana and Varela developed a particular elegant approach of abstraction, generally referred to as the concept of autopoiesis. They propose that the characteristic to all living systems is that they are continually self-creating, that is, they are continually "auto-poieting." This concept considers humans as autonomous creative systems which modify in interaction with their environment. Autopoiesis is a process of operational closure, a general dynamic pattern that can be recognised without the necessity of input/output analysis. We will give a short introduction which will serve as basis for a further elaboration of autopoiesis with respect to human social behaviour. For an extensive discussion of the concept of autopoiesis is referred to Maturana and Varela (1980, 1992) and Capra (1996).

THE GENERAL PATTERN OF LIVING SYSTEMS IS AUTOPOIESIS

The concept of autopoiesis abstracts the wide variety in the way living systems interact with their environment to a general pattern that is common to all living systems. This abstraction is realised by distinguishing clearly between the organisation and the structure of systems. The organisation of a system refers to its set (network) of relations which defines the system's identity, i.e. as a member of a specific class. The structure of a system refers to its physical components and their interrelations that actually constitute the system. It is the structure that makes the system's organisation concrete (see Figure 3.1) (Maturana & Varela 1992). Although living systems differ from each other in their structure, they are alike in their organisation.

The general pattern of a living system is its continuous process of maintaining its organisation, while its structure continually changes in recurrent interaction with its environment. This pattern is called autopoiesis; a living system is an autopoietic system. In contrast with physical systems, characteristic for an autopoietic system is that its organisation and structure are recursively dependent: the organization is essential for the structure and vice versa. According to Maturana and Varela, the key distinction between physical and living systems is the difference between relations among static components and relations among processes of production of components. When these processes of production stop, the entire organization stops. In other words, autopoietic systems must continually regenerate themselves to maintain organization. There is no separation between process and product, the being and doing of an autopoietic system are inextricably linked. The concept of autopoiesis is illustrated in example I with a unicellular, the most simple autopoietic system (Maturana& Varela 1992).

Example I

A bacterial cell is a concrete first-order autopoietic system, like every other unicellular. The structure of the bacterial cell consists of components (cell molecules) that continually interact with each other in a molecular network of ongoing chemical transformations (cell metabolism). The components in this network continually produce and transform one another. They are related in a network of ongoing interactions. The cell's molecular network produces its own components which in turn form the network that produces those components. Some of the components of the cell's network form the membrane, a limit to the network of transformations. However, the membrane is not solely a limit to the network, because it participates therein like other cellular components. The membrane both limits the network of transformation and is part of that network of transformations. The bacterial cell is characterized by an organizational closed production of a molecular structural network which defines its own boundary and in which it is impossible to separate product (the cell membrane) from process (cell metabolism).

In an autopoietic system all network components are produced by other components in the network, that is, the system is organisationally closed. Organisational closure implies self-organisation in the sense that the system's acting is not imposed by environmental influences, but is established by the system itself. Autopoietic systems are consequently characterised as autonomous systems (Capra 1996).

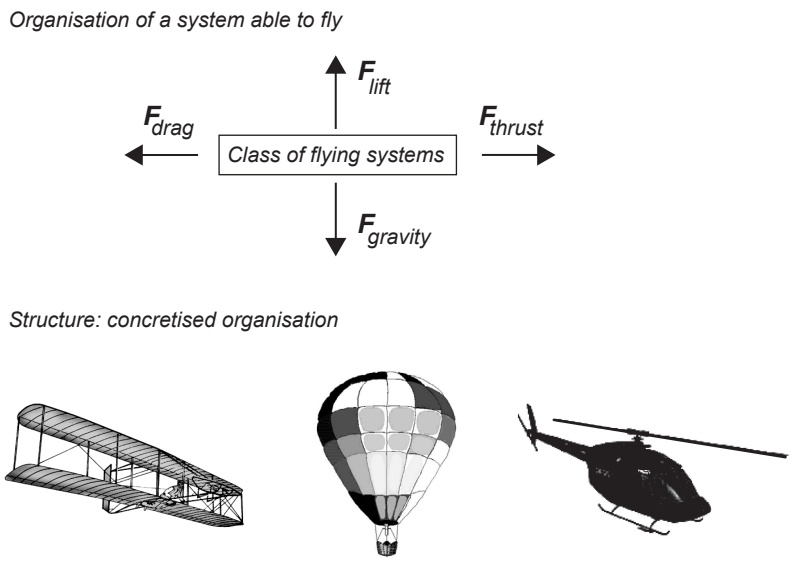


Figure 3.1 *The difference between organisation (top) and structure (bottom) of a "system that is able to fly." The organisation of a system is the set of relations that characterises the system as belonging to the class of systems that can fly. For a system to fly it should satisfy the dynamic relations between the drag, thrust, gravity and lift. The organisation does not identify the components of the system. The structure determines how the organisation of a "system able to fly" is actually constituted. The structure is constituted by the actual relationships among the physical components. The organisation can be constituted in many different structures, such as a Wright-glider, a balloon or a helicopter.*

STRUCTURAL CHANGE OF LIVING SYSTEMS

Essential for the continuing existence of an autopoietic system is its recurrent interacting with its environment. As it keeps interacting with its environment the autopoietic system maintains and regenerates itself; the autopoietic system is continually changing its structure. This structural change is a continuous process of adapting the system's network of self-producing components to environmental influences. The domain of structural changes that are available for a system at a certain moment is determined by its current structure. Which structural change the system's structure actually undergoes is subordinate to its maintenance of organisation. The system's structural changes represent its interactions with its environment. The domain of interactions defines the set of en-

vironmental influences a system can process. Only some of the environmental influences impinging on humans trigger structural changes. There will be influences having no effect whereas other influences are directly fatal for humans.

In the view of autopoiesis, structural change is determined by the system's current structure and only triggered by the environmental influences. Instructive interactions are therefore not possible since this assumes that structural change is determined by environmental influences, whereas in the view of autopoiesis structural change is determined by the structure. Although structural change is solely determined by the system it does not mean that the system is isolated from the environment. Structural change is what the environmental influences mean to the system. The only way to determine what environmental influences mean to the system is to observe its structural change. Note that with structural change the "influences from the environment" and the "influences on the environment" coincide.

Structural coupling of living systems

As long as the system continues to exist, its recurrent interacting leads to a historical development of mutual harmonized structural changes of both the system and its environment. The historical development of structural changes in a system is called ontogenesis. A system's ontogenesis at any moment is realised by its current structure. Mutual harmonized ontogeneses of the system and its environment is called structural coupling. In structural coupling the system and its environment mutually specify each other. Structural coupling implies that there is inevitably mutual influencing, but how this influencing takes place is determined by the structure of both the system and its environment (Capra 1996; Maturana & Varela 1992).

For an observer there are two distinctive cases of structural coupling: a system coupling with its environment in general, and a system coupling with another specific system. For the internal dynamics of a system, however, this observer distinction does not exist. In the process of structural coupling between two (or more) autopoietic systems each involved system is, with respect to the internal dynamics of the other(s), a source of triggers for structural changes and vice versa. Structural coupling between two first-order autopoietic systems, leading to a new operational closed process of which they are components, is called a second-order autopoietic system (see Figure 3.2). A first-order system's existence is subordinate to the maintenance of organisation of its entailing second-order system and it disintegrates when its structural coupling is broken. An example of

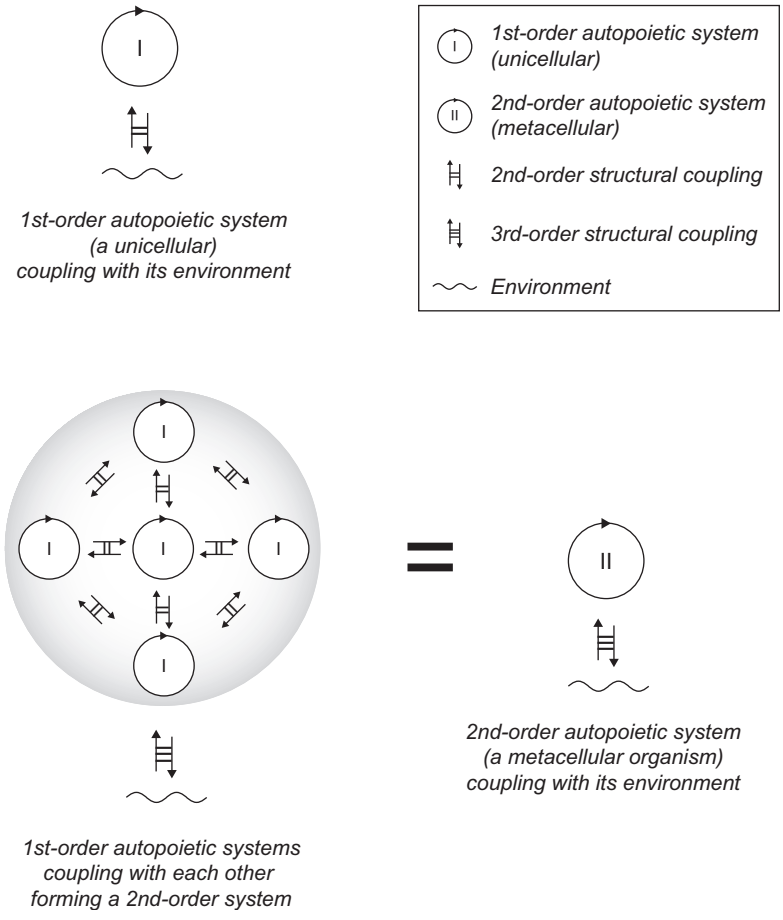


Figure 3.2 *Diagram illustrating the concept of autopoiesis for unicellular and metacellular organisms on organisational level. An autopoietic system creates itself by coupling structurally to its environment. The upper left corner illustrates a 1st-order autopoietic system coupling to its environment. The lower left corner illustrates a 2nd-order system consisting of five 1st-order autopoietic systems coupling with each other. 1st-order autopoietic systems which maintain their organisation by structural coupling among each other and at the same time establish a new autopoietic coherence constitute a 2nd-order autopoietic system or a metacellular system. Every living organism, except unicellulars, can be seen as a metacellular. Humans are thus metacellulars. In this classification of structural coupling, which differs on purpose from Maturana and Varela's classification (1980, 1992), there is no distinction made between a system coupling to its environment or to another system because for its internal dynamics there is no distinction either.*

a first-order autopoietic system, given in example I, is a unicellular. A second-order autopoietic system is for instance a human being.

A third-order autopoietic system is defined by Maturana and Varela (1980) as a system constituted by third-order structural couplings among second-order autopoietic systems (see Figure 3.3). Third-order structural coupling among second-order systems is called social coupling. A third-order system is therefore also called a social system. Social coupling is the mutual harmonized ontogeneses of the participating second-order systems. Consequently, a social system constituted by humans should include humans who are realised through their social coupling and whose individual ontogenesis occurs fundamentally as a component of their entailing social system. In this view a human's existence is subordinate to its entailing social system and disintegrates when its social coupling is broken.

The structural coupling of first-order systems is classified by us as a second-order coupling, independent whether the system is coupled structurally to its environment in general or to another first-order system. This classification differs on purpose (see discussion in Section 3.4) from the classification of Maturana and Varela (1980, 1992). They classify the structural coupling of a first-order system to its environment as a first-order coupling and the structural coupling of a first-order system to another first-order system as a second-order coupling.

3.4 Discussion on autopoiesis and human social behaviour

To be able to use the concept of autopoiesis for explaining human erroneous behaviour, it has to be applicable to humans coupling socially, that is, to human social behaviour. For being able to prove the applicability of this concept, it is necessary to know whether or not humans form a third-order autopoietic system, i.e. a social system, through social coupling. The question whether or not humans coupling socially constitute a new third-order autopoietic system has been discussed elaborately by different authors (Capra 1996). However, this question cannot be resolved unless we are able to rise above the third-order system of which we are essential components. As long as we are humans this is impossible. Maturana and Varela necessarily had to leave the question of social systems open, still they reason that social coupling is not essential for a human's existence.

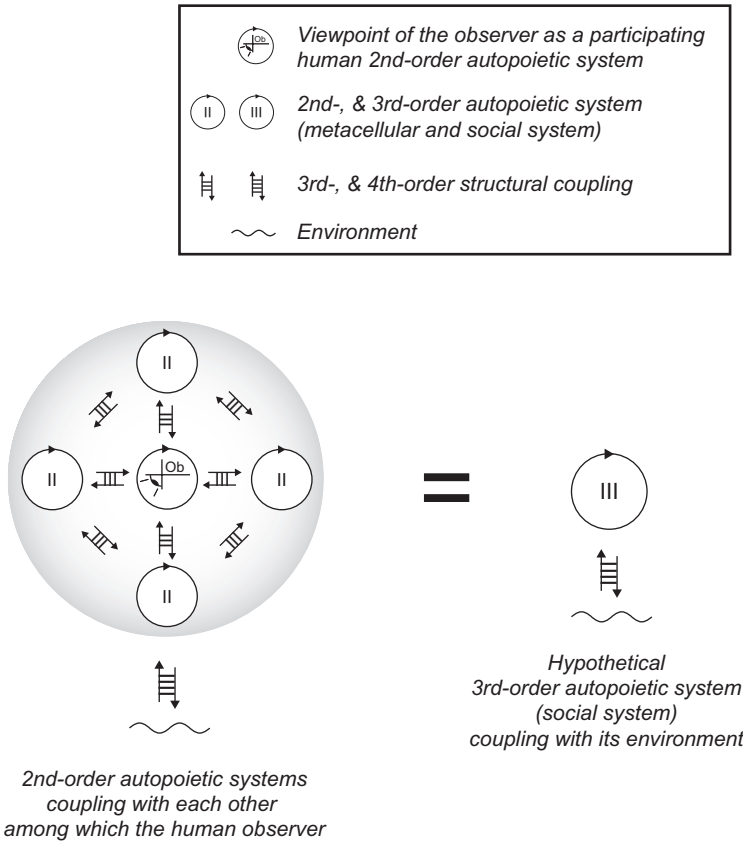


Figure 3.3 *Diagram illustrating the concept of autopoiesis on organisational level for a social system, i.e. a third-order autopoietic system. A social system is defined by Maturana and Varela (1980, 1992) as a system that is constituted by social coupling among second-order autopoietic systems (e.g. humans). According to this definition humans who participate in a social system realise themselves through their social coupling. In this perspective, the existence of humans depends on social coupling. Consequently, human social behaviour is determined by autopoiesis as well as biological behaviour, since for the internal dynamics there is no difference between coupling with the environment or coupling with another human. Note that we (observers) can only view the social system from within, just as we can only view the edges of the universe from within, therefore we can only speak of a hypothetical third-order autopoietic system. We can reason there could be a third order system, we cannot prove it (see also the discussion in Section 3.4).*

In this section we evaluate their reasoning before we will reason that social coupling is essential for a human to continue to exist.

Human's existence is independent of resigning from social institutions

Maturana and Varela define a human social system as a system constituted by third-order couplings among humans (see Figure 3.3). Consequently, a human social system should include humans who are realised by their social coupling and whose individual ontogenesis occurs fundamentally as a component of their entailing social system. For a social system to be an autopoietic third-order system it should satisfy the following two conditions (Varela 1979):

1. a social system should be materially self-limiting (like the self-limiting process of a bacterial cell (see Section 3.3))
2. a human as component should disintegrate when decoupled (resigned) from its social system

Maturana and Varela argue, based on the second condition, that social coupling among humans cannot be of the same order as structural coupling among first-order systems, because humans continue to exist when they resign from a social institution, whereas they should disintegrate when social coupling is essential to their existence.

In this reasoning Maturana and Varela categorise social institutions, enterprises and clubs as social systems. However, these systems cannot be a social system conform their first condition of material self-limitation. Consider a component in a hypothetical third-order system that is coupling socially. Then this component is not able to discern the boundary of the third-order autopoietic system because it can only view this third-order system from within (Varela 1979). This means that a social institution, in which the participating humans are considered as autonomous components, cannot satisfy the criterions for a third-order autopoietic system. Consequently, resigning from a social institution is not the same as resigning from a social system.

Social coupling is essential for a human's existence

The hypothesis that social coupling is essential for a human's existence can be found by turning around the reasoning of Maturana en Varela: because a human does not dis-

integrate when he has resigned from a social institution, this implies that its social coupling is not broken. This reasoning is consistent with our argument that a social institution cannot be a third-order autopoietic system. As described in Section 3.3, a living system's existence depends on the way the system couples to its environment. Consequently, human's existence emerges from social coupling to its environment. For a living system to be identified as a human system, another human is necessary. In other words, a human exists only in relation to another human. Or as Von Foerster (2003) has put it: "identity equals the coupling between other humans and I." In which identity is equivalent with existence. In this view therefore, a human's existence emerges from social coupling among humans. As long as a human continues to exist its social coupling continues. From our viewpoint social coupling is inseparably connected with the existence of the system "human." Social coupling is only a specialised form of structural coupling that enables the human to enlarge his domain of interactions.

In accordance with the concept of autopoiesis, the internal dynamics of a human do not distinguish between structural changes triggered by social coupling with other humans or structural changes triggered by structural coupling with the environment. Consequently, we may conclude that the concept of autopoiesis applies for social coupling as well. In social coupling the participating humans are mutually involved in maintaining their respective organisation. And therefore, we may conclude that social coupling is essential for human's existence. Both, human biological behaviour as well as social behaviour are determined by autopoiesis. All human coordinated actions, which are in fact structural changes, are subordinate to their Maintenance of Organisation (MO).

3.5 Human social behaviour revisited with autopoiesis

For explaining human operator behaviour in all its varieties, we argued that we needed a human behaviour model which generates two main characteristics of human operator behaviour, i.e. autonomy and creativity, and which includes both observer and operator. In section 3.4, we argued that the concept of autopoiesis meets both conditions.

An autopoietic system is literally a self-creating system which is characterised in that it maintains organisation (MO) while it continually undergoes structural changes in interaction with its environment. The condition of autonomy is captured by this self-organisation of its structure. It is the structure of a living system which determines the system's structural change, that is, its acting and not the environmental influences. Its structural

changes are identical to the environmental influences that have had an impact on the system. The creativity is captured by the system's continuous structural changes. At any point the structure of a living system is a momentary state of its historical development of structural changes (ontogenesis). The process of structural coupling makes a self-organising system a self-creating system, that is, a living system. It is therefore a developing or learning system. For instance, a dog that is kicked will act to the kick with structural changes conditioned by its momentary structure but subordinate to its maintenance of organisation (Bateson 1979). The dog may respond with barking or biting, or it may hide behind a chair. This response (i.e. structural changes) cannot be undone because it will be incorporated in its structure, which will in turn affect the dog's future responses.

According to the common view, operator behaviour is acting that belongs to the operator involved, in other words, the operator shows behaviour irrespective of the observer who perceives independently the operator's behaviour. The common view defines behaviour as the actions of an operator in response to environmental influences. This view on behaviour denies the mutual influencing among observer and operator established in their social coupling. According to the autopoietic view, operator behaviour is what operator acting means to the observer, i.e. operator behaviour is the structural change in the observer. Not all operator's acts trigger structural change in the observer. What the operator's acting means to the observer is determined by the observer's structure which determines its possible domain of structural changes and which is subordinate to its maintenance of organisation (MO). In the view of autopoiesis, social or communicative behaviour is co-ontogenetic acting which has been established in social coupling among observer and operator.

The concept of autopoiesis together with our reasoning that social coupling is essential to human existence provides a theoretical framework for human social behaviour. Moreover, with our extension of the concept of autopoiesis with regard to social behaviour we have found the critical process that can explain human social behaviour, namely Maintaining Organisation (MO). Note that the critical process MO is a theoretical process that reveals the connecting pattern between all varieties in human coordinated actions. MO itself cannot be observed or measured directly, it is an abstract theoretical process which can be recognised through stable dynamic behaviour of humans, i.e. through Eigen behaviour.

3.6 Human behaviour rephrased as Maintaining Sensing Influence (MSI) conditioned by social representing

In this section we will elaborate on human acting in social coupling, i.e. human social behaviour. Therefore, as a human specific case of social coupling the term social representing is introduced and explained. And as a human specific case of structural change, the terms Sensing Influence (SI) and Noting Influence (NI) are introduced and described.

To be able to illustrate the process of social representing as simply as possible the focus of this section will be on social coupling in the primary unit of analysis of a social system: two humans coupling with each other, the observer and the operator. The focus taken is illustrated in Figure 3.4. This primary unit of analysis of social representing can be extended to one human coupling to many other humans or many humans to many humans. The principle remains the same although the complexity increases.

Human social coupling: social representing

Social coupling is common to all second-order systems that couple structurally to other second-order systems. In section 3.3 we have defined social coupling as: the mutual harmonised historical development of structural changes of the involved systems. With it we defined structural changes as: the meaning to the system of environmental influences. When substituting the definition of structural changes in the definition of social coupling the following definition of social coupling emerges: Social coupling is the mutual harmonised historical development of what environmental (and each other's) influences mean to the involved systems.

To distinguish social coupling among humans from social coupling among second-order systems in general, we will adopt the term social representing as introduced by Zeegers (1988) and Mattheeuws (1990). They have defined social representing as: "jointly developing of evident meanings." Social representing is a recursive process; it creates meaning and has been created by meaning. Social representing is therefore ontological determining (Mattheeuws, 1990), i.e. the basis of a human's existence. Social representing is an ongoing human process of mutually developing meaning and since this is an ongoing process of structural change it is impossible to escape from. Whatever someone's actions are, when sensed by an observer, they mean something to the observer. Watzlawick (1967) wrote: "one cannot not behave"; in our words: "one cannot

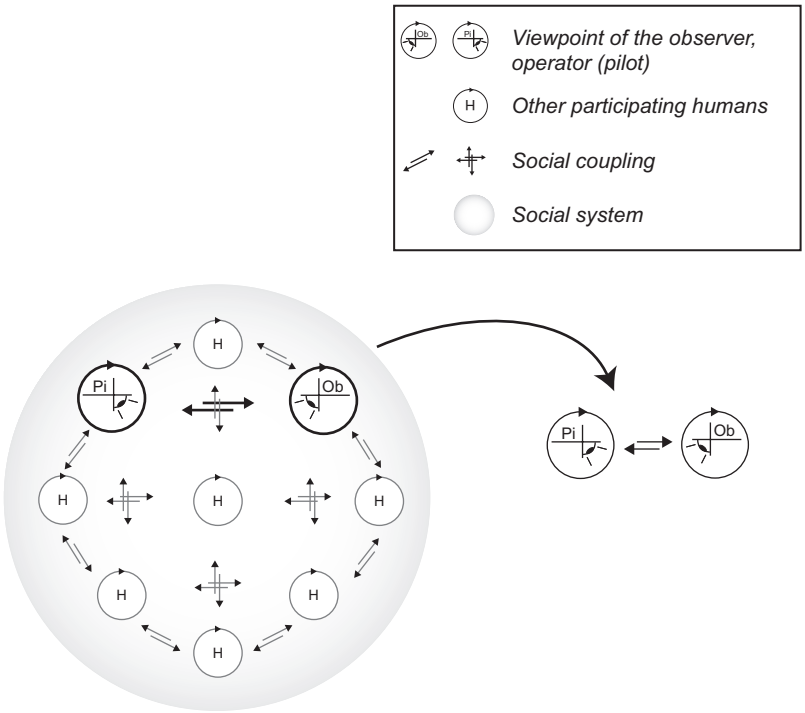


Figure 3.4 *Focus of Section 3.6. To be able to illustrate the process of social representing as simply as possible the focus will be on social coupling in the primary unit of analysis in a social system: two humans coupling with each other, the operator and the observer. Note that we do not reduce the system to the observer and the operator, because we do not renounce the social coupling of the two humans with others, the social coupling with others remains present due to the structure of both the observer and operator, which represents their respective historical development.*

not represent socially." Social representing differs from social coupling in that it incorporates a semantic form of social coupling. This semantic form appears when coordinated actions lead to coordinated actions that refer to themselves. For example, the term "chess playing" is a semantic coordinated action that refers to the set of coordinated actions that constitute chess playing.

From the ongoing process of social representing, social representations emerge. They are the constituted evident meanings about established appropriate patterns of actions.

Due to the recursiveness of this process of social representing, social representations are continually subject to changes. Social representations have partly been codified explicitly into rules of appropriate behaviour, like for example the rules of the criminal code, traffic rules and the rules of law. Although these rules are based on social representations they can serve only as basic directive for what someone's acting means to its environment. Because social representations are forever on the go, someone's acting is, in general, interpreted according to the spirit instead of to the letter of the agreed general codes. In court for example, the defendant's acts for which he stands trial, is mainly interpreted by jurisdiction, which is a recursive process of interpreting humans' doings and goings.

With social representations, expectations about appropriate coordinated actions emerge, i.e. about the way someone is supposed to behave. Expectations emerge based on the recurrent interactions that have proved themselves to be successful in social representing.

Structural change: Sensing Influence and Noting Influence

In the process of social representing observer and operator continually influence each other, by which both change structurally. Seen from the viewpoint of the observer, three forms of structural change can be distinguished (see Figure 3.5 in which the viewpoint of the operator is illustrated as well):

- **Influence**:= the observer's actions triggering structural changes in the operator (meaning *to* the operator *of* the observer's actions).
- **Sensing Influence (SI)**:= structural changes in the observer triggered by the operator's actions (meaning *to* the observer of the operator's actions).
- **Noting Influence (NI)**:= Reflecting of the observer on his SI, that is, the meaning *to* the observer when he reflects on: the effect he expected of his influence and the effect of the operator's actions he actually senses.

Influence is a hypothetical constructed effect which is only discernable when it is sensed by an observer; influence implies sensing influence. Influence and sensing influence are inextricably linked. In other words, influence cannot exist without sensing influence. Influence and sensing influence coincide in structural change (similar to Goudsmit 1998).

Sensing Influence (SI) is structural change in a human triggered by influences from the environment. SI is what environmental influences mean to a human. SI is not imposed

by the environment. All living systems continually sense influences from their environment. Characteristic for humans is that they can put their SI into words. Although semantic expression is a human specific form of SI, it is comparable with other forms of SI of living systems. SI is subordinate to Maintaining Organisation (MO). As long as the human continues to exist, SI remains present, that is the human remains sensitive to environmental influences.

Noting Influence (NI) is a reflexive form of SI. With Noting Influence the idea of our "self" emerges. When humans "note influence" they consider what (returned influence) they actually sense with regard to what they expected of their own influence. Reflection is Sensing Influence of Sensing Influence, it is an act of turning back upon ourselves. When the observer has NI: he reflects on his own influence. When the observer does not consider his own influence with regard to the returned influence he senses, there is no reflection anymore: the observer has no NI. In the case where the observer has no NI, he only "senses influence", that is, he only considers the influences from the environment as if his influence does not contribute any more to the influence returned. NI can be lost, whereas SI cannot be lost as long as the living system continues to exist. When NI has been lost the observer reverts to SI due to the concept of maintaining organisation.

Human behaviour explained with the critical process MSI

In Section 3.5 we defined human (operator) behaviour as the structural change (i.e. the meaning of environmental influences) in the observer in response to the operator's acting. Human operator behaviour therefore involves both operator and observer. Although the structural change in the observer is triggered by the operator's actions it is determined by the observer's structure. The observer's structure is subordinate to his MO and has emerged in social coupling.

In this section we have introduced Sensing and Noting Influence to distinguish human structural change from structural change in general and social representing to distinguish human social coupling from social coupling in general. In social representing, observer and operator continually represent their structural changes, being the meaning to them of environmental influences. Social representing is a human specific case of behaving, because behaving has been redefined as: what acting means to the observer.

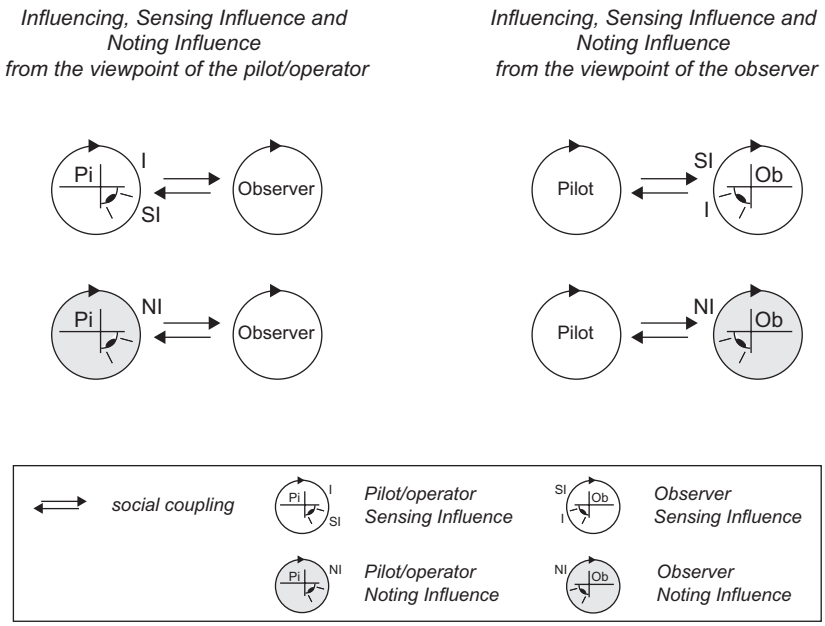


Figure 3.5 *Influencing (I), Sensing Influence (SI) and Noting Influence (NI) illustrated from the viewpoint of the operator (left) as well as from the viewpoint of the observer (right). In the concept of autopoiesis both operator and observer continually participate in a process of mutual Influencing and Sensing Influence. Where the observer, for instance (right), reflects on the effect he actually senses of his influence and the effect he expected based on his social representations, we speak of Noting Influence. Noting Influence is a reflexive form of Sensing Influence in which our "self" emerges. Note that in this diagram the common environmental interactions are left aside, although they remain present. This is done because we focus on humans coupling socially.*

For our human behaviour model we will use **Maintaining Sensing Influence (MSI)** as the critical process for continuing existence of a human. In other words, by Maintaining Sensing Influence (MSI), under all circumstances, a human system maintains its organisation. MSI is not a matter of choice but a matter of necessity with regard to MO. How SI is maintained is determined by the human's structure. The human's structure determines which influences are sensed and how they are sensed. A human maintaining SI is continuously changing structure, triggered by environmental influences. MSI is a case of MO.

Instead of MO conditioned by social coupling, we will continue using MSI conditioned by social representing as the critical process to explain human behaviour.

3.7 Erroneous behaviour explained with MSI

According to the concept of autopoiesis inconsistencies in the operator's behaviour arise from a change in the operator's context generating his acting while the context of the observer has not changed accordingly. Inconsistent behaviour reveals that there is a discrepancy in context between observer and operator.

In this section we will discuss, as observer, inconsistent behaviour of the operator in respect to maintaining SI, the critical process to which humans act. Therefore we start by putting ourselves in the operator's place. In other words, we imagine that we can describe what the operator does, although there is nobody to observe. Then we discuss the effects the operator's actions may have on the observer who tries to explain what happens to the operator via the observable actions of the operator. Then we will distinguish three classes of contexts with which we will define erroneous behaviour in respect to inconsistent behaviour.

Context change is MSI

The common meaning of context is: that what assigns meaning. With the definition of social representing from Section 3.6, context can be reformulated as: the historical development of meaning to the human (operator) of environmental influences. Although context is a personal matter, it does not exist independently since it emerges in social representing. A discrepancy in context means a notable difference in development of what environmental influences mean to the operator and the observer respectively. Since meaning of environmental influences to the human is an equivalent of either SI or NI, context is a historical development of a human's SI. In the view of autopoiesis, SI is a necessity for the existence of a human.

MSI can be realised through either NI or SI. In case the operator has NI, it implies that for every situation the operator considers as well all possible effects as the effects he actually senses with respect to his actions. The operator reflects on his SI, no matter what the returned effects are. As long as the operator maintains NI, he is optimally creative and flexible in adapting his actions to the returned effects. In other words, although

the operator may sense unexpected effects from his actions in the first instance, he does know what to do to achieve expected effects. In adapting to unexpected responses, the operator creatively adapts to what the environmental influences mean to him by considering what his actions may have meant to the environment. In other words, he creatively changes his context while maintaining NI.

Where the operator has lost NI, this implies he reverted to SI. In the case of lost NI, the operator does not reflect anymore on his SI. Reverting to SI can occur in two ways:

1. There is a match between the effect the operator expected and the effect he actually senses, though it has become so evident it is not noted anymore by the operator. In other words, the operator senses expected effects although he does not note them anymore. This occurs when the operator becomes accustomed to the ways of appropriate acting, i.e. habituation and experience.
2. There is a mismatch between the effect the operator expected and the effect he actually senses, and the operator is not able to adapt creatively to the new situation. In other words, the operator senses unexpected effects to his actions but he does not know what to do to achieve an expected effect from his environment. The operator can sense effects as unexpected in two ways:
    a. The operator senses “no effect” on his acting at all, whereas he expects a positive effect. Undoubtedly, there is a positive effect to the operator’s acting, though this effect is so evident that the operator does not sense it anymore. In this case, with lost NI, for MSI there is nothing left than to provoke a negatively sensed effect. The operator gets into situation 2b.
    b. The operator senses a negative effect on his actions.

When the operator senses unexpected effects and he does not know what to do to achieve expected effects in response to his acting, the operator senses that the others do not note that his acting is essential to his MSI. Consequently, the operator’s existence is threatened. Threatened existence is averted through changing context such that the operator senses his acting again as appropriate for meeting his expectations irrespective of the response of the others. In changing the context in this manner, the operator changes the meaning of his acting, rather than that his behaviour becomes appropriate.

NI is desired for an operator to be optimally creative and adaptive in responding to effects that are in the first instance unexpected. As the operator reverts to SI, the operator is

limited in his creativity compared to his creativity when he has NI, that is, his domain of interactions is limited to his intuitive creativity.

Observer context and (in)consistent operator behaviour

Inconsistent operator behaviour means that the operator and the observer no longer share their context. This implies that the observer's context has not changed in congruence with the context of the operator. Inconsistent operator behaviour is therefore always relative to the context of the observer.

Ideally, when the observer maintains NI, inconsistent operator behaviour does not occur. After all, an observer who continually has NI is able to consider all possible effects on the operator of his (and environmental) influences and the actual effects the operator returns on his influences, irrespective of the operator's context and irrespective of whether or not the operator has NI or SI. With NI the observer is optimally creative and flexible in adapting his expectations with respect to the operator's responses, that is, the observer is creative and flexible in adapting his context to the context of the operator.

When the observer has lost NI, it does not necessarily mean that the operator shows inconsistent behaviour. When the observer has lost NI but the operator still shows consistent behaviour in the view of the observer, it either implies that the context of the operator remained congruent with the context of the observer or that the discrepancy in context is not yet sensed by the observer. In the first case the observer and operator both lost NI and changed context congruently, i.e. there is relatively no discrepancy. It is also possible that the observer lost NI but the operator still maintains NI and therefore is able to adapt to the observer's context (see Figure 3.6).

When the observer has lost NI and the operator shows inconsistent behaviour in the observer's view, it means that the operator has lost NI as well and they both changed context differently. They both maintain SI in a different context. See Figure 3.6, the worst case scenario.

The evident, situational and reference context

The following three contexts can be defined to distinguish between consistent and inconsistent behaviour:

1. the evident context,

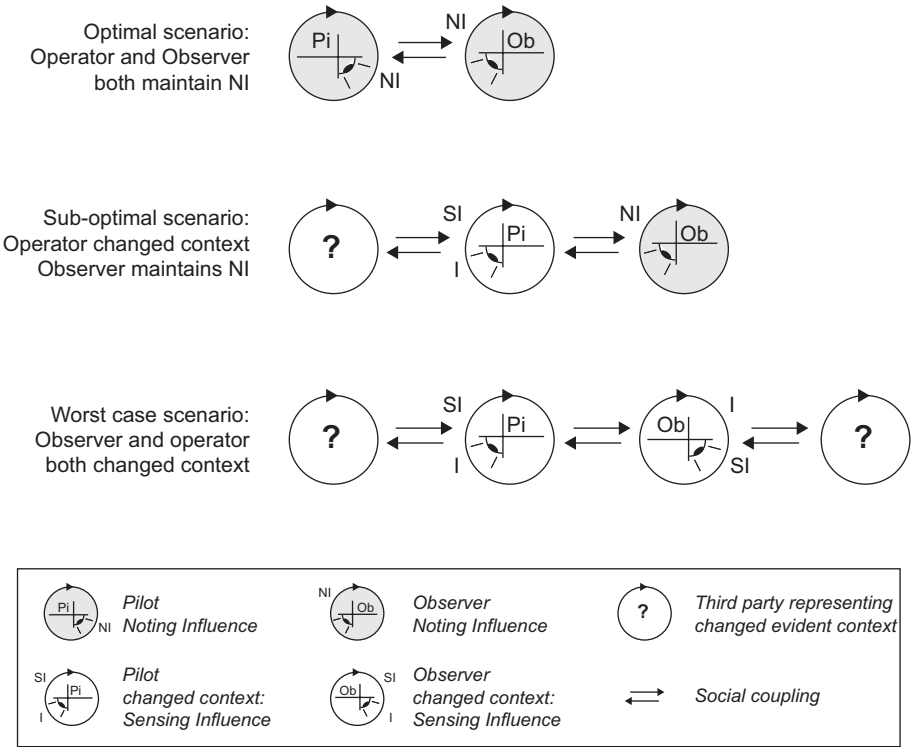


Figure 3.6 *Erroneous, creative and expected behaviour of the operator from the perspective of the observer. When the operator's acting and observer's acting are tuned to one another their mutually coordinated actions match their expectations. Then they reflect on the effect they expect and the effect they get, they reflect on their Sensing Influence. In other words, they "note the effect of their influence." This is the optimal scenario, as long as both maintain NI, erroneous behaviour does not occur. The sub-optimal scenario is the case when the operator loses his NI and changes context with which he maintains SI and the observer maintains his NI. In respect to the organisational expectations (the reference context) and consequently the observer, the operator acts erroneously. The worst case scenario occurs when both the observer and the operator have lost NI. Two things can happen: the observer and operator change context synchronously: the observer does not speak of erroneous behaviour since he errs as well, or the operator and the observer change context asynchronously: the observer speaks of erroneous behaviour.*

2. the situational context,
3. the reference context.

Evident context represents the tacit habitual based coordinated actions of appropriate behaviour, the so-called self-evidences. Self-evidences are the tacit unreflected habits that have come about in social representing, i.e. the unreflected social representations. They consist of all those mutually coordinated actions that in the historical development of meaning have become obvious, regular and acceptable, to a human. Self-evidences are indisputable for the humans involved and they are a necessity for acting. Self-evidences roll along without NI. In the evident context new patterns of acting are harmonised and developed without NI. In general, self-evidences are unnoticed for the humans involved, they become only notable for an observer when the operator changes his self-evidences. Asking for a reason for self-evidences will provoke incomprehension (I act like this because everybody acts like this). It is, for instance, a self-evidence that this PhD-thesis is written from left to right and from top to bottom. One does not ask for the reason. It is therefore considered to be unnecessary to instruct how to read it, since every body knows that automatically. When we do want to include an instruction, the intangibility of self-evidences appears when we realise that to read this instruction at the beginning, we need an instruction that reads how the instruction must be read, which needs an instruction... ad infinitum. The origin of self-evidences is hidden due to the recursiveness of the process of social representing, and therefore cannot be revealed by back-tracing.

Situational context represents all those situations for which there exist no established rules of appropriate behaviour yet. In the situational context new patterns of appropriate behaviour still have to be discovered. The situational context distinguishes itself from the evident context in that it is new for the human(s) involved. An extreme example of a situational context is the first time the Americans landed on the moon. A situational context has no explicitly defined stopping rule. In the situational context NI is more difficult to maintain, since it is not certain for an operator what to expect anymore. The operator's NI will change depending on the adjustment of expectations with regard to the situation. A situational context is a momentary context, it does not continue to exist, the discovered pattern of appropriate action will drift to an evident context or will be formulated explicitly as a rule of appropriate behaviour within the reference context.

Reference context represents the explicitly specified rules which constitute the reference in anticipated situations. The reference context is a formalised context about ap-

propriate behaviour which is accepted by the participating humans. It can be laid down based on self-evidences or based on new discovered rules of appropriate behaviour from a situational context. The reference context is a typical context with respect to the evident and situational context. It serves to pass on appropriate patterns of acting for daily situations, without being necessary to discover them by experience as in the evident or situational context. In aviation, for instance, the reference context consists of the standard operating procedures and federal regulations. The participants, i.e. the pilots and air traffic controllers, have accepted these rules as guidance for their behaviour. In daily life the reference context consists of for example the generally agreed codes like the criminal code and traffic rules. The reference context by definition equals the situation for which it is defined. It is a specific extraction of the social representations applicable for a defined situation.

In the reference context NI is high, because acting according to the specified rules results in expected effects by definition. After all, social rules refer to established agreed coordinated actions and to satisfy social rules implies that the actions expected and the actions sensed should always match. However, explicit defined social rules are inferred from social representations. This has two consequences: firstly, the specified rules are not applicable for all situations. Specified rules are established on situations from the past. They may fall short for new situations. Secondly, there are behaviour patterns that are not translated into explicit social rules. Hence, not all appropriate behaviour patterns are specified explicitly in rules.

Erroneous behaviour reveals a change to an evident context

Helmreich (2001) defined erroneous behaviour as an (in-)action that leads to a deviation from organisational expectations. Other currently used definitions of human error resemble this definition (Gerdes 1997; Kohn *et al.* 2000; Lipshitz 1997; Woods *et al.* 1994). Taking the term organisational expectations as an equivalent of the reference context, this implies that in case of:

- creative though erroneous behaviour, the operator changed from a reference context (organisational expectations) to an evident context to maintain SI, while the observers did not change their context synchronously.
- creative though desired behaviour, the operator changed to a situational context in respect to the reference context to maintain SI.

Erroneous behaviour is a subset of the set of inconsistent behaviour. When the operator shows erroneous behaviour this implies that he does not know what to do to satisfy the organisational expectations. According to the model of Maintaining Sensing Influence, not being able to satisfy the organisational expectations threatens the operator's existence. The operator averts this by changing to an evident context with which he senses his acting again as appropriate irrespective of responses from the observer. That is, irrespective of how the observer judges the operator's behaviour. Only when an operator shows erroneous behaviour we, observers, may conclude that he lost NI and changed to an evident context to maintain SI.

Note that we may not conclude that an operator who lost NI and changed to an evident context (by definition) consequently shows erroneous behaviour in the view of the observer. It is possible that the change to an evident context is not yet visible for the observer because the operator still acts (although without NI) according to the expectations of the observer. And it is possible that the observer and the operator both synchronously have changed to an evident context, in which case there is no discrepancy in context between the observer and operator.

3.8 Discussion

To be able to explain human operator behaviour in all its varieties, irrespective of human error or human success, the concept of autopoiesis has been introduced and adapted. This concept of autopoiesis synthesizes the boundless variety of specific human behaviour to a generic process characteristic for human behaviour. The concept of autopoiesis sets out from the assumption that in every stable dynamic feedback system Eigen behaviour arises. For living systems and therefore for humans this Eigen behaviour is autopoiesis or self-creating. From autopoiesis a critical circular closed process has been abstracted: Maintaining Sensing Influence (MSI). The MSI-model assumes that human behaviour is determined by Maintaining Sensing Influence. MSI is the process that is generic to all human behaviour. The way MSI is realised is a personal matter and can be done in many different ways.

Non-objectivity or the involvement of the observer

According to Koningsveld (1976) humans neither perceive independently an existing objective reality (objectivity), nor imagine independently a subjective reality (subjectiv-

ity); because reality is empirically established *in social representing by the participating humans*. The same reasoning accounts for the observer observing human operator behaviour. Operator behaviour is not an objective phenomenon that emerges from the operator and is perceived independently by the observer. Neither is operator behaviour imagined independently by the observer. Operator and observer do not behave independently from each other, they mutually specify each other's behaviour.

In the MSI-model this mutual involvement of the observer and operator is accounted for via the recursive process of social representing, the human case of behaving.The MSI-model reformulates the way human operator behaviour can be viewed. In the MSI-model operator behaviour is: how the observer senses the operator's influences. Behaviour is what the operator's actions mean to the observer. Operator behaviour does not belong to the operator but it belongs to the observer. Through this reformulation of operator behaviour we, as observer, are included in our behaviour model. The process of MSI applies for both the operator and the observer. Via behaving MSI is realised and vice versa. There is a double closure. The way an observer senses the operator's influences is subordinate to the observer's MSI. This makes is possible to explain human operator error and human operator success with the same process MSI. Human operator error is a discrepancy between what the operator's acts mean to the observer and what the operator's acts mean to the operator. In other words, there is a discrepancy between the context of the observer and the context of the operator. According to the model of Maintaining Sensing Influence, the following statement applies: what we observers say (about the operator), is what we observers are.

With the MSI-model we, observers, become participants in the process we examine. As participants in this self-referent process we cannot rise above this process and therefore cannot prove that we are subordinate to the critical process MSI. This does not imply that the MSI-model is not acceptable for explaining human behaviour. It only implies that the acceptability of the MSI-model depends on its applicability.

No Predictions

We modelled human behaviour by starting from the generic Eigen behaviour and abstracting its description to the operational closed process of Maintaining Sensing Influence. By this way of modelling we have abandoned the idea that human behaviour can be described as if the operator pursues a defined external objective, as the current behaviour models imply. This does not imply that the MSI-model is incompatible with current

behaviour models. On the contrary, it is supported by them since they are specific cases of the MSI-model. When the observer's NI and the operator's NI reign, and their reference contexts are synchronous, the observer can describe the operator's behaviour by a set of causal relationships between his objectives and actions. In other words, the operator's behaviour can be described as if it is determined by a set of external objectives. Rule-based behaviour and skill-based behaviour are examples of such operator acting. It appears to the observer as if the operator's behaviour is predictable. Still, this does not imply that the operator's behaviour is determined by these external objectives. The regularity in his behaviour has emerged in social representing, that is, from a mutually harmonised history of coordinated actions.

When abandoning the idea that human behaviour is determined by a set of defined objectives to which the operator is supposed to act, one is forced to abandon the idea that human behaviour can be predicted. One speaks of a prediction when the state of a system observed is claimed to be followed by a consequent (observable) state resulting from the system's structure (Maturana and Varela, 1992). However, in the case of living systems the structure of the system is continually changing subordinate to MSI, which makes it impossible to predict the future state. Therefore, the model of Maintaining Sensing Influence cannot predict human behaviour. This is in agreement with two characteristics of humans, namely autonomy and creativity, which implies that prediction is impossible.

No free choice

Due to the MSI-model based on the concept of autopoiesis, humans are not free to choose from several possibilities offered by the environment. The "choice" humans make is always subordinate to their essential process of Maintaining Sensing Influence (MSI). However, the idea that humans are free to choose how to behave contributes very much to their Noting Influence (NI) and consequently to their MSI, particularly in the Western world. So the idea of free choice is not only of great importance for establishing human NI. It also threatens their NI, when their freely chosen behaviour does not meet their expectations.

3.9 CONCLUSIONS

Our objective was to develop a behaviour model that is generic to all specific autonomous creative behaviour of humans, successful and erroneous behaviour. The model we have developed considers all human behaviour as subordinate to their critical process of Maintaining Sensing Influence (MSI). In other words, any form of human behaviour is determined by the process of MSI, irrespective of it being a success or error. MSI in turn depends on the current structure of the human which has emerged by the ever ongoing process of social representing.

The MSI-model satisfies the conditions discussed in Section 3.2: it is a generic behaviour model that is able to generate all specific human behaviour patterns. Since we have taken advantage of the characteristic common to all stable dynamic closed systems, Eigen behaviour, and have reformulated it in the critical closed process MSI, the condition of autonomy is automatically fulfilled. The characteristic of creativity is fulfilled through the structural coupling of the human operator with his environment. The involvement of the observer is captured by social representing, a human specific case of behaving.

How do we know the MSI-model is an added value for remedying human errors in real life situations? Prerequisite for the model to be an added value to the existing behaviour models is that it does not contradict the current insights in human error, moreover it should even be supported by them. That this requirement is fulfilled is argued briefly in the discussion. In the next chapter the MSI-model will be applied to two accident cases from aviation to evaluate whether the wide variety in human error and the many influences involved can be abstracted to the critical process of MSI. To illustrate the added value of the MSI-model to remedying persistent human error, it is applied in a real life setting of students who show erroneous behaviour with regard to their expected academic progress.

REFERENCES

Bateson G (1979) *Mind and Nature*. New York: Elsevier-Dutton.

Capra F (1996) *The Web of Life, A New Scientific Understanding of Living Systems*. New York: Anchor Books.

Cook RI, Woods DD (1994) Operating at the Sharp End: the Complexity of Human Error.

In: Bogner SM, ed. *Human Error in Medicine*. Lawrence Erlbaum Associates, 255-310.

Dreyfus HL, Dreyfus SE. (1986) *Mind over machine: the power of human intuition and expertise in the era of the computer*. Oxford, Basil Blackwell.

Dreyfus HL (1997) Intuitive, deliberative, and calculative models of expert performance. In: Zsambok CE and Klein GA, eds. *Naturalistic Decision Making*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 17-28.

Endsley MR (1995) Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, **37:** 32-64.

Flach JM. (2004) The Human as a Critical Component in an Adaptive Meaning Processing System. In: Thissen W, ed. *Proceedings of the International Conference on Systems, Man and Cybernetics,* xx-xx. The Hague.

Gerdes V. (1997) *Identification and Analysis of Cognitive Errors; Application to Control Room Operators.* PhD Thesis. Delft University of Technology.

Goudsmit AL. (1998) *Towards a negative understanding of psychotherapy.* PhD Thesis, Rijksuniversiteit Groningen.

Helmreich RL, Wilhelm JA, Klinect JR, Merritt AC (2001) Culture, error and Crew Resource Management. In: Salas E, Bowers CA, and Edens E, eds. *Improving Teamwork in Organizations*. Hillsale, NJ: Erlbaum, 305-331.

Kohn LT, Corrigan JM, Donaldson MS (2000) *To Err is Human; Building a Safer Health System*. Washington DC: National Academy Press.

Koningsveld, H (1976) *Het verschijnsel wetenschap, een inleiding tot de wetenschaps-filosofie*. Amsterdam. Boom.

Lipshitz R (1997) Naturalistic Decision Making Perspectives on Decision Errors. In: Zsambok CE and Klein GA, eds. *Naturalistic Decision Making*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 151-160.

Mattheeuws A (1990) Het spel der vanzelfsprekendheden. *Systeemtheoretisch Bulletin*, **8(4):** 255-287.

Maturana HR, Varela FJ (1980) *Autopoiesis and Cognition; The Realization of the Living*. Dordrecht, the Netherlands: D. Reidel; Publishing Company.

Maturana HR, Varela FJ (1992) *The Tree of Knowledge; The Biological Roots of Human*

Understanding. Boston, Massachusetts: Shambhala Publications.

McRuer DT, Jex HR (1967) A review of Quasi-Linear Pilot Models. *IEEE Transactions on Human Factors in Electronics*, **8:** 231-249.

Naikar N, Saunders A. (2002) Crossing the Boundaries of Safe Operation: Training for Error Detection and Error Recovery. In: Johnson CW, ed. *Proceedings of the 21st European Annual Conference on Human Decision Making and Control*, 8-14. Glasgow, Glasgow University.

Rasmussen J (1983) Skills, Rules, and Knowledge: Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-13:** 257-266.

Reason J (2000) Human error: models and management. *British Medical Journal*. **320:** 768-770.

Reason J (1990) *Human error.* Cambridge University Press

Varela FJ (1979) *Principles of Biological Autonomy*. North Holland/New York: Elsevier.

Varela FJ, Thompson ET, Rosch E (1991) *The Embodied Mind; Cognitive Science and Human Experience*. Cambridge MA: MIT Press.

Von Foerster H (2003) *Understanding understanding*. New York: Springer.

Watzlawick P, Beavin JH, Jackson DD (1967) *Pragmatics of Human Communications*. New York: Norton & Company, Inc.

Weick KE (1995) *Sensemaking in Organisations*. Thousand Oaks, CA: Sage Publications.

Wickens CD (1992) A Model of Human Information Processing. In: *Introduction to Engineering Psychology and Human Performance*. New York: HarperCollins Publ., 16-21.

Winograd T, Flores F (1986) *Understanding Computers and Cognition; A New Foundation for Design*. Norwood NJ: Ablex.

Woods DD, Johannesen LJ, Cook RI, Sartert NB (1994) *Behind Human Error: Cognitive Systems, Computers, and Hindsight*. Columbus: CSERIAC.

Zeegers W (1988) *Andere tijden, andere mensen, de sociale representatie van identiteit*. Amsterdam, the Netherlands: Bert Bakker.

Chapter 4

RE-EXAMINATION OF TWO CASES OF PERSISTENT HUMAN ERROR[1]

The system and contextual approach to failures has led to an increased safety and performance in aviation. However, this increased safety made human failures more visible in unsafe outcomes. Today, unsafe outcomes develop because of an accumulation of small human failures. These sequences of human failures that accumulate despite the remedial measures have been characterised as fixation or preoccupation. Fixation is characterised as holding on to a course of actions or plan whereas they should have been changed due to evidence at odds with them. In examinations of these accidents, the Safety Board ascribed fixation to poor situation awareness (SA) of the crew. However, poor SA cannot be an explanation for fixation, because it is a result. Sensing Influence (SI) on the other hand, can give an explanation. This is related to the axiom that maintaining SI is the essential process underlying all results, irrespective of error or success. To show that SI can offer an added value with respect to the construct SA, two cases of crew-involved accidents examined by the Safety Board will be re-examined based on the model of Maintaining Sensing Influence (MSI-model).

1.This chapter has been submitted as: "*Getting a new perspective on persistent human failure: a re-examination of two cases of aviation accidents by applying the MSI-model,*" by M. van der Vlugt, P.A. Wieringa, and J. Dankelman

> *"... I do not always do what I am told to do. And this is because when people tell you what to do it is usually confusing and does not make sense. For example, people often say "Be quiet," but they don't tell how long to be quiet for."*
>
> Mark Haddon, 2003

4.1 INTRODUCTION

In aviation accidents and near misses rarely appear at once full-sized, they are in general the end result of an accumulation of small human errors that are not trapped by the available measures that should have remedied them or rendered them harmless (Helmreich 2000; Reason 2000). Several studies into aviation accidents revealed a typical persistent form of human (decision) error, fixation. Fixation is characterised as holding on to a course of actions in the face of cues such as alarm warnings or alerts, or crew suggestions, that indicate a change is necessary (De Keyser et al. 1990; Woods *et al.* 1994; Xiao and Mackenzie 1995; Orasanu *et al.* 1998). This form of persistent human error is also known as tunnel vision or cognitive lockup (Moray and Rotenberg 1989) and preoccupation (NTSB 2000, 2001).

The National Transportation and Safety Board analysed 37 crew-involved accidents from 1978 to 1990 (NTSB 1994). A study of Orasanu, Martin and Davison (1998) examined the results of the Safety Board and found that 25 of the 37 accidents could be ascribed to so called tactical decision errors. The study stated that common to decision errors was that the crews showed a tendency to maintain their originally selected course of actions despite evidence at odds with them. Examples included deciding to continue the flight in the face of a system malfunction, continuing an unstable approach and flying into deteriorating weather (Orasanu and Fischer 1997). They concluded that in most of these cases the crews exhibited a poor situation awareness. Situation Awareness (SA) is defined as an understanding of the state of the environment (Endsley 1995b).

Flach (1995) has cautioned against considering SA as explanation of human (erroneous) behaviour. In our view this is related to the fact that SA is a construct (Endsley 1995a), an output (state) of complex information processing. The operator's level of SA is derived from whether or not the operator is able to achieve an agreed goal. This implies that SA and achieving this goal cannot explain each other causally, and vice versa. Therefore a deviation from an agreed goal cannot be explained with a poor SA or a loss of SA of the operator. Unless the process of persistent human errors is understood, i.e. what

causes loss of SA, it will be difficult to remedy these errors instantaneously. The MSI-model we have been developing in Chapter 3 explains persistent human error such as fixation via the process Maintaining Sensing Influence (MSI), the basis of all human behaviour, irrespective of their level of SA. As a phenomenon description, SA invites to research causal relations between human-machine systems and human performance (Flach 1995).

In this chapter it will be shown that for explaining persistent human error, the process MSI can offer an added value with respect to the construct SA. Therefore, two cases of crew-involved accidents that have been examined by the Safety Board will be re-examined based on the MSI-model: the controlled flight into terrain of flight 801 of Korean Air, and the runway overrun of flight 1420 of American Airlines (NTSB 2000, 2001). In both cases the actions of the crew (human failures) were considered to be decisive for the accident. The case-specific explanations of human failures obtained by the NTSB are given together with the explanation obtained by applying the MSI-model. The case-specific MSI explanations will be generalised. Next, the difference is discussed between SA and SI. It is shown that with the MSI-model, starting points can be found to develop remedial measures additional to the existing measures.

4.2 Understanding human failures via Maintaining Sensing Influence

In this section a short outline of the MSI-model is given together with its perspective on active failures. The MSI-model connects the wide variety of human failures through the key process common to all human behaviour: Maintaining Sensing Influence (MSI). For an extensive treatment of this model is referred to Chapter 3.

Outline of the Maintaining Sensing Influence (MSI) model

The MSI-model is founded on two axioms. The first axiom is that specific behaviours of each individual human are subordinate to their continuation of existence, or in the words of Maturana and Varela, to their “Maintaining Organisation” (MO) (1980, 1992). As long as a human continues to exist it Maintains Sensing Influence (MSI). Sensing Influence (SI) is the way humans assign meaning to environmental influences that affect them. A special reflexive form of SI is Noting Influence (NI). In case of NI, a human reflects on the effect he expects to happen and the effect he actually senses in response to his

influence. Comparable existing terms that describe NI as phenomenon are wariness (Weick 1999) or mindfullness (Varela, Thompson & Rosch 1991). NI can be lost, contrary to SI, which cannot be lost. The second axiom of the MSI-model is that a human's existence (as a human) only emerges in social coupling with other humans. In social coupling, humans mutually interpret each other's coordinated actions (i.e. they mutually sense influence) and act according to their interpretation of these coordinated actions. With social coupling mutual expectations about appropriate behaviour emerge. Although interpreting (i.e. Sensing Influence) is mutually congruent because of social coupling, it is in essence a personal matter due to humans' autonomy. The historical development of a human's SI is called context. Context assigns meaning to coordinated actions.

Based on these two axioms, three contexts have been distinguished. Firstly, the reference context, which is formed by the explicitly defined established rules of appropriate behaviour. In aviation, for example, these rules are the FAA regulations and company procedures. In the reference context, these rules of appropriate behaviour guide one's actions. In this context, NI rules because expectations and experienced effects match by definition. Secondly, the situational context, in which there are no established rules of appropriate behaviour yet, because the situation one faces is new. In this context NI can be threatened depending on whether or not the situation develops according to one's expectations. Contrary to the reference context, the situational context has no explicitly defined stopping rule of appropriate behaviour. And finally, the evident context, which is formed by the tacit habitual established rules of appropriate behaviour between the involved humans. These habitual rules have emerged unnoticed and are so common that they are not reflected upon, i.e. they are self-evident. These rules of tacit habits are the so-called self-evidences. In this context NI is lost and SI rules.

Human failures reveal a context discrepancy

Active failures in the view of the MSI-model implies that the observer and the operator do not share anymore a synchronous context, in other words, they do not share anymore their meaning about their coordinated actions. There is a context discrepancy between observer and operator.

In theory, when the operator maintains NI, active failures do not occur. Continually Noting Influence implies that for every situation the operator faces, he both has to predict the effects of his course of actions and has to evaluate the actual effects. This continually maintaining NI is highly demanding and therefore hard to maintain for the operator

(Weick, 1995, Wickens 1993). When the operator has lost NI, he has changed to an evident context and the likelihood of occurrence of active failures increases. Loss of NI occurs in two ways:

1. when becoming accustomed to the ways of appropriate behaving, i.e. habituation and experience,
2. when facing new situations in which the effects of actions are uncertain.

Although loss of NI does increase the likelihood of committing an active failure, it does not necessarily mean that an active failure will actually be committed. It is possible that the operator acts seemingly normal, that is, as if he has NI. Therefore, lost NI does not necessarily have to be recognised by an observer. As long as the operator acts habitually according to the general rules of appropriate behaviour, the observer may not recognise that the operator has lost NI. In practice, it is not unusual for humans to act without NI, when they are accustomed to the situation. For example, after work we sometimes drive home on the "automatic pilot," i.e. without NI. When arriving home we wonder if we paid attention to the traffic lights. Did we drive through a red light or not? When arriving home safely without breaking the rules of appropriate behaviour (i.e. without committing active failures), our driving home without NI was not necessarily recognisable by an observer. An observer cannot detect lost NI. Only when an operator commits an active failure, we as observers may conclude, based on the concept of MSI, that the operator has lost NI and changed autonomously to an evident context that maintains his SI. Thus, in the view of the MSI-model, an active failure is, for the observer, a visible effect of a change to an evident context with which the operator maintains SI.

4.3 Case 1: Flight 801 of Korean Air

The re-examination of the case of flight 801 of Korean Air focuses on the decision-making of the crew in the approach and landing phase of the flight. For a detailed examination the reader is referred to the full report of the Safety Board (NTSB 2000). The last section of the final approach is shown in Figure 4.1.

Short outline of the case

In 1997, Korean Air flight 801 crashed into terrain about three miles southwest of Guam International Airport. The National Transportation and Safety Board determined that the probable causes of the accident were (NTSB 2000):

- the captain's failure to adequately brief and execute the nonprecision approach,
- the first officer's and flight engineer's failure to challenge the captain's execution of the approach.

In general there are two ways to execute an approach: a precision approach and a nonprecision approach. In the case of a precision approach the guidance for the crew is the glideslope, a lateral and longitudinal guidance path displayed on the flight director. In the case of a nonprecision approach two possibilities exist, a visual approach when vision is sufficient and an instrument approach when vision is insufficient. For a visual approach the main guidance are the runway lights. For a nonprecision instrument approach the guidance for the crew is to maintain a required vertical speed, which is displayed on the flight director and which is set manually.

The glideslope was out of order, consequently the captain was left two possible though not required courses of actions: an instrument nonprecision approach and a visual nonprecision approach. Both courses of actions were not optimal for the captain for the following two reasons. Firstly, against an instrument approach was the unusual location of the distance measuring equipment (DME) at Guam. Normally, the DME is located at the airport, but at Guam the DME is located 3.3 miles outside the airport making the instrument approach at this airport more complex than a usual nonprecision instrument approach (NTSB, 2000). In addition, an instrument approach at Guam should be carried out via a step-down profile, although a gradual descent profile was allowed and was even preferred with a Boeing 747-300. This made the nonprecision instrument approach even more demanding. The complexity of a nonprecision instrument approach is supported by the statements of the Air Line Pilot Association (ALPA) and a British Airways Boeing 777 captain who was a member of the Controlled Flight Into Terrain (CFIT) Awareness Task Force (NTSB, 2000). Secondly, against a visual approach was the unsettled weather. Due to the unsettled weather the "decision" for a visual approach might only postpone the necessary change to the complex nonprecision instrument approach. In favour of the visual approach is the extra time gained because the aeroplane has to be stabilised at 500 feet instead of 1000 feet. In addition, a visual approach is simpler than a nonprecision instrument approach especially at Guam.

Considering the unsettled weather conditions, the nonprecision instrument approach should have been preferred to a visual approach though not specifically required. Nevertheless, in the last minute of the approach the crew was still attempting a visual ap-

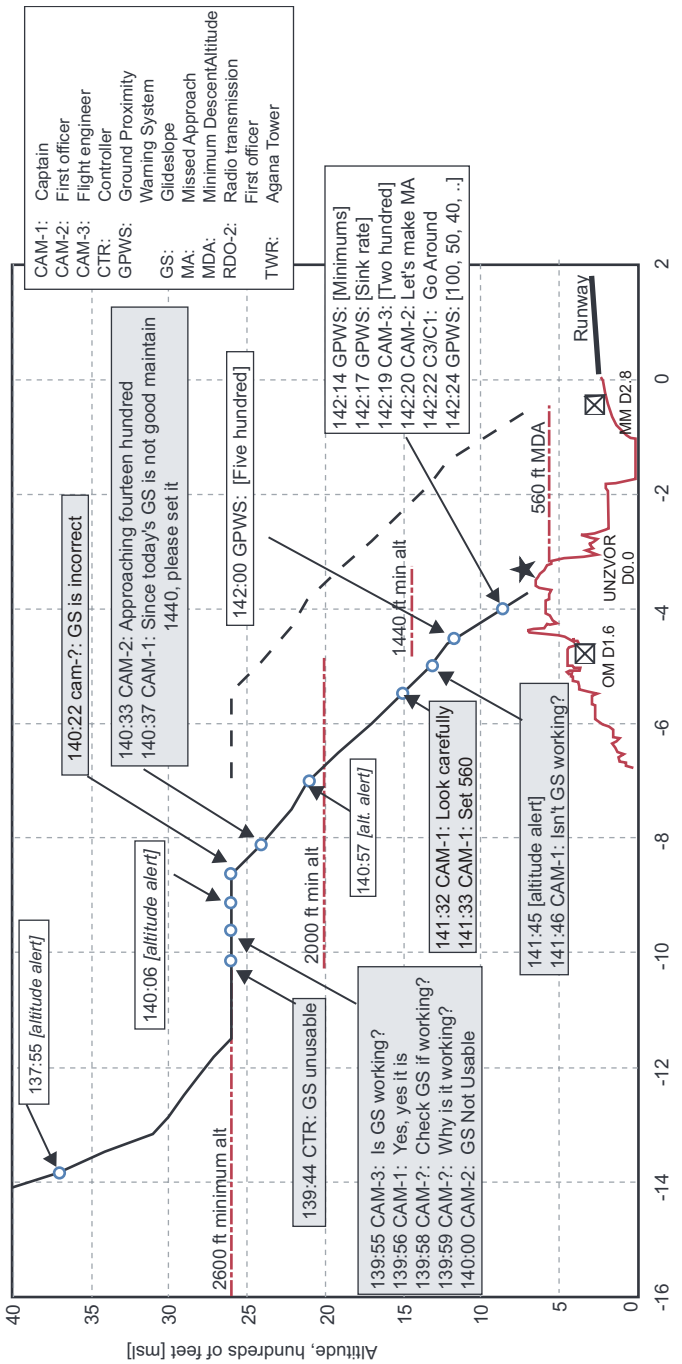


Figure 4.1 *The actual descent profile of Flight 801 of Korean Air together with the required levels of the step-down profile and the descent profile possibly presumed by the crew. The Cockpit Voice recorder information adapted from the NTSB (2000). The apparent confusion about the glideslope and the warnings of the Ground Proximity Warning System (GPWS) indicate that the crew lost NI and changed to an evident context.*

proach and looking for the runway when they were warned by the ground proximity warning system (GPWS) for 500 feet shortly before fligth 801 crashed into terrain. Looking at the descent profile the Board assumed that it was possible that the crew members might have believed that the plane was 3.3 miles closer to the airport (NTSB 2000).

Explanation for the crew's contribution given by the NTSB.

The Safety Board concluded that the captain had a strong expectation that a visual approach was possible. The board also concluded that the captain likely experienced increased stress as the approach progressed without encountering the expected visual conditions, because the nonprecision instrument approach was inadequately prepared. The Safety Board also concluded that the captain lost awareness of flight's 801 position, as a result of the captain's confusion and preoccupation with the glideslope status, he failed to properly check the aeroplane's position. The NTSB ascribed this failure to the findings that pilots under stress might have a degraded performance level.

In addition, the Safety Board concluded that also contributing to the accident was the failure of both the first officer and the flight engineer to challenge the captain's execution of the approach. A study of the Safety Board found that the failure of first officers to challenge the captain's errors, especially tactical decision errors, is a frequent factor involved in CFIT accidents. Nevertheless, the precise reasons for this failure could not be given by the Safety Board.

Explanation for the crew's contribution based on the MSI-model.

The accident of flight 801 has been categorised as a controlled flight into terrain, which means that the crew's failures were seriously contributing to the accident. According to the MSI-model, the captain and crew lost NI in a situational context, and maintained SI by changing to an evident context when they had to decide for an approach due to their uncertainty about the possible consequences.

Considering the two suboptimal courses of actions and the sequence of crew failures, it seems that the captain was acting in two minds. In the first instance the captain aimed for a visual approach and kept open a possible change to an instrument approach, if this would be necessary. A possible change from visual to instrument approach was kept open by reckoning with both the rules of a visual approach and the rules of the nonprecision instrument approach. Acting in two minds itself is not incorrect. However, in this

case the captain's first mind was aiming for a visual approach which was complemented with a nonprecision instrument approach. Considering the unsettled weather conditions which could make a visual approach impossible, the captain's first mind should have been aiming instead for a nonprecision instrument approach according to the NTSB, which could be complemented with a visual approach as long as permitted by the weather conditions. But the nonprecision instrument approach threatened the captain's SI in the first place. The visual approach became the objective itself with which the captain maintained his SI.

There are two indications that support the hypothesis that the captain may have lost his NI due to his uncertainty of the consequences of both courses of actions, and maintained SI by changing from a situational context to an evident context . The first indication for a lost NI is that the captain mainly briefed about a visual approach, whereas he only briefed some parts of an instrument approach in case a visual approach would be impossible. The instrument part was insufficiently briefed according to the company rules. The second indication that the captain had lost NI, is his hope that the glideslope was still operative. He first stated that glideslope was working and then he questioned, "isn't glideslope working?" (see Figure 4.1 at 0139:56 and 0141:46 respectively). This contradicted both his own approach briefing in which he informed the crew that the glideslope was inoperative and the statement of the control tower that glideslope was unusable (see Figure 4.1 at 0139:44).

The consequence of maintaining SI in an evident context is that signals referring to the actual situation do not trigger the captain to change course of actions because this would mean that he should turn back to a situational context which threatened his SI in the first place. From the NTSB report and the CVR transcript it becomes clear that in the last minute the salient warning signals of the GPWS referring to a serious problem in the actual situation have been glossed over. It is important to notice that the signals were not ignored in the sense that the crew did not pay attention to them. On the contrary, the signals were perceived by the crew, but the signals were interpreted in favour of the course of actions the crew took. In other words, the GPWS signals which are situation related signals were interpreted by the crew according to their evident context. As long as the crew maintains SI in an evident context, signals that only refer to the state of the aeroplane (i.e. situational context) and do not refer to the state of the crew, do not initiate a change in course of actions of the crew.

4.4 Case 2: Flight 1420 of American Airlines

The second case taken from the NTSB to be re-examined, concerns flight 1420; a crash after a runway overrun (for details see: NTSB 2001). The critical moments with regard to the decision-making of the crew occurred in the final approach phase. A top view of the approach path of flight 1420 together with the location of the thunderstorms is shown in Figure 4.2.

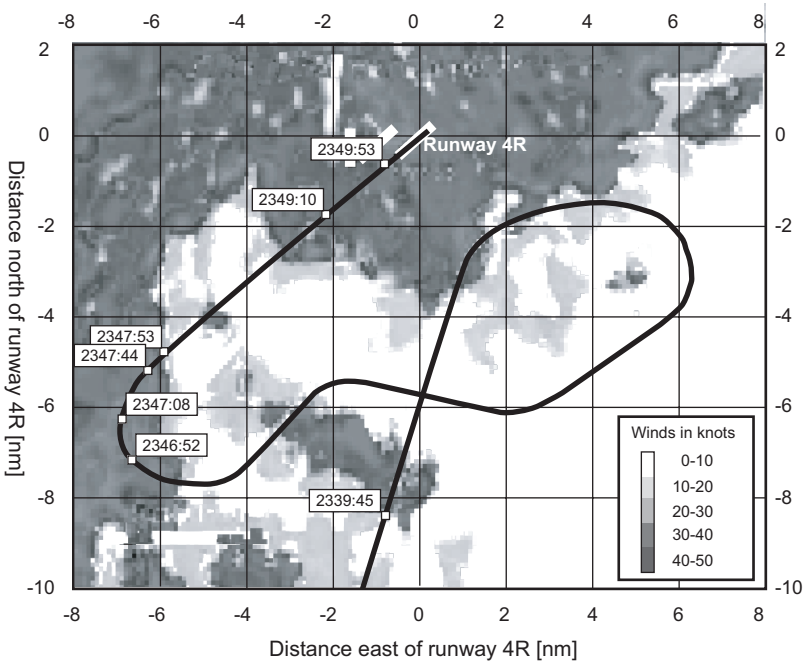
A short outline of the case

On June 1999, American AIrlines fligth 1420 crashed after it overran the end of the runway during landing at Little Rock National Airport, Arkansas. According to the Safety Board, the probable cause of the accident was the crew's failure to discontinue the approach when severe thunderstorms and their associated hazards to flight operations had moved into the airport area (NTSB 2001). A contributing factor to the crew's failure to discontinue the approach was the company's procedure for a stabilised approach that contained unclear criteria when an approach should be discontinued.The Safety Board also concluded that the lack of spoiler deployment was the single most important factor that the crew could not manage to stop the aeroplane within the runway length. It should be noticed that this conclusion implies that the approach itself, although not perfect, was not causal to the accident.

In the approach phase, the crew faced three possible courses of actions:

- hold and wait for better circumstances,
- diverting to an alternate airport,
- expediting the landing to beat the thunderstorms.

The developing weather (heavy rain, dropping visibility and increasing crosswinds) would have been sufficient to hold and wait, or to divert. According to the Safety Board, flight 1420 would have had enough fuel for these two possible courses. Against these two courses was that the situation at the moment the crew entered the final approach phase did not require executing one of them according to the Standard Operating Procedures (SOPs) and the regulations of the FAA. Against the plan of expediting the landing was the uncertainty about how the weather would develop. A possible factor favouring the plan of expediting the landing could have been the company duty time limitation. The delayed departure time was only half an hour within the latest possible departure



2254:xx Decision for expediting landing, no discussion about the two alternatives holding or diverting

2234:11 Controller giving weather report stating winds from 280 dg at 28 knots gusting to 44 knots

2239:45 First wind shear alert provided by controller, reporting centerfield wind 340 dg at 10 knots indicating that the winds from 330 dg are 25 knots

2246:52 Captain: "we're going right into the storm."

2347:08 Controller giving weather report of a wind from 350 dg at 30 knots gusting to 45 knots (i.e. crosswind component at 23 knots gusting to 34 knots). No discussion between the crew that the winds reported exceeded the maximum crosswind component of 15 knots

2347:44 Captain calls out "gear down," indicating that 2nd half Before Landing Checklist must be accomplished, e.g. spoiler lever, autobrakes, and flaps. None had been performed correctly.

2347:53 Second wind shear alert provided by the controller stating a wind of 350 dg at 32 knots gusting to 45 knots. Still no discussion about the exceeded maximum crosswind component.

2349:10 Another wind report by the controller of 330 dg at 28 knots. Captain: "this is a can of worms."

2349:57 First Officer calls quietly "go around." No actions upon this call by the first officer self nor the captain

2350:20 Touch down

Figure 4.2 *Case overview of the last phase of Flight 1420 of American Airline. Above the Flight 1420's approach path to runway 4R is plotted against the information of the weather radar 3 minutes before touchdown at 2350:20 local time (adapted from NTSB 2001). Below the communication of the crew selected from the cockpit voice recorder transcript.*

time. A delay in landing of more than half an hour would have led to exceeding the duty time limitation. However, the Safety Board concluded that there was no evidence indicating that this factor played a role in the crew's plan to expedite the landing (NTSB 2001). In addition, expediting the landing was suggested by the dispatcher.

The crew "decided" for expediting the landing, a course of actions that was backed up by the rules. When executing the plan of expediting the landing, the last escape in the final approach phase is to execute a missed approach in the case that the approach is not stabilized according to the company criteria or when other criteria threatening a safe operation are exceeded. By accepting a short approach the crew increased the time pressure in which they had to evaluate a significant amount of weather information simultaneously. In this phase of high workload, the crew missed the decisive criterion of the exceeded cross wind limitations for executing a missed approach. In addition, the crew performed incorrectly the "Before Landing Checklist," they did not arm the autospoiler nor verified its status. After touchdown the crew failed to verify that the spoilers had automatically deployed and inherently the captain failed to do it manually. Flight 1420 overran the runway and crashed.

Explanation for the crew's contribution given by the NTSB

The Board concluded that the situation should have heightened the crewmembers' awareness that they might not be able to continue the approach safely. The Safety Board concluded as well that the crew's level of performance was degraded with respect to the level of performance that would have been expected considering their experience and qualifications. The degraded performance level of the crew was ascribed by the Safety Board to possible stresses to the crew due to the rapidly deteriorating weather conditions. The difficulty is that weather conditions cannot be controlled and therefore remain an uncertain factor. In addition, the Safety Board stated that stress can also induce the crew to maintain their original course of actions until there is clear and overwhelming evidence at odds with them. In this specific case the Safety Board found that the company's criteria for a stablised approach did not clearly indicate when to initiate a missed approach.

As a remedial measure for this specific problem, the Safety Board recommended a definition of detailed parameters for a stabilised approach indicating clearly when a missed approach should be performed. Although this may be successful for many situations, it will never be wholly successful. Strictly viewed it is possible to pursue perfectly defined

criteria for a stabilised approach that minimise the chance of entering unclear situations. However, the chance of entering unclear situations always remains because there is no neat end to the definition of real life situations. What is the maximum deviation allowed from the perfect approach? Should these criteria incorporate all weather conditions? Do the criteria account for all situations or are there exceptions, and what are these exceptions, etc.? However, even if the crew had had the additional decisive criteria for executing a missed approach, it is not certain that the crew would have executed it, because they did not notice the exceeding crosswinds limitations which were decisive for initiating a missed approach.

Explanation for the crew's contribution with the MSI-model

Considering the accident is primarily ascribed to crew failures, we may conclude based on the MSI-model that the crew lost NI. The starting point for any MSI-explanation is that the crew always maintains SI. The crew maintained SI by changing to an evident context, for they committed human errors, they "ignored" situational signals that should have urged them to alter their course of actions.

The question, "how did the crew lose their NI?" can be answered by analysing the courses of actions they faced: holding, diverting, or expediting. Holding or diverting were not required for the situation they faced. Consequently, the consequences when holding or diverting would have been the captain's responsibility. They would not be backed up by company rules if anything were to go wrong, or if they turned out to be not necessary at all when the weather had not deteriorated after all. However, the consequences for expediting the landing were uncertain as well, due to the rapidly deteriorating weather conditions. According to the MSI-model, uncertainty about the consequences of all three courses of actions led to a loss of crew's NI threatening the crew's SI. Because SI cannot be lost the crew maintained SI by changing to an evident context.

An indication that the crew "decided" without NI for expediting the landing is the first officer's statement that there was at that moment no discussion of delaying or diverting the landing (NTSB 2001). In other words, it was self-evident that they decided for expediting the landing. With the MSI-model this can be interpreted as: "expediting the landing implies a safe landing," instead of: "a safe landing may imply that the landing has to be expedited." In other words, the dependency between safe landing and expediting landing was turned around. It is important to note that deciding for expediting the landing itself would not necessarily have led to the accident, because there was still the escape

through a missed approach procedure. It was the manner in which the crew decided for expediting the landing, that is without NI, that contributed to the accident. It was possible that when the crew decided for expediting the landing with NI they could still have performed the “Landing Checklist” correctly and could have armed the autospoiler, or otherwise noted that the spoilers did not deploy and so could have corrected it manually, resulting in a stop within the runway length. In the view of the MSI-model the incorrect performing of the “Landing Checklist” is only a confirmation of lost NI which already had become manifest in the way the crew decided for expediting the landing.

4.5 EXTENDING THE MSI-EXPLANATION OF BOTH ACCIDENTS TO AVIATION IN GENERAL

In aviation an exhaustive reference context has been defined: the Standard Operating Procedures (SOPs). The SOPs define exactly the appropriate courses of action for known situations. The SOPs have been developed to shape the crews actions so as to minimise their chances to encounter situations that may initiate human failures (Reason 2000, Kaempf and Orasanu 1997). From the perspective of the MSI-model, the SOPs reduce the chance that the crew encounters a situational context that may threaten their NI. Due to the SOPs the crew has only to assess the situation, from which the appropriate course of action follows. Because the crew does not have to generate and evaluate possible courses of action, the SOPs support the crew in their effort to maintain NI. The SOPs are generally acknowledged as the basis for safe flight operations. A drawback of the exhaustively defined SOPs is that the scope for personal interpretation of situations has shrunk to such an extent that the crews can be hampered when the situation requires creative solutions that go beyond the defined SOPs.

Re-examination of the two accident reports carried out with the MSI-model reveals that active failures were manifestations of lost NI in situations where the SOPs were insufficient, followed by a change to an evident context with which the crew maintained SI. Both cases show that active failures occurred when the crew faced a specific situation for which there was no suitable SOP prescribing the obvious course of actions. This means that the crews had to go beyond the SOPs.They found themselves inevitably in a situational context in which their NI was challenged when they had to decide for uncertain consequences.

When a crew has to act beyond the SOPs it has two consequences. Firstly, there is no end rule anymore that explicitly defines when a change in the course of action is necessary. Whereas the SOPs represent generalised rules of safe operation (reference context) adopted (shared) by the participants, the situational context represents particular rules for safe operation that are not explicitly specified. Thus if the crew's actions are beyond the SOPs and are leading to a probable unsafe outcome it is hard for them to define the point of safe return. Secondly, when the crew recognises that their actions are probably leading to an unsafe outcome they cannot fall back on the SOPs because they did not satisfy the situation then, and they certainly do not satisfy the situation anymore at that moment. When the crew loses NI at this moment, they can only maintain SI by changing to an evident context. Changing to an evident context means that they assign a positive meaning to their failures. Changing to an evident context is not a matter of choice, it is a matter of necessity regarding the MSI-model. Note, that acting beyond the SOPs does not necessarily mean that the crew loses NI. As long as the crew maintains NI and acts according to a situational context, they act creatively and are an added value for safe flight operations.

The crew's inability to act to evidence at odds is explained by the MSI-model by changing to an evident context, because the contradictory evidence refers to the reference context. When acting in accordance with an evident context, active failures are assigned with a positive meaning that does not correlate anymore with the situational context. Therefore, evidence at odds that relates to the reference context will be assigned with a positive meaning by the crewmembers as long as they are in an evident context. For instance, driving through a red light, is always accompanied by assigning a positive meaning to this violation of traffic rules, like "since there is nobody else around and I am in a hurry due to an emergency, I am allowed to drive through the red light." Note also, that once the crew has entered a situational context, falling back on the reference context (the SOPs) is not possible anymore because the situation the crew faced already required acting beyond the SOPs.

Following the MSI-model, human errors are the visible manifestations of a change to an evident context which does not synchronise with the observer's context. Maintaining SI in an evident context means that human errors and evidence at odds are interpreted in favour of self-evident rules that do not match the observer's situational or reference context. The human errors of the crew, for instance its failure to cross check the aeroplane's position, not noticing that the "before landing checklist" is not completed, and

not noticing the exceeding of crosswind limitations, appear to the observer as violating company rules or ignoring contradictory evidence.

4.6 THE ADDED VALUE OF SI WITH REGARD TO SA

The perspective taken to re-examine both cases corresponds with the findings of the NTSB and with previous findings of investigations into human factors, which pointed out that in the case of human failures:

1. Operators acted outside the reference context (for instance the SOPs), because it was inappropriate for the situation they faced (Cook and Woods 1994, Reason 1995, Kaempf and Orasanu 1997, Woods *et al.*1994)
2. Operators did not act in accordance to contradictory evidence (Wickens 1993, Reason 1995, De Keyser 1990, Xiao and Mackenzie 1995)

The crew's inability to act on contradictory evidence is explained by the Safety Board with poor SA, because if the crew had sufficient SA then they would have acted in accordance with the contradictory evidence. SA is related to the level of understanding of the state of the environment. To improve SA, the measures among others proposed by the Safety Board, were to define more precisely the criteria for a stabilized approach, through which the level of understanding of the state of the environment should improve. Actually, these measures only aim at reducing the uncertainty of the decision-situation (work-situation). They do not remedy *human* error. By improving the work-situation human error is only indirectly influenced. Moreover this approach will fall short for remedying persistent human error because of the impossibility to define in advance all possible work-situations the crew possibly may encounter. There will always be potentially critical situations not included in the anticipated work-situations (Flach 2004).

SI on the other hand, does give starting points for developing measures for remedying human error instantaneously, because MSI is the process essential to all specific behaviour, irrespective of error or success. This implies that the state of SA can be explained with the process of MSI. According to the MSI-model, SA can be interpreted as SI that is coupled with an observer preferred goal. SA is therefore an observer interpretation of the actual result achieved by the operator, compared with the result desired by the observer (i.e. the observer preferred goal). The actual result achieved by the operator is the outcome of the process of achieving the observer preferred goal.

Assuming that the observer has NI and observes the operator's actions, there are two views on the operator's performance:

1. The operator has SA in the view of the observer when the SI of the observer and the SI of the operator are maintained in synchronous contexts.
2. The operator has poor SA in the view of the observer when the SI of the observer and the SI of the operator are maintained in asynchronous contexts.

In undesired situations that emerged because of human error, the behaviour of the operator is the only 'factor' that can be changed instantaneously. By understanding the MSI-process of the operator in relation to the MSI-process of us, the observers, it is possible to develop additional remedial measures that can establish instantaneously a recovery from human error. Premise for establishing a successful change to desired behaviour is that the alternative offered to the operator is able to restore his NI in a situational context. Chapter 6 will consider the starting points for developing remedies.

In the past concentrating on the role of the operator in human error led to blaming him for his errors (Reason 2000). Our renewed focus on his role concerns his essential contribution in operating the system, for better or for worse. Making errors is human they say, as is recovering from them.

4.7 Conclusions

Before measures can be developed to remedy persistent human error, it is necessary to put a finger on their essence. According to the model of Maintaining Sensing Influence, all human failures are manifestations of a loss of Noting influence of the crew in the situational context that coincide with a change to an evident context with which the crew maintains SI. This hypothesis is tested by applying the MSI-model to the two accident cases of KAL 801 and AA 1420. The model illustrated in both cases that human failures happen and persist in happening when the crew has to act outside the Standard Operating Procedures (i.e. in a situational context), and when the consequences of the course of actions are unclear or less than optimal with regard to the organisational expectations, i.e. safety. When the crew maintains SI in an evident context, it means that they interpret their actions and contradictory evidence in favour of self-evident rules. In the case of human error these self-evident rules do not match the observer's situational or reference context. The human errors of the crew, appear to the observer as violating company rules or ignoring evidence at odds. Moreover, when the crew maintains SI in an evident

context, they cannot fall back on the Standard Operating Procedures forming the reference context since these did not satisfy in the first place.

Loss of Situation Awareness (SA) did not satisfy for explaining fixation, because fixation is a state of which SA is derived. Maintaining Sensing Influence on the other hand can explain fixation, because MSI is the essential process underlying all human behaviour. SA can be viewed as SI that is coupled with an observer preferred goal. So loss of SA becomes SI as soon as SA is operator coupled.

To apply the model of Sensing Influence without knowledge of the outcome requires a further validation of this model and a further assessment and sophistication of the methodology based on new cases of which the outcome is unknown beforehand. Nevertheless, the viewpoint presented in this chapter based on the process Maintaining Sensing Influence by changing context could give new insights for developing additional means and improving existing means to remedy human failures.

REFERENCES

Cook RI, Woods DD (1994) Operating at the Sharp End: the Complexity of Human Error. In: Bogner SM, ed. *Human Error in Medicine*. Lawrence Erlbaum Associates, 255-310.

Endsley MR (1995a) Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors,* **37(1)**: 32-64.

Endsley MR (1995b) Measurement of Situation Awareness in Dynamic Systems. *Human Factors,* **37(1)**: 65-84.

Flach JM (1995) Situation Awareness: Proceed with Caution. *Human Factors,* **37(1)**: 149-157.

Haddon M (2003) *The Curious Incident of the Dog in the Night-Time*. USA: Vintage Books.

Helmreich RL (2000) On error management: lessons from aviation. *British Medical Journal*, **320:** 781-785.

Kaempf GL, Orasanu J (1997) Current and Future Applications of Naturalistic Decision Making in Aviation. In: Zsambok CE and Klein GA, eds. *Naturalistic Decision Making*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., 81-90.

Keyser Vd, Woods DD (1990) Fixation Errors: Failures to Revise Situation Assessment in Dynamic and Risky Systems. In: Colombo AG and Bustamante ASd, eds. *Systems Reliability Assessment*. Dordrecht: Kluwer Academic Publishers, 231-252.

Maturana HR, Varela FJ (1980) *Autopoiesis and Cognition; The Realization of the Living*. Dordrecht: D. Reidel; Publishing Company.

Maturana HR, Varela FJ (1992) *The Tree of Knowledge; The Biological Roots of Human Understanding*. Shambhala Publications.

Moray N, Rotenberg I (1989) Fault Management in Process Control: Eye movements and action. *Ergonomics*, **32:** 1319-1342.

National Transportation and Safety Board (1994) *A Review of Flightcrew-involved Major Accidents of U.S. Air Carriers, 1978 through 1990.* NTSB/SS-94/01. Washington DC, NTSB.

National Transportation and Safety Board (2000) *Controlled Flight into Terrain, Korean Air Flight 801 Boeing 747-300 HL7468 Nimitz Hill Guam.* NTSB/AAR-00/01. Washington DC, NTSB.

National Transportation and Safety Board (2001) *Runway Overrun during landing, American Airlines Flight 1420, McDonnell Douglas MD-82, N215AA, Little Rock Arkansas.* NTSB/AAR-01/02. Washington DC, NTSB.

Orasanu J, Martin L, Davison J (1998) Errors in Aviation Decision Making: Bad Decisions or Bad Luck? In: *Proceedings of the Fourth Conference on Naturalistic Decision Making* Warrenton Virginia.

Orasanu J, Fischer U (1997) Finding Decision in Natural Environments: The View from the Cockpit. In: Zsambok CE and Klein GA, eds. *Naturalistic Decision Making*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., 343-357.

Reason J (1995) A systems approach to organizational error. *Ergonomics*, **38:** 1708-1721.

Reason J (2000) Human error: models and management. *BMJ*, **320:** 768-770.

Varela FJ, Thompson ET, Rosch E (1991) *The Embodied Mind; Cognitive Science and Human Experience*. Cambridge MA: MIT Press.

Watzlawick P (1989) *Munchhausens haren en Wittgensteins ladder: uit de greep van de werkelijkheid*. Deventer: Van Loghum Slaterus.

Weick KE (1995) *Sensemaking in Organisations*. Thousand Oaks, CA: Sage Publications.

Weick, KE, Sutcliffe KM, Obstfeld D. (1999) Organizing for high reliability processes of collective mindfullness. *Res Organizational behavior,* **21:** 23-81.

Wickens CD, Huey BM (1993) *Workload Transition: Implications for Individual and Team Performance.* Washington, D.C., National Academy Press.

Woods DD, Johannesen LJ, Cook RI, Sarter NB (1994) *Behind Human Error: Cognitive Systems, Computers, and Hindsight*. Columbus: CSERIAC.

Xiao Y, Mackenzie CF (1995) Decision Making in Dynamic Environments: Fixation Errors and Their Causes. In: *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, 469-473.

Chapter 5

MSI, A NEW CONCEPT TO RECOVER FROM UNDESIRED DECISION BEHAVIOUR

In critical situations decision-makers (i.e. pilots) may perform errors. The number of errors could be reduced if it would be possible to recover from undesired human behaviour. Therefore the system theoretical concept of Maintaining Sensing Influence is introduced. It assumes that whatever a human (the actor) decides or does, he always maintains sensing influence. If the actor's Sensing Influence (SI) threatens to disappear, he may change to another context by which his SI is maintained. When the observer uses the situational context, he will value the actor's behaviour as undesired. Having identified the context change, SI becomes a criterion by which the observer can understand the actor's undesired behaviour. This understanding can be used to recover desired behaviour. The concept has been tested successfully in a student counselling setting. The student setting appears to be comparable to the pilot setting and, therefore, it may be useful to evaluate this concept in the pilot setting as well.

***Keywords**: aviation, undesired behaviour, decision-making, autopoiesis, Maintaining Sensing Influence.*

This chapter partly appeared in:
Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (pp. 2574-2579). The Hague, The Netherlands. 2004.

The following terms have been changed to bring them in line with the terms that are currently used in this thesis:

Context shift/switch:	Context change
Critical variable MSI:	Essential process MSI
Frame of Reference (FoR):	Reference context

5.1 INTRODUCTION

Although it is acknowledged that pilots are indispensable for enhancing safety due to their possibility to adapt to new situations (desired variability in human behaviour), at the same time they can be a threat, because it can also result in undesired variability in behaviour. A famous example of desired variability in crew behaviour is Flight 232 of United Airlines in which the crew faced a catastrophic engine failure causing all three of the aircraft's flight control systems to fail, making the aircraft almost uncontrollable (NTSB 1990). Although this situation was never expected to happen, the crew managed to control the aeroplane by using alternate thrust. A famous example of undesired crew behaviour, on the other hand, is Flight 173 of United Airlines in which the captain was busy managing a landing gear malfunction and preparing for a possible landing emergency (NTSB 1979). The captain did not monitor the aircraft's fuel state, resulting in fuel exhaustion to all engines. These two examples indicate that system safety could be enhanced if humans are able to recover instantaneously from undesired behaviour.

The majority of the studies aiming at enhancing system safety, define undesired variability (i.e. (decision) errors, failures) as a deviation from a standard process which increases the likelihood of an unsafe outcome (Lipshitz 1997). This definition is insufficient to identify instantaneously undesired behaviour. Although standard operating procedures are a thorough blue print of situations that have occurred or are expected to occur, there still exist situations where standard procedures do not satisfy. In these situations deviating from the standard procedures can be desired to avoid unsafe outcomes. Consequently, outcome knowledge is necessary to identify deviant behaviour as desired or undesired; knowledge we do not possess until afterwards.

To understand both desired and undesired variability in decision behaviour we need the criterion on which humans behave. The concept of autopoietic systems, developed by Maturana and Varela (1992), supplies us with the necessary basis for such a criterion. It adheres to the idea that living systems maintain themselves. The concept has been adapted to be able to explain human undesired behaviour.

This paper introduces in short the developed Maintaining Sensing Influence concept (Van der Vlugt and Wieringa 2004). It provides a criterion with which undesired behaviour can be explained. The concept assumes that in the context of the actor, he always decides (behaves) consistently, however, in the context of the observer this does not have to be the case. The concept has been tested with success in a student counselling set-

ting, resulting in an intervention method to establish a recovery from undesired behaviour. A field study in a full-scale simulator has been conducted to compare the cockpit setting with the student setting. Similarities and differences between the two settings are discussed, together with the difficulties to adapt the student interventions into aircrew interventions.

5.2 THE DIFFICULTY WITH DECISION-MAKING

Aviation is par excellence a domain of decision-making. Much effort has been put into developing standard operating procedures and accompanying checklists (Kaempf and Orasanu 1997), to cover all imaginable situations by standardized decision procedures. The standard procedures often simplify the situations to if-then decisions: if “x” occurs, then do “y.” However, there are situations as Kaempf and Orasanu sketched (1997) in which a straight application of the procedures is ambiguous, as shown by the following example. In the “rejecting take-off” procedure the criterion to continue or terminate takeoff safely is the airspeed V1, which is the prescribed maximum speed for safely stopping, calculated on aircraft type, weight, etc. If during acceleration a malfunction occurs below V1 then the aircrew should terminate the takeoff. If a malfunction occurs above V1, then the aircrew should continue. Yet, in practice more factors than accounted for in the criterion may come into play, such as the runway condition and wind conditions. In these situations the crew may deviate from the prescribed criterion using another criterion.

Another problem is that the standard procedures implicitly assume that humans are rational decision-makers, i.e. that humans decide on logical reasoning based on some clear objective criterion. Granted that humans are rational decision-makers, this implies that they are able to choose an objective criterion on which the pros and cons of the outcomes of the variety of decisions are weighed. However, also for choosing this criterion, another criterion is needed on which again the pros and cons of the possible criteria have to be weighed and so on. In other words, the assumption of rational decision-making leads to an infinite regression of deciding that never leads to actions. Nevertheless, it is known that humans decide all the time on some “final criterion.” But what is the “final criterion” the actor uses when the observer speaks of undesired (decision) behaviour?

5.3 Concept for explaining (undesired) decision-making

To be able to recover instantaneously from undesired behaviour, the final criterion on which humans decide should be identified. Therefore the conceptual framework of Maturana and Varela (1992) is introduced and adapted to human operator behaviour.

Concept of Maintaining Organization

A theory supplying us a final criterion for behaving is the theory of autopoiesis described by Maturana and Varela (1972). Autopoiesis stems from the Greek words auto = self and poiesis = creating. Maturana and Varela characterize autopoietic systems as systems that continually create themselves, i.e. autopoietic systems are living systems. Living systems are systems that continually maintain themselves.

When explaining living systems, it is important to distinguish between organization and structure:

- **Organization:** An ordered set of relationships characterizing the system. It makes the system a member of a specific class. Organization represents identity.
- **Structure:** The actual constitution of the system. It refers to the system's concrete components and their interrelations.

In case of living systems, the distinction between organization and structure is essential, because their structure is continually undergoing changes while their organization is maintained. The variety of behaviour of living systems relate to the ongoing structural changes. According to this theory, the essential process Maintenance of Organization (MO) is the "final criterion" on which living systems behave. Note that this criterion is not an external objective criterion (goal, reference), but a subjective one, because the system (actor) itself is subordinate to its Maintenance of Organization.

Essential to the concept of maintaining organization is the way living systems interact with their environment. By interacting recurrently with its environment a living system couples structurally to its environment. Structural coupling triggers structural changes in the system (Maturana and Varela 1972). When the structural coupling is with another living systems, it is called social coupling, which is of our particular interest.

Humans couple socially through valuing their experiences via "languaging." This process has been called social representing and has been defined as creating evident meanings (i.e. "valuing") in community (Zeegers 1988). In social representing social representations (social rules) emerge, which in turn affect social representing (Mattheeuws 1990). The process of social representing is an ongoing human dance of mutually valuing. It is impossible to escape from this dance. Whatever someone is doing, when sensed, it will automatically be valued. Or as Watzlawick (1967) would put it: "One cannot not behave."

Concept of Maintaining Sensing Influence (MSI)

In case of social representing, at least two parties are involved, the observer and the actor. Starting from the observer, the recursive process of social representing can be described as: valuing of the observer of the valuing of the actor of the valuing of the observer etc. Taking the actor's viewpoint three kinds of effects can be defined (see Chapter 3):

- **Influence**:= effect *of* the actor's valuing on the observer (or, more generally, the actor's influence on the environment)
- **Sensing Influence (SI)**:= effect *on* the actor of observer's valuing (or, more generally, the environment's response sensed by the actor)
- **Noting Influence (NI)**:= effect *on* the actor when he connects his expected effect of his influence with the observer's valuing (or, the matching by the actor between the expected effect of his influence *on* the environment and the actual effect *of* the environment's response sensed effect on the actor)

Sensing Influence (SI) is the triggering of structural changes "in the actor" by influences from the environment. All living systems sense continually influences from the environment.

Noting Influence (NI) is a specialized form of SI that is unique for humans. Humans can compare their expected effect of their influence on their environment with the actual sensed effect of their environment. For NI two extremes can be distinguished: firstly expected and sensed effect match: NI is optimal; secondly the expected and sensed effect mismatch: NI has lost, that is the actor does not sense any more the effect of his influence. When NI has lost SI automatically takes over.

The concept of Maintaining Sensing Influence (MSI) is one to one coupled with structural changes facilitating the essential process of Maintaining Organization. It assumes that the final criterion on which humans behave (decide) is their SI. This concept can be used to explain human behaviour as always consistent with regard to SI.

Maintaining Sensing Influence through context change

Whereas NI can be lost, SI cannot be lost; it is always maintained as long as the actor stays alive. When the actor's NI is optimal, desired behaviour is shown, that is, the expectations of the observer and the actor match. When the actor's NI has lost, he changes *autonomously* to a context in which his SI is maintained, consistent with the concept of maintaining organization. Whether the actor's behaviour is valued as undesired depends on the context of the observer. Undesired behaviour is thus not an objective fact, it is an observer's value and it only implies that the observer and the actor do not share the same context (see Maturana and Varela (1992) and Chapter 3). Behaviour can be valued to one of the following contexts:

- **Reference context**: the explicitly defined rules which are directives for acting in anticipated situations. In aviation these are the standard operating procedures but also for example the general received codes like the criminal code and traffic rules. The reference context by definition equals the situation for which it is established and therefore NI is maximal.
- **Situational context**: less explicit rules formed in a history of successful recurrent interactions (including unwritten social rules) that apply when the reference context does not satisfy any more because the situation is more complex than has been anticipated by the reference context. The NI is still high and therefore the likelihood of a safe outcome is high. This context represents those situations in which the operators take a well-considered decision based on other rules than the rules of the reference context.
- **Evident context**: unwritten social rules that apply when NI is lost in an actual situation. When NI has lost, the actor changes autonomously (intuitively) from either the foreseen reference context or situational context to a context of self-evidences, being the social representations. The actor's actions are determined by the unwritten social rules and do not relate anymore to the actual situation since this threatens his SI. This behaviour not necessarily has to be inadequate to meet the requirements of the situational context or the reference context. However, the likelihood of an unsafe outcome is considerable.

According to the concept of MSI, undesired behaviour implies that the actor has lost NI and maintains SI by changing context while the observer did not take into account the change of context of the actor.

HYPOTHESES

A number of hypotheses can be derived from the concept in which maintaining sensing influence (MSI) is used as essential process (the final criterion) for explaining undesired human behaviour:

1. Both actor and observer behave (i.e. “value acting”) according to their own context.
2. Inconsistent behaviour of an actor is an observer value; it implies that the actor and the observer do not share the same context.
3. Inconsistent behaviour indicates that the actor has lost NI in either the reference context or situational context and now optimizes his SI in an evident context.
4. When NI is lost, positive valuing effects are lost with it and only inconsistent behaviour will be experienced by the actor as valued; although it is negatively valued by the observer.
5. Inconsistent behaviour indicates that the actor does not know what to do to regain NI, that is, he does not know how to change to the reference or situational context.
6. Persistence of inconsistent behaviour reveals the inability of the observer to find or understand the context with which the actor’s behaviour is valued as desired.
7. Regaining NI will start when the actor senses a positive valuing when he admits openly that at this very moment he does not know what to do, although he should, considering his professional expertise.
8. When the observer values positively the actor’s failing to meet the expectations, it means that they share the same context again, the prerequisite for the actor to regain NI.

5.4 Implementation of the concept in the student counselling domain

Since implementing the model in the field of aviation was not possible, the model has been implemented in the field of student counselling. The study conducted was a qualitative study since the final criterion, maintaining SI, is a subjective criterion. In this study the student is the actor and the student adviser and the student counsellor are the observers. In this study the student adviser used the methods currently accepted to recover study progress. The student counsellor applied the MSI-concept to recover the student's desired behaviour.

Justification for student counselling

At first sight, the problem of students lagging behind seems to differ extremely with the pilots erring persistently. However, the essence of the problem of students lagging behind shows significant similarities with pilots maintaining their original plan or course of actions despite contradictory evidence. The difference between both domains is mainly context specific.

In aviation, through thorough selection, education, and training, crews are considered to be qualified to carry out successfully the tasks that are required for operating an aeroplane safely. In academic education, also through thorough selection, education and training, students are considered to be qualified to carry out successfully the study programme that is required to obtain a masters degree (or specifically for the TU Delft: to become an engineer.) Viewed in this perspective the "becoming an engineer" decision, like the takeoff decision in aviation, seems comparable unambiguous: the student should stop when his study progress is insufficient, or the student should continue when his study progress is sufficient. Like in aviation the student's decision situation is much more complex in practice. Two complicating factors in this seemingly unambiguous decision situation are, that students are given the opportunity to take their exams several times and that a lack of study progress does not influence their possibility to continue to follow the study programme as a student. This means that the university system allows their students to show persistent erroneous behaviour, that is, to persist in lagging behind. In other words, the student may stick to his original plan to become an engineer despite the contradictory evidence that he will not succeed when continuing in this manner.

When students are confronted with a lack of study progress despite their qualifications and their ongoing efforts to regain study progress, they either are contacted by the student adviser of their department, or they contact their adviser themselves, to discuss which measures can be taken for regaining study progress. When the measures suggested by the student adviser are not successful, they are advised to consult a student counsellor at university level. In these cases it is assumed that the interventions methods used by the student adviser did not satisfy to regain study progress or to facilitate the decision to abort their study.

MSI-BASED ANALYSIS OF THE LACK OF STUDY PROGRESS THAT PERSISTS

Students who consult the student counsellor have no idea about the cause of their lack of study progress or present the presumed cause of their lack of study progress also as the solution. As cause of insufficient study progress students frequently mention their loss of motivation. The student counsellor's first step is an analysis of the persistence of the student's problem by asking the student if he possibly reasons as follows: "My problem is that I do not make study progress because I have no motivation, thus my problem is solved when I regain my motivation, but I only can regain motivation when I make study progress." When he agrees, the next step is to show the student that in this reasoning loss of motivation has become an effect instead of a cause and that when the effect is considered to be the cause there is the danger of circular reasoning.

Next to study related behaviour, students mention that they, for example, sleep late, regularly drink coffee, go to the pub, visit their parents every weekend, etc. Viewed from a situational study context this behaviour is undesired *but only* when study progress falls short of expectations. However, viewed from the evident context of being a companionable house-mate, being socially involved by going to the pub with your house-mates, and being a good son or daughter by visiting your parents frequently, it is desired behaviour. To regain study progress, the student adviser told the students to drink coffee with their house-mates less often and go to the pub less frequent, because this behaviour was considered to be the cause of their stagnating study progress. This is advice that failed in the case of students that consulted the student counsellor, because the behaviour patterns are effects and not causes.

When a student adviser agrees with the way a student presents his problem, as circular causation of which causes and solutions are reversible, it is not amazing that regaining study progress cannot be guaranteed. When the student is then referred to a student

counsellor, it also is not amazing that the student presents his problem in the same way or even in a more hopeless way and with even lesser faith that it can and will be solved. For solving these problems a more sophisticated intervention method is required.

MSI TO EXPLAIN UNDESIRED BEHAVIOUR

For explaining undesired behaviour, the first working hypothesis is that the study context used to value the student's problems did not result in interventions that led to a regain of study progress. Whatever the student did, for example, making a study plan or starting early, nothing met his expectations; consequently the student lost NI. Following the MSI-concept, losing NI implies the student changes context to maintain SI. In this new context the social activities are valued. Regarding this context the student shows desired behaviour. However, when the adviser does not take into account this context, he values the student's behaviour as undesired.

The question to be answered is: why did the student change context, making the formerly successful student to behave in an undesired way? Based on the MSI-concept, the following explanation for the student's behaviour can be given. A formerly successful student who behaves undesirably does not note his influence any more, that is, he does not sense any valuing of the environment to his efforts to maintain study progress. In the situational study context he values his successful progress as self-evident considering his capabilities. When students value their efforts as self-evident, that is, as unnoticed, they lose their NI and autonomously change to an evident context that maintains their organization by optimizing SI. Through behaving in an evident context their efforts will be noticed, that is valued, again. Although the value is (perceived as) negative, it is a value linked to the student's behaviour and the only thing left when success becomes self-evident.

When a student shows undesired study behaviour many people are invited to value it and many do so: parents, friends, student adviser and lecturer. When modest forms of undesired study behaviour do not provoke sufficient negative valuing, students have to intensify their efforts by showing more extensive forms of undesired behaviour. The MSI-concept relates escalation of undesired behaviour with the ongoing inability of the observer to find, understand and communicate the context with which he can value this behaviour as appropriate for meeting the self-evidences ruling in that context.

The problem sketched is related to the student's history of success and relates to the idea of the observer that formerly successful students can show appropriate behaviour in certain circumstances, because they have showed it before. For students at TU Delft who were formerly successful, it is self-evident that they can start studying at the right time. However, when they actually cannot start, it guarantees loss of NI.

Regaining NI, MSI based interventions

To regain NI, there is one thing a student can successfully do in the situational context. A student who consults a student counsellor senses that he does not know what to do to master his problem. When asked, the student will say as expected something like "That is correct, I really do not know what to do. I tried it all, but nothing worked". Based on the MSI-concept, it was hypothesized that the student is back in the situational context and regained his NI, because, consistent with the MSI-concept, maximum NI relates to a congruence of expected and experienced effects within a reference context or situational context. The student achieves a congruence of expected and experienced effects in the situational context by saying: "I do not know what to do to master my problems." Because when you do not know what to do, you might expect an absence of desired goals. Moreover, the student has succeeded in asking the student counsellor if he knows what to do to regain NI.

Then the student counsellor rewards the student for his initiative to come and consult a professional for a problem the student does not know the cause of nor the solution for. The rewarding is essential because in this field of human behaviour people are confronted with multiple suggestions that they should know the cause and the solution of their problems.

Results using the MSI-concept

After regaining NI, specific methodical interventions (which do *not* contain behaviour proposals the student has already performed before without success) were introduced through which students could start applying their capabilities again to tasks which were directly related to realizing their desired goals (De Booij 2003). Over the last two years 43 students were involved in the study. In a maximum of four sessions with the counsellor the students recovered desired behaviour. In accordance with the concept of MSI the assumption is that as long as the sessions are not successful the counsellor has failed to find and communicate the context in which the student maintains SI. Although in most

cases the problem was solved within the first two sessions (each of approximately two hours), some of the students felt uncertain with the result. For those students two or three extra sessions were offered.

The nature of the MSI-method implies that no hard evidence can be given for its success. The fact is, that students referred by their student adviser to the student counsellor regained study progress, sometimes after a stagnation of more than a year. Before the availability of the MSI-method, the student counsellor involved has had the experience of coaching students who persisted in their behaviour. He counselled, for example, a student who took a whole year to make a planning for completing his MSc-programme, but the counsellor could not establish that the student took the exams planned. Disappointed, the student stopped consulting the student counsellor.

With two particular cases the added value of the MSI-method is illustrated. The first case concerns a student who consulted the student counsellor because, after a long period of unsuccessful efforts to regain his study progress, he decided to get a job, with only the final year of his master's programme to go. It appeared that the student did not know anymore what to do to complete it. His solution was to find a job where he could prove his capacities gained during his student life. The student counsellor confronted the student with his opinion that he was convinced that the student was going to obtain his degree instead of getting a job. He proposed the student to execute the essential MSI-based intervention for recovering his formerly successful study behaviour. After doing so the student indeed continued his master's programme and completed it with success. The second case concerns a student who started to consult the student counsellor from her first year at TU Delft, with intervals of about a year. After four years a real problem emerged, the student nearly always scored a mark lower than was required. When a 6.0 was required, she scored 5.8, when a 7.5 was required she scored 7.0. It was agreed that her method of studying should satisfy for obtaining at least the required mark. Therefore, the student had no idea what else to do to obtain the required marks. When the student counsellor suggested the MSI-method as a solution, the student explained that she not able to do it, due to personal reasons. The student counsellor respected this, but told her that she most likely would not regain study progress. When, in addition her physical condition declined seriously, her parents advised her to abort her study. Although physically limited, the student persisted in taking exams, without obtaining the required marks. One day she told the student counsellor, she would carry out the MSI-method. In the next academic year she completed 125% of the annual required study progress and in the end she obtained her master's degree.

The success of the sessions based on the MSI-concept can to a large extent be ascribed to the benefit of the way student counselling is carried out at TU Delft. Student counsellors at TU Delft only get started when consulted. The students decide if and when they visit a student counsellor. By passing the doorstep of the counsellor's room students express non verbally that they do not know what to do to master their problem, which implies that half the work is already done. Note that although the student's initiative to consult a professional for his problem is a necessary condition for regaining NI, it is not a sufficient condition since the department advisers, who did not use the MSI-concept, could not establish a regain of NI. This means that, when the student has taken the first step to regain his NI, the counsellor needs an additional set of interventions that are based on the MSI-concept to lead the student back to the situational context in which the regained NI is maintained.

5.5 Student counselling versus aircrew interventions

Although the set of interventions developed for the student context proved to be successful, an additional set of interventions suitable for the cockpit setting has to be developed. The student setting and cockpit setting differ in several ways. To examine similarities and differences in the decision-making process between both settings, an observational study was conducted. Seven full flight mission simulator sessions were observed and video-taped to analyse the aircrew decision-making processes.

Reducing the decision-making process to its essence, it was concluded that for both settings the basic decision process is the same, that is, they are both based on "if-then" rules. The reference context of the cockpit setting (standard operating procedures), is much more explicit and elaborately defined than the reference context of the student setting. The effect of this difference is that many more students show undesired behaviour (and "crash") than aircrews do.

In the student setting, the main factor influencing behaviour is considered to be motivation, whereas, in the cockpit setting the main factor is the aircrew's professionalism. As argued before, motivation is not a cause of study progress but an effect. This applies to "professionalism" as well. However, professionalism has extra complicating self-evident meanings. Being a professional implies self-evidently always knowing what to do, because they are employed to know. It is therefore not self-evident at all that professionals are able to admit openly that they do not know what to do when they are supposed

to know. Furthermore, in aviation (and other professions) professionalism of aircrews is related to acting conform the determined standards that maximize passenger safety. Professionalism of the aircrew (self-evidently!) is considered to be essential for safe operations, whereas students can permit losing NI in their study context, because they do not jeopardize immediately others well being. Students are given time and opportunities to discover the crucial step to regain NI themselves, pilots are not, since the consequences can be catastrophic.

When a pilot, who is expected to know what to do to recover from critical situations, is able to evaluate his complex situation as too complex for him to handle this very moment, he is behaving maximal professional. Instead of losing his NI and changing to undesired behaviour, he now can ask for advice and thereby maintains his NI. This view on professionalism is not commonly accepted in aviation and, therefore, acceptance and implementation of this view may be a crucial step to further enhance safety in aviation.

5.6 Significance of the MSI-concept

In case other interventions did not lead to the expected results, the two main advantages of the MSI-concept are:

1. it provides an explanation for undesired behaviour additional to the existing explanations such as problems with attention, and loss of Situation Awareness,
2. it provides starting points to develop an intervention strategy for recovery from undesired behaviour.

Loss of NI is a mental construct within the MSI-concept, which is defined in a way which seems to relate with loss of Situation Awareness (SA) as introduced by Endsley (1995). According to Endsley, SA is the result of complex information processing. It is the actual result achieved by an operator assessed in terms of an established result desired by an observer. SA is ascribed to an actor by an observer and therefore observer related. Noting Influence (NI) on the other hand is actor related. It is based on the process of Maintaining Sensing Influence (MSI) instead of on an observer-based desired result, and accounts for all situations where expectations and experiences match. When expected and experienced effect mismatch, that is when an actor has lost NI, MSI takes over through context change and ‘undesired’ behaviour optimizing SI. In other words loss of SA implies that the context of the actor mismatches with the context of the observer. Loss of SA only refers to actor behaviour which is not shown (as experienced

by the observer), but should be shown (as expected by the observer). The construct SA provides no explanation for the actual shown and undesired behaviour. A commonly mentioned explanation for loss of SA is stress. Stress, as does SA, refers to a result of complex information processing and does not provide an explanation since stress is a behavioural manifestation as well. However, within the MSI-concept stress can be an explanation for undesired behaviour, through loss of NI, but only when stress is viewed as the result of a mismatch between expected and experienced effects.

What is discussed for the relation of an observer with NI and an actor without NI, is identically applicable to the relation of an observer with NI who is observing an observer who has lost NI observing an actor, without or with NI.

The main limitation of the MSI-concept is that it cannot predict *when* undesired behaviour will occur because it adheres to autonomy, literally meaning self-ruling, of humans. The "self-rule" (i.e. process) where humans are continually subordinate to their maintenance of SI.

Another limitation of the MSI-concept is that it does not provide a ready-made "if-then" recipe to realize a recovery from undesired behaviour. The MSI-concept provides an explanation for undesired behaviour when it has occurred from which interventions can be derived to establish a recovery from it. Depending on the professional area (e.g. cockpit, student counselling, anaesthesia, etc.) this is a starting point to develop area specific interventions.

5.7 CONCLUSIONS

To realize an instantaneous recovery from undesired decision behaviour, we have introduced a concept that supplies us with the "final criterion" on which humans behave (decide). The concept is based on the essential process of Maintaining Sensing Influence (MSI) in which SI is the effect of the influence of the observer (environment) on the actor. The MSI-concept comes into play when undesired behaviour occurs unforeseen and when the actual situation cannot give a reason for it. The MSI-concept serves as the last defence layer of the system to enhance system safety.

The concept has been used to develop an intervention technique for student counselling at Delft University of Technology. Although the decision process of students is in essence similar to the decision process of aircrews, there are two essential differences as well.

The first difference is the time span for interventions. In the cockpit setting the time span is only a fraction of that in the student counselling setting. The second difference is the consequence of undesired behaviour of students versus aircrew. Undesired behaviour of students is rarely life threatening for their environment whereas undesired behaviour of an aircrew can be fatal for the passengers and for themselves. In practice this means that the student is given the time and opportunity to experience the effect of their undesired behaviour and to discover the effect of establishing a recovery from it based on the MSI-concept handed by the student counsellor. For an aircrew the process of recovery from undesired behaviour should preferably be experienced in an additional training programme.

REFERENCES

De Booij JWP (2003) Studeren doe je niet alleen, de communicatieve betekenis van studeerproblemen. *Systeemtheoretisch Bulletin.* **21(3):** 91-102.

Endsley MR (1995) Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors,* **37(1)**: 32-64.

Kaempf GL, Orasanu J (1997) Current and Future Applications of Naturalistic Decision Making in Aviation. In: Zsambok CE and Klein GA (eds.) *Naturalistic Decision Making* Mahwah, New Jersey: Lawrence Erlbaum Associates Inc.. pp. 81-90.

Lipshitz R (1997) Naturalistic Decision Making Perspectives on Decision Errors. In: Zsambok CE and Klein GA (eds.) *Naturalistic Decision Making* Mahwah, New Jersey: Lawrence Erlbaum Associates Inc. pp. 151-160.

Mattheeuws A (1990) Het spel der vanzelfsprekendheden. *Systeemtheoretisch Bulletin*, **8(4):** 255-287.

Maturana HR, Varela FJ (1980) *Autopoiesis and Cognition; The Realization of the Living*. Dordrecht, The Netherlands: D. Reidel; Publishing Company.

Maturana HR, Varela FJ (1992) *The Tree of Knowledge; The Biological Roots of Human Understanding*. Boston, Massachusetts: Shambhala Publications.

National Transportation and Safety Board (1979) *United Airlines Flight 173*, *McDonnell Douglas DC-8-61, N8082U, Portland Oregon.* NTSB/AAR-79/07. Washington DC, NTSB.

National Transportation and Safety Board (1990) *United Airlines Flight 232, McDonnell*

Douglas DC-10-10. NTSB/AAR-90/06. Washington DC, NTSB.

Van der Vlugt M, De Booij JWP, Dankelman J, and Wieringa PA (2004) Training Aircrews to recover from Human Error, a Proposal for a Training Outline. In: *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (*pp. 188-193). The Hague, The Netherlands.

Watzlawick P, Beavin JH, Jackson DD (1967) *Pragmatics of Human Communications*. New York: Norton & Company, Inc.

Zeegers W (1988) *Andere tijden, andere mensen, de sociale representatie van identiteit*. Amsterdam, The Netherlands: Bert Bakker.

Chapter 6

TRAINING AIRCREWS TO RECOVER FROM HUMAN ERROR, A PROPOSAL FOR A TRAINING OUTLINE

In aviation the crew's decision behaviour is critical. General acceptance that undesired behaviour (human error) is inevitable in the end has led to the acknowledgement that safety could be enhanced when a recovery from undesired behaviour can be realized. The current training programs do not satisfy this purpose due to the danger of negative training. This paper proposes an outline for a training that assists the crew to recognize undesired behaviour and to establish a recovery from it. For the outline the concept of Maintaining Sensing Influence (MSI) is used. It assumes that whatever a pilot does, he is continually guided by the essential process of MSI. According to the MSI concept undesired behaviour implies that the pilot changed context to maintain his SI, while the instructor did not take account of the context change. Undesired behaviour is not an objective fact, it is a value of the instructor implying that the instructor and pilot do not share the same context. Having identified the change in context, MSI becomes a criterion by which the instructor can understand the pilot's undesired behaviour and can assist the pilot to recover desired behaviour.

Keywords: *aviation, aircrew training, decision-making, recovery from human error, Maintaining Sensing Influence*

This chapter appeared in:
Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (pp. 2574-2579). The Hague, The Netherlands. 2004.

The following terms have been changed to bring them in line with the terms that are currently used in this thesis:

Context shift/switch:	Context change
Critical variable MSI:	Essential process MSI
Frame of Reference (FoR):	Reference context

6.1 INTRODUCTION

According to Wiegmann and Shappell, safety measures aiming at technical system errors have been much more effective than those aiming at reducing human error. Research into aviation accidents pointed out that approximately 80% of the accidents are attributable to aircrew errors (Wiegmann and Shappell 2001). NASA observed that aircrews were more successful at managing technical error than managing human error (FAA 1998). Many training programs to date have emphasized training for a high technical proficiency. It is therefore not surprising that the rate of accidents attributable to system error has declined much more than the rate attributable to human error, since pilots serve as a backup system defence in case the system fails. But who or what backs up the human defence when it errs?

For a further reduction of the accident rate, we should aim at reducing the accident rate attributed to human (pilot) error since pilots are essential to a safe outcome. The consensus is that we not only must aim at human error prevention but at error recovery as well. Not only because it is generally acknowledged that human error is inevitable in the end but also because in many cases after a human error occurred there are still opportunities to establish a recovery from it (Naikar and Saunders 2002).

Without the intention to blame aviation accidents merely on pilot error, for it is well established that accidents result from a combination of circumstances, we still would propose to focus on aircrew decision-making. It is acknowledged that aircrew decision-making is a critical component in aviation and that it could and should be trained. Training of aircrew decision-making has become commonplace and proved to have a positive impact on reducing aircrew decision errors. Nevertheless, it is also recognized that the present training programs have limitations with regard to recovery from human error and still give room for advancement, since they are based on classical decision strategies and emphasize the prevention of the occurrence of human errors in the first place (Kaempf and Orasanu 1997; Naikar and Saunders 2002).

In this paper an outline is introduced to develop a training approach that assists aircrews to recognize their decision error (i.e. human error) and establish a recovery from it. The approach outlined is based on a model which assumes that the essential process of humans is maintenance of organization (Capra 1996; Maturana and Varela 1972, 1992). According to this model all human (decision) behaviour is guided by Maintaining Sensing Influence (MSI). The training approach proposed has been based on a MSI-based in-

tervention technique that has been specifically developed for student counselling (Van der Vlugt et al. 2004). To identify the needs for aircrew training several training sessions were observed. The findings are integrated in a proposal for a training outline.

6.2 PRESENT AIRCREW TRAINING

There is general acceptance that humans can threaten system safety due to their proneness to make errors which can lead to critical situations. Nevertheless, they are indispensable to enhance system safety as well due to their ability to recover from critical situations. Acceptance that human error is inevitable in the end (Reason 2000), has led to the acknowledgement that for a further enhancement of safety, aircrews should be trained in recognizing human error and recovering from human error.

In theory, the ideal training should be a simulator session that provokes the crews to make errors with which they cross the boundaries of safe operation and that assists the crews to discover ways to recover from these crew errors (Naikar and Saunders 2002). In practice commercial airliners turn out to be rather reserved to incorporate such a training in their training program because of the danger of negative training. Negative training means that the crew has a bad experience due to the course of the training (for example a crash or a near crash) which shatters the confidence of the crew. This reservation is related to the experience of instructors that such a training approach to date has nothing to offer on what happens with the crew when it errs and does not find a way to recover from it. It is therefore not surprising that current training programs aim at training the crew to prevent the occurrence of an error in the first place.

For crews to train how to deal with decision errors, the Crew Resource Management (CRM) approach has been used frequently to date. The basis for the CRM approach is the crew-based error model of Helmreich (see Figure 6.1). It describes a taxonomy of crew-based errors in which a crew-based error has been defined as “an (in)action that leads to a deviation from organizational expectations” (Helmreich *et al.* 2001). The organizational expectations can be formulated in short as maximizing the probability towards a safe and efficient flight. Although training programs based on the CRM approach have proven to be effective to enhance safety (FAA 1998), they are limited for training of recovery from human errors since CRM tends to focus on the error manifestation and its evolution. The crew-based error model used in CRM also does not explain why a crew ends up in an error loop nor why it breaks through this error loop nor why it avoids

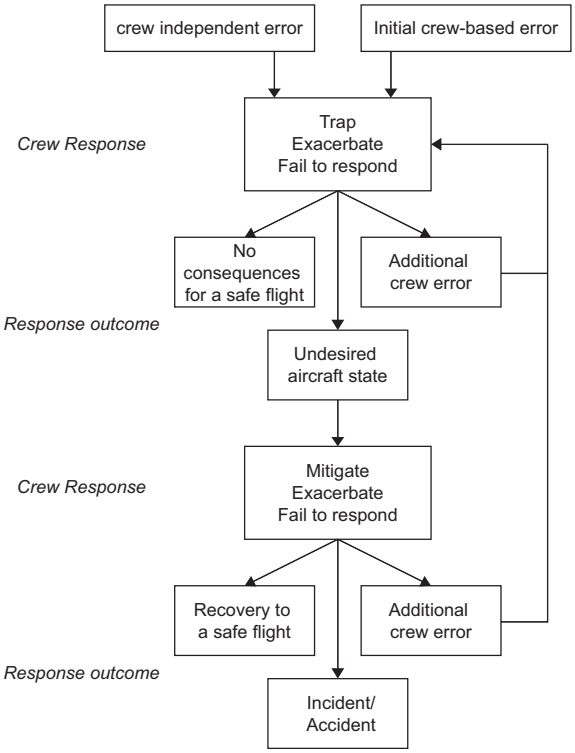


Figure 6.1 *Crew-based Error Model adapted from Helmreich (2001). A Crew-based error is defined as a crew (in)action that leads to a deviation from organizational expectations.*

the error loop at all. In other words, the model does not give the final criterion according to which the crew behaves. At first sight the criterion seems to be aiming for a safe flight. However, if this is the criterion, what makes crews deviate from this aim, this criterion? These deviations cannot be explained by the deficiencies of the crew because such an explanation tells how the crew does not act and very little about how it actually acts.

The objective of this paper is to explore issues for developing an additional training approach to redirect the training efforts so that it addresses particularly recovery from undesired (decision) behaviour. Therefore the essential process is necessary which accounts for all human behaviour. Undesired behaviour actually implies that it does not satisfy the criterion defining desired behaviour. The training outline that will be proposed

in this paper uses a concept that assumes that all human behaviour is subordinate to the process of Maintaining Sensing Influence (MSI), that is, to human survival, how undesired this behaviour may seem with regard to system safety (Van der Vlugt *et al.* 2004. The MSI-criterion enables the instructor to let the crew err and train to recognize their undesired (decision) behaviour and to find a way to recover from it. With the MSI-criterion negative training can be avoided although the crew crosses the boundaries of safe operation.

6.3 Model used for developing the training outline

The model we will be using to develop the training approach adheres to the viewpoint that humans are faultless systems of which their essential process is maintenance of organization. All human behaviour is subordinate to the process of maintaining organization by Maintaining Sensing Influence (MSI).

Concept of Maintaining Sensing Influence

In case of human behaviour, at least two parties are involved, that is, the observer (i.e. the instructor) and the actor (i.e. the pilot). Starting from the observer, the recursive process of humans behaving can be described as: valuing of the observer of the valuing of the actor of the valuing of the observer etc. Taking the actor's viewpoint three kinds of effects can be defined:

- **Influence**:= effect *of* the actor's valuing on the observer (or, more generally, the actor's influence on the environment)
- **Sensing Influence (SI)**:= effect *on* the actor of observer's valuing (or, more generally, the environment's response sensed by the actor)
- **Noting Influence (NI)**:= effect *on* the actor when he connects his expected effect of his influence with the observer's valuing (or, the matching by the actor between the expected effect of his influence on the environment and the actual effect of the environment's response sensed effect on the actor)

Sensing Influence (SI) is the triggering of structural changes "in the actor" by influences from the environment. All living systems sense continually influences from the environment.

Noting Influence (NI) is a special form of SI that is unique for humans. Humans compare the effect they expect of their influence on their environment with the effect they actually sense of their environment. For NI two extremes can be distinguished: firstly, the expected and sensed effect match: NI is optimal; secondly the expected and sensed effect mismatch: NI is lost, that is the actor does not sense any more the effect of his influence. When NI is lost SI automatically takes over.

The concept of Maintaining Sensing Influence (MSI) is directly coupled with the structural changes facilitating the process of maintaining organization. It assumes that the essential process to which humans behave (decide) is Maintaining SI. This concept is used to explain human behaviour as always consistent with regard to maintaining SI.

Maintaining SI through context change

Whereas NI can be lost, SI cannot be lost; it is continually maintained as long as the actor stays alive. When the actor's NI is optimal, desired behaviour is shown, that is, the expectations of the observer and the actor match. When the actor has lost his NI, he changes *autonomously* his context in which his SI is maintained, consistent with the concept of maintaining organization.

Whether the actor's behaviour is valued as undesired depends on the context taken by the observer. Undesired behaviour is thus not an objective fact, it is an observer's value and it only implies that the observer and the actor do not share the same context (Maturana and Varela 1992). See also Chapter 3. Behaviour can be valued to the following contexts:

- **Reference context:** the explicitly defined rules which are directives for acting in anticipated situations. In aviation these are the standard operating procedures. The reference context by definition equals the situation for which it is established and therefore NI is maximal.
- **Situational context**: less explicit rules formed in a history of successful recurrent interactions (including unwritten social rules) that apply when the reference context does not satisfy any more because the situation is more complex than has been anticipated by the reference context. The NI is still high and therefore the likelihood of a safe outcome is high. This context represents those situations in which the actors take a well-considered decision based on other rules than the rules of the reference context.

- **Evident context:** unwritten social rules that apply when NI is lost in an actual situation. When NI has been lost, the actor changes autonomously (intuitively) from either the foreseen reference context or situational context to a context of self-evidences. The actor's actions are determined by the unwritten social rules and do not relate anymore to the actual situation since this threatens his SI. This behaviour not necessarily has to be inadequate to meet the requirements of the situational context or the reference context. However, the likelihood of an unsafe outcome is considerable.

Following the concept of MSI, undesired behaviour implies that the actor has lost NI in the reference context or situational context and maintains SI by changing to an evident context while the observer does not take into account the actor's change of context.

For example, in according to the Federal Aviation Administration (FAA) a stabilized approach is the safest profile in all (reference context) but special cases (situational context), and is essential to a safe approach and landing. A stabilized approach is characterized by a constant angle (3 degrees) and a constant rate of descent (not greater than 1000 feet per minute (fpm).) Following the reference context a flight in visual conditions should be stabilized at 500 feet height above touchdown (FAA 2000). In other words, for a stabilized approach, there should be no warnings. A sink rate warning below 500 feet for example indicates that the rate of descent is greater than 1000 fpm. Disregarding a sink rate warning below 500 feet is undesired according to the reference context, because it indicates that the plane is beyond the boundaries of a stabilized approach. In special cases (the situational context) disregarding the warning is only desired, when it is expected and prepared in a special approach briefing. Disregarding the warning unprepared implies that at least one of the crew members behaves according to an evident context at that moment.

6.4 HYPOTHESES FOR MSI-CONCEPT

A number of hypotheses have been derived from the concept that assumes that maintaining sensing influence (MSI) is the essential process (the final criterion) for explaining "undesired" human behaviour:

1. Both actor and observer behave (i.e. "value acting") according to their own context.

2. Undesired behaviour of an actor is an observer value; it implies that the actor and the observer do not share the same context.
3. Undesired behaviour indicates that the actor has lost NI in either the reference context or situational context and now optimizes his SI in an evident context.
4. When NI is lost, positive valuing effects are lost with it and only undesired behaviour will be experienced by the actor as valued; although it is negatively valued (by the observer)
5. Undesired behaviour indicates that the actor does not know what to do to regain NI, that is, he does not know how to change back to the reference context or situational context.
6. Persistence of undesired behaviour reveals the inability of the observer to find or understand the context with which the actor's behaviour is valued as desired.
7. Regaining NI will start when the actor senses a positive valuing when he admits openly that at this very moment he does not know what to do, although he is supposed to, considering his professional expertise.
8. When the observer values positively the actor's failing to meet the expectations, it means that they share the same context again, the prerequisite for the actor to regain NI.

6.5 IDENTIFYING NEEDS TO TRAIN RECOVERY FROM HUMAN ERROR

For the student counselling setting an intervention strategy has been developed based on the concept of MSI that assists the students to recover from their undesired behaviour (Van der Vlugt *et al.* 2004). The student counselling study showed that for a recovery from "*persistent* undesired" behaviour, it was necessary for the student to experience himself his "undesired" behaviour, that is, his loss of NI, and to experience the successful way to regain NI. Although it is generally known that undesired behaviour cannot be prevented, *persistent* undesired behaviour can be averted when a recovery from undesired behaviour can be established. Since it is irresponsible in aviation, with regard to the safety, to let the crew experience a loss of NI in an actual flight, we have to turn to a training program that evokes a loss of NI of the crew.

Maintaining NI is essential for positive training

To assess the state of the art of the current training programs two types of aircrew training programs were attended, the type qualification program for newly qualified commercial pilots and the compulsory type recurrent program for commercial airline pilots. Starting point for the present programs is that crews should gain a positive experience of the training. A positive experience will be affected by both the scenario and the instructor's interventions.

A scenario which is too easy most likely will not lead to decision errors and will not challenge the pilots. The program is then experienced by the crews as a bit dull and a waste of (expensive simulator) time. A scenario that is too difficult on the other hand, can lead to too many decision errors, which can leave the pilots with a negative experience of training. However, a (too) difficult scenario in which the crew makes decision errors has not necessarily to be experienced as negative. The experience of the crew depends largely on the interventions of the instructor, for the instructor supervises the crew flying the aeroplane. When the instructor is able to explain in all circumstances what the crew did or how they decided, crews will experience their training as positive, i.e. the instructor is able to assist the crew to learn from their decision errors.

To prevent negative experience of a training it is therefore crucial that the aircrew's decision behaviour stays "within control" of the instructor. In other words, the instructor should always be able to explain to the crew what they did, how undesired this decision behaviour might be regarding the context of safe operation. Following the MSI-concept, to "stay in control" implies that the instructor should maintain his NI since he is subordinate to the same process of MSI as the crew is; only when the instructor's NI is optimal, he is able to explain the crew's behaviour.

An example of negative training

The importance of maintaining NI of the instructor for the crew to experience a positive training is illustrated by two particular examples of instructor interventions that occurred during a type-qualification training program. Two "aspirant" aircrews in the same stage of their program were challenged to operate on the maximum of their abilities. In theory they should know how to handle the situation they faced, since they passed their theory exam and they were briefed before the session what to expect and what to do. The practice in the simulator proved to be more unruly. The situation the crews faced required

action "x," if action "x" was not taken in time (a few minutes) the situation would deteriorate steadily due to an accumulation of related system errors. It appeared that both crews were not able to take action "x" in time. The first instructor let the situation deteriorate and watched how the crew coped with it. The crew managed to solve many of the problems, although not totally satisfactorily to return to safe operation. When it was clear the crew was not able to return to safe operation, the instructor "froze" the simulator, discussed the situation with the crew and what they did, complemented them on what they solved and discussed what they experienced. After this interruption the instructor repositioned the simulator and let the crew do the situation all over again. In contrast, the second instructor tried to intervene instantaneously when he realized that the crew would not take action "x" in time. He first advised, then instructed, and even commanded the crew to take action "x" by pushing button "a1, then a2 and then a3." The tension built up very quickly, resulting in more confusion by the crew. In the end the instructor did not succeed to get the crew to take action "x." Just before the plane was to crash the instructor "froze" the simulator.

The first crew experienced a positive training although they were challenged to operate at and even beyond the maximum of their capabilities during which they trusted the instructor completely. The second crew on the other hand, experienced a negative training and had to regain their shattered confidence. Both crews lost their NI and changed from the reference context to an evident context that maintained their SI. However, the difference between a positive or a negative experience was made by the interventions of the instructor. In the second case not only did the crew lose NI but the instructor did lose his NI as well. The interventions of the second instructor were still reference context based. The crew did therefore not respond as the instructor expected, since he expected the crew to behave as if they were performing in the reference context. The instructor's expected effect of his interventions and experienced effect mismatched, he lost NI and changed to a context of self-evidences just like the crew did.

6.6 Outline for training for recovery

The observations showed that, for the pilot to learn to recover from decision errors, the flight scenario is as important as the instructor's interventions. In practice it turned out that in most cases both the scenario and the instructor's interventions were on the safe side, because going too far beyond the boundaries of safe operation is judged as possibly threatening for a positive experience of the aircrew.

Additional to existing training programs, aircrews should also be exposed to scenario's which are designed to evoke procedural deviations. For the instructor such deviations indicate that an aircrew is (possibly) challenged to operate on the limits of its professional capabilities. When a procedural deviation occurs two situations can arise: first, the procedural deviation of the one pilot is corrected successfully by the other; second, the procedural deviation is not corrected and can lead to a point of no return in safe operation. This last situation is the focus of our interest of our training approach. This training approach would require the instructor to leave this error uncorrected to give the crew a further opportunity to learn to recognize that they do not note the effect of their influence any more (they lost NI) and changed context, and to experience the evolving situation.

For the instructor, the challenge is to identify the context change and to intervene such that the pilot regains his NI in the situational context. For these purposes the MSI-concept provides the instructor with a conceptual framework with which he can explain the (decision) behaviour of the crew. Firstly, the MSI-concept provides the instructor with indisputable clues for the identification of a context change. Secondly, he knows what crew members are up to, that is, Maintaining Sensing Influence by changing context, and thirdly the MSI-model provides him with intervention techniques which he can apply to let the crew recover from their context change.

When debriefing the session the instructor should focus on the context change, that is, the deviation from the expected course of actions during the session and the evolving situation. The session should always be video-taped because audiovisual feedback is considered to be an effective means for crews to learn from their behaviour, since it allows them to view themselves from another perspective. Feedback of crews that received a tape of their session we attended were positive in this respect. In practice nevertheless, audiovisual feedback is scarcely used by the companies we visited.

When looking back at the moment of context change the instructor should firstly try to evoke a free recall of the crew's experience at the moment of context change. What and how did they decide and was it desired within the actual situation. Then the next question for the crew is whether they can imagine other situation(s) in which this decision would have been desired? If they cannot it is important that the instructor can offer them one or more situations in which this decision would have been desired.

To discuss what happened, the starting point to bear in mind at all times for the instructor is that the pilot was momentarily not able to do what he should have been able to do

due to his professionalism. After all, if he would have been able to act as expected, he would have acted as such. It is fundamentally wrong for the instructor to ask why the pilot *could not* do the prescribed actions. Especially in cases where things did go wrong, a 'why question' is likely to impose a moral context.

A constructive question has to be context related and the instructor has to know the right answer, otherwise teaching the crew from what happened is impossible. The question for the instructor to ask is if the pilot can agree with the instructor's opinion that the pilot did not know what to do at that very moment, based on the reference context. As soon as the pilot agrees, which is unavoidable finally, the instructor assures the pilot that in such cases the most professional and desired thing to do is to say to relevant others: "at this very moment I don't know what to do". Firstly, this statement is correct about the pilot functioning in the reference context respectively situational context, so his NI is optimized and secondly this statement is a prerequisite for being able to perform as expected.

When pilots deviate from the procedures (reference context) they have changed to a situational or a self-evident context. Behaviour following the context change will show which context is dominant. Although the behaviour preceding the context change is of main interest for training recovery, behaviour following a context change, in particular after a change to an evident context, is important to show the pilot how a situation can evolve to an incident or accident when he does not recognize his context change (audiovisual feedback: seeing is believing!). For instruction purposes, the pilot behaviour in the evident context, can be evaluated by linking it to other situations where it would have been desired. By this "contextualizing," the pilot understands his undesired behaviour and can experience a positive rewarding influence from his undesired behaviour. Moreover, by having him supplied with the possibility to perform the desired behaviour next time, he will profit optimal from his training.

6.7 Future implementation

To enhance system safety further, training for recovery from human error should be addressed more directly. Part of the training required to meet this goal may be accomplished without the development of additional training resources. For example, by briefing and debriefing the existing training sessions with (anecdotal) cases may open discussions about the value of decisions in different contexts and will support the crews in explaining

their decisions. This corresponds with the present philosophy of CRM and Naturalistic Decision-making (Helmreich et al. 2001; Kaempf and Orasanu 1997; Reason 2000; Zsambok and Klein 1997). However, this is not sufficient, since experiencing themselves a loss of NI and recognizing the evolving situation is essential to be able to recover from human error in the future, provided that the instructor intervenes according to the MSI-concept.

The success of this training approach will depend on the instructors who apply it. Therefore, instructors, scenario developers and course designers should be educated in the MSI-concept and trained in applying the technique based on this concept.

The training approach should be included as regular part of the recurrent training program, just as CRM has been integrated in the regular program to date, but only after a special spin-off training has been given to the pilots addressing specifically recovery from human error. The spin-off training requires a special error provoking scenario which should be carried out in a full scale simulator since it approaches closest to the actuality. Audiovisual feedback to debrief the session and the behaviour of the crew is essential to benefit maximal from this training approach.

6.8 Conclusion

The training approach we have been developing to support the aircrew to establish a recovery from human error is based on the assumption that human errors will occur on occasion. Since humans in these occasions are the most flexible link in the situation as well, the emphasis in training should also be on recognizing human error and establish a recovery from it.

Our study shows that the MSI-concept and the accompanying intervention technique, originally developed for and implemented successfully in the area of student counselling (Van der Vlugt *et al.* 2004), can be used as starting point to develop an additional training program that supports aircrews to realize a recovery from human error in an actual flight. From observations of simulator training sessions at three different companies it appears that the training method proposed is not regularly integrated in existing training programs. Nonetheless, some of the instructors elicited aspirant aircrews to recover from their human errors. The challenge is to make the technique more concrete and generally applicable so it can be incorporated formally in existing training programs.

REFERENCES

Capra F (1996) *The Web of Life, A New Scientific Understanding of Living Systems.* New York: Anchor Books.

FAA (1998) *Advisory Circular: Crew Resource Management training.* AFS-210/AC120-51C. Washington DC: Federal Aviation Administration.

FAA (2000) *Advisory Circular: Standard Operating Procedures for Flight Deck Crewmembers.* AFS-210/AC120-71. Washington DC: Federal Aviation Administration.

Helmreich RL, Wilhelm JA, Klinect JR, and Merritt AC (2001) Culture, error and Crew Resource Management. In: Salas E, Bowers CA, and Edens E (eds.) *Improving Teamwork in Organizations.* Hillsdale New Jersey: Lawrence Erlbaum Associates Inc. pp. 305-331.

Kaempf GL, Orasanu J (1997) Current and Future Applications of Naturalistic Decision Making in Aviation. In: Zsambok CE and Klein GA (eds.) *Naturalistic Decision Making* Mahwah, New Jersey: Lawrence Erlbaum Associates Inc.. pp. 81-90.

Maturana HR, Varela FJ (1972) *Autopoiesis and Cognition; The Realization of the Living.* Dordrecht, The Netherlands: D. Reidel; Publishing Company.

Maturana HR, Varela FJ (1992) *The Tree of Knowledge; The Biological Roots of Human Understanding.* Revised ed. Boston Massachusetts: Shambhala Publications.

Naikar N, Saunders A (2002) Crossing the Boundaries of Safe Operation: Training for Error Detection and Error Recovery. In Proceedings of: *21st European Annual Conference on Human Decision Making and Control* (Johnson CW ed.). Glasgow, pp. 8-14.

Reason J (2000) Human error: models and management. *British Medical Journal*, **320**: 768-770.

Van der Vlugt M, De Booij JWP, Dankelman J, and Wieringa PA (2004) Maintaining Sensing Influence, a new concept to recover from undesired behavior. In: *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (*pp. xx-xx). The Hague, The Netherlands.

Wiegmann DA, Shappell SA (2001) *A Human Error Analysis of Commercial Aviation Accidents Using the Human Factors Analysis and Classification System (HFACS).* DOT/ FAA/AM-01/3. Washington DC: FAA.

Zsambok CE, Klein GA (Eds.) (1997) *Naturalistic Decision Making*. Mahwah New Jersey: Lawrence Erlbaum Associates.

Chapter 7

DISCUSSION

"If I have seen further than most men, it is because I stood on the shoulders of giants."

Isaac Newton, 1676

7.1 HUMAN ERROR MANAGEMENT

The objective of this research was to develop a behaviour model that could give starting points for designing additional measures for remedying persistent human error. In error management the system defences occupy a key position with regard to reducing accidents, including those ascribed to human error. High reliability domains incorporate many system defences that in most cases trap human error, or mitigate its effects. Nevertheless, there are still inadequacies in the system defences considering that fixation and other forms of persistent human error still occur and once in a while lead to an incident or accident. In these cases, it is the persistence of human error that has been unanticipated. This section discusses how the MSI-model can be of additional value to error management. And how the MSI-model can be implemented in the future. Finally, some conclusions are drawn.

EXISTING APPROACHES TO HUMAN ERROR MANAGEMENT

Until today, in error management human error has been approached in two ways, individually and organisationally. The classical individual approach views human error as originating in the human operator. It focuses on the errors of individual operators and isolates the acts of the human operator from its context (which implicitly is interpreted as a specific situation, see Boy (1998)). Their remedy to human error is directed at reducing the fallibility of the individual human operator, who is viewed as free to decide between safe and unsafe acts. In 1983 Rasmussen noted, for instance, that the operator's role at that time was to make up for the holes in the engineer's work (Hollnagel & Woods 2006). Making the right decision, as prescribed by the reference context, is the main purpose of training.

The organisational approach views humans as intrinsically fallible so human operator errors are inevitable in the end. Its remedy to human error is directed at trapping the operator's fallibility or at least at reducing the effect of the operator's fallibility. When an operator-involved accident or incident occurred, it does not focus on who erred but how and why the system defences failed (Reason 2000). In the organisational view human

error represents a temporary inability of the operator to cope effectively with the complexity of the system (Hollnagel & Woods 2006). The central idea is that human error is a consequence of the conditions under which the human operators work. The organisational approach aims at embedding in the system defences the unexpected and rich variety (fallibility) of human behaviour. By optimizing the work conditions through system defences, human errors or at least their effects can be remedied (Helmreich *et al.* 1999). The assumption is that the operator is able to act as expected, when the system defences are providing cues that unambiguously suggest to the operator that a change in course of actions is necessary, and/or prevent the operator from taking a wrong course of actions. The system defences are incorporated to trap the consequence of the operator's fallibility.

Although the solution of both approaches to remedying human error differs, two main similarities can be identified as well. The first similarity is that both approaches adhere to the idea that it is the operator who has to make the right decision in the end. Both approaches improve the operator's competencies for making the right decision by providing training programs. The individual approach has laid the foundation for improving the operator's competencies. The organisational approach took advantage of the individual approach by integrating additional system defences with the required training programs, which have reduced the accident rate further. The second similarity is that both approaches assume that human operators are able to make the right decision when needed, provided that they are protected against their fallibility by means of a professional training program that, in case of the organisational approach, is combined with sophisticated system defences.

A NEW APPROACH TO HUMAN ERROR MANAGEMENT: MSI-APPROACH

In this thesis a third way to view human error has been introduced, the MSI-approach. This approach challenges the assumption that human operators are able to make the right decision when needed, because this assumption has been shown to be wrong in practice. The ability to make the right decision by reasoning is not given to humans when they are only given different but appropriate possibilities to act. Every decision needs an evalutation criterion that exceeds the possibilities the human operator is facing. When the organisation does not provide such a criterion or if the criterion supplied falls short for the situation the human operator is facing, then the operator uses his private final criterion. Unfortunately, both the individual and the organisational approach do not provide this final criteria underlying all human acting, that is, they do not provide the final

criterion that explains both human error and human success. The idea that humans are able to choose reasonably, when given the opportunity, seems so self-evident that it has not been discussed until now as the main fallibility of all humans (e.g. human operators).

With the MSI-approach a final criterion has been introduced on which all humans act continually without exception: Maintaining Sensing Influence (MSI). This approach abstracts the rich variety of concrete specific human behaviour to the abstract MSI-process underlying all human behaviour. The MSI-process is not a reasonable criterion, it is a necessary criterion for all humans to continue to exist. Strictly viewed, MSI is a sophisticated representation of survival, i.e. social survival. In the MSI-model humans do not decide by reasoning; they simply act by their necessity of Maintaining Sensing Influence. As long as the human operators act according to the expectations of an observer (i.e. organisational context), their acting that is subordinate to the MSI-process is fully congruent with their acting explained by the individual or organisational approach. But as soon as operators deviate from the expectations that are based on the organisational context, the MSI-model observes two different kinds of behaviour. It observes erroneous behaviour with regard to the organisational context (the reference or situational context) but appropriate behaviour with regard to the private self-evident context.

Connecting the MSI-approach with the existing approaches

With the MSI-approach the focus balances between the classical individual approach and the organisational approach. In contrast with the classical individual approach, the MSI-approach focuses on the individual operator within his private context that is subordinate to his essential process Maintaining Sensing Influence. The MSI-approach views human operator behaviour as existing only in relation with the observer and/or other humans. Consequently, human operator failures only occur in relation to an organisational context, that is, to organisational expectations. The MSI-approach agrees with the organisational viewpoint that human operator errors are conditioned by the situation which in turn is created by the organisation.

According to the organisational approach (Reason 2000), persisting human errors become visible when they have slipped through the holes in the defence layers and have led to an incident or accident. An accident or incident is explained with a momentary line up of the holes in the system defences that permit a sequence of human errors (see Figure 7.1). In error management the remedy has been to repair the holes such that the

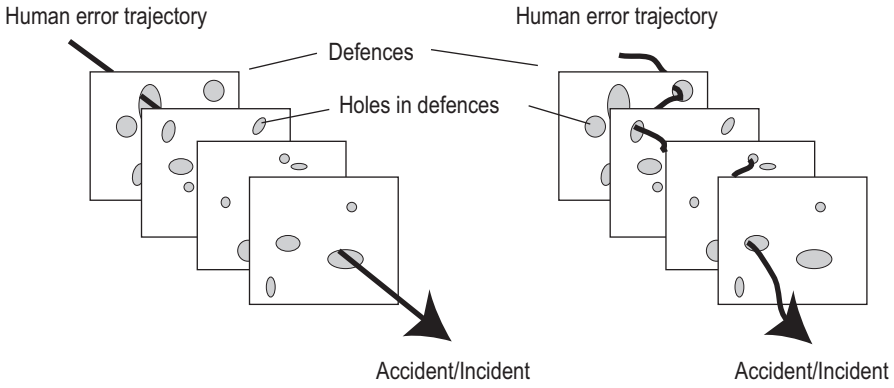


Figure 7.1 *Illustration of difference in view between the organisational approach of Reason (left) and the approach based on the MSI-model (right), with regard to a sequence of human operator errors. In the view of Reason an accident happens when the holes in the defences momentarily line up permitting an opportunity for a sequence of human errors. According to the MSI-model the holes do not line up but the operator will find the holes in the defences once he has changed to an evident context to maintain SI, since the holes are manifestations of self-evidences.*

human errors can be trapped before they lead to an unsafe outcome. However, the followers of the organisational approach also acknowledge that repairing all the holes in the defences is unfeasible, because defences have many holes which are continually opening, shutting, and shifting their location (Reason 2000).

With the MSI-model, a crew-involved accident is now explained as: human operators who find, by necessity, the holes in the system defences when they have changed to an evident context to maintain SI (see Figure 7.1). The MSI-model postulates that operators change by necessity to an evident context when the operator is uncertain about the consequence of a decision, which threatens his SI. As long as the operator maintains SI in a private self-evident context, he shows persistent human error to the observer who maintains SI in the organisational context. This view is supported by Wickens (1992) who observed that “human operators tend to seek (and therefore find) information that confirms their chosen hypothesis and to avoid information that could reject it, which produces inertia that favours the hypothesis initially formulated.” Note, that inertia means unwillingness. In the view of the MSI-model ignorance of information at odds cannot correlate with unwillingness, but correlates with a, for the operator, necessary change

to an evident context to maintain his SI. When an 'organisational' observer perceives a continually moving and changing of the holes in the system defences, it can be seen as a manifestation of different operators who each apply a different private evident context in one organisational context.

The MSI-approach also explains why it is not only infeasible but actually impossible to repair all the holes in the system defences:

1. Creative autonomous behaviour shown actually by an operator is a manifestation of self-evident social rules that emerged in social representing. When acting autonomously and creatively, human operators make sense of the many influences impinging on them drawing not only on experience (acquired skills and rules) but also on knowledge and common sense. Such behaviour is impossible to translate into a reference context completely, because it has accumulated in a history of successful recurrent interactions in a vast number of encountered situations (Dreyfus 1997; Winograd and Flores 1986).
2. Self-evidences are hard to define because they are continually (though slowly) changing due to the ongoing process of social representing. The moment we think we have defined them they may have changed already.
3. Self-evidences are invisible as long as the operator acts in accordance with the organisational expectations (reference context). They only become apparent when a crew acts abnormally with regard to these expectations (Watzlawick 1989).

The contribution of the MSI-approach to Human Error Management

It is always the human contribution that on the one hand improves safe operation and on the other hand jeopardizes safe operation. Both the individual and the organisational approach have contributed largely to safe operation of humans in complex dynamic systems and consequently to safe operation of complex dynamic systems as a whole. When, despite the training and the system defences, human error occurs and persists, the contribution of the MSI-approach to human error management becomes significant.

The MSI-approach focuses on the question: what made an operator do something that is more important to him (otherwise he would not do it) than what he is expected to do. The presupposition of this approach is that when an operator errs, he has the profes-

sional capacity to do what is expected, otherwise there would be no operator error. An operator shows erroneous behaviour when he maintains SI in his individual evident context by showing self-evident behaviour. Applying the MSI-approach to erroneous behaviour means that we have to find the operator's self-evidences in that particular situation that are expressed by his evident behaviour that prevails over his expected behaviour with which he can not maintain his SI. Note that the organisational approach also presupposes that the operator is able to do what he is expected to do, but they conclude that an *operator* error is due to a poor working condition. However, this actually implies that, according to our definition, the error is not an operator error any more.

For the success of the MSI-approach, being an individual approach, the organisational approach is a prerequisite. Using the analogy of Reason (2000), the individual self-evidences are like mosquitoes. They can be swatted one by one, but they still keep coming. The organisational approach concentrates on the organisational self-evidences, by draining the swamps (i.e. by designing system defences). Tightening up the organisational holes in the defences reduces the chance of defeating them, and more important, it improves the chance to recognise human failures before they lead to an accident or incident. The highly procedural domain of aviation is a good example of this mechanism, where the Standard Operating Procedures form a reliable and firm system defence. The more difficult it is for the operator to slip through the holes of the system defences, the more pregnantly his self-evidences underlying human failures will show.

7.2 Procedural system defences: rule management

The re-examination of two crew-involved accidents in Chapter 4 pointed out that essential to the accidents was the uncertainty of the crew about the possible consequences of different courses of actions. In human error management, rule management is one of the key system defences that deals with these uncertainties. Rule management represents the procedural part of the reference context. Two approaches can be distinguished (Grote 2004):

- the "minimising uncertainties" approach,
- the "competently coping with uncertainties" approach.

In the minimizing uncertainty approach, the key assumption is that humans themselves are the uncertainty with respect to system safety. A way to minimise this uncertainty is to develop prescriptive operating rules restricting as much as possible the operator's

creative autonomy. In other words, the uncertainty, induced by the human operator himself, is minimised by trying to make the reference context watertight. However, it has been noticed that a reference context that is too tightly designed can be counterproductive as well (Reason 1995; Grote 2004). They make the operator rely on the tightly coupled reference context in situations that may require creative actions of the operator that go beyond the actions prescribed by the reference context.

Because of the possible negative side-effect of the minimising uncertainties approach, Grote (2004), Weick (1995) and Klein (1991) suggested that it is necessary to shift the attention to the 'coping with uncertainties' approach. With regard to rule management this implies the development of a maximally loosely coupled reference context that ensures that the individual operator uses his autonomous creativity when coping with uncertainties. This 'coping' approach is based on trust in human competence (Grote 2004; Grote and Künzler 2000). Note that in the minimising approach the uncertainty refers to the operator and that in the coping approach uncertainty refers to the situation or the working condition. This corresponds respectively to the classical individual and organisational approach. To achieve a more loosely coupled reference context, Grote (2004) proposes systematic rule management focusing especially on the design of process rules. Process rules define the way in which decisions about a course of action must be arrived at (Hale and Swuste 1998). This proposal is based on the assumption that more process rules may be helpful for crews to optimize their autonomous creativity, and on the observation that a generally low level of process rules exists in aviation.

When it is decided to loosen the reference context, the question arises how to optimize between a too loosely and a too tightly coupled reference context; a question that has already been posed by others among which Dekker (2003), and Woods and Roth (1988). And how to design such process rules, a question that has been investigated by Grote (2004).

An example of an organisational process rule mentioned by Grote is: "do not use personal leadership when you have standard procedures." This rule justifies itself, but does not give a guidance what to do when, in case of an emergency, an operator is using his personal leadership although standard procedures are available and preferable (ruling her or him out with an overruling action rule?). Moreover, who has to decide when the standard procedures are not satisfying any more and process rules are more appropriate? The operator who is using his personal leadership?

The MSI-approach could give an answer to the question how to optimize between a loosely coupled and a tightly coupled reference context. The criterion used for defining the optimum is the operator's SI. In those situations for which the autonomous creativity of humans is essential to safe operation, process rules should be designed such that the operator's Noting Influence (NI) is maximised, the reflective form of SI. The characteristic of those situations is that a minor change of work conditions could have a major negative influence on the NI of the operator.

7.3 Engineered system defences: Man-machine interaction

In this thesis the minimum social coupling that is considered was human-human interaction. How is this translated into human-machine interaction? According to the MSI-model a machine, for instance an aeroplane, is nothing more and nothing less than a materialised set of social representations (rules of appropriate behaviour). An aeroplane, consisting of technical devices, automatic devices and user interfaces, represents the engineered part of the reference context. This is comparable with the Standard Operating Procedures that form the procedural part of the reference context. The aeroplane represents one large nested reference context communicating to the crew the rules or appropriate behaviour.

This engineered (cockpit) environment is limited in its interaction with the crew because by engineering the social representations into technical (sub)systems, they become fixed. That is, they do not keep pace with the autonomously developing social representations amongst humans. As Norman (1990) has put it, the interaction between the automatic devices in the cockpit and the crew is primitive and leaves much room for improvement. According to Norman and Hutchins (1995) the issue is, whether or not a user interface or an automatic device can interact with the crew in a way the crew interact amongst themselves. Because the technical devices do not develop autonomously in interaction with humans they have to be updated when they are not in line any more with the ruling social representations.

7.4 Recommendations for further research

The research described in this thesis had two main objectives, to develop a new conceptual behaviour model that explains persistent erroneous behaviour. And to formulate

starting points for developing additional remedial measures against this form of behaviour. The main starting point is the search for self-evidences, also called the unreflected social representations (see Chapter 3), which in a particular situation condition human erroneous behaviour. Future research should be aimed at finding the self-evidences that are ruling when an operator errs. Although behaviour originating from self-evidences is highly individual, the self-evidences themselves are not.

Training for recovery from human error in aviation

Although aviation was mainly used as means to develop the MSI-model, it can benefit from the findings of the model presented in this thesis as well. Future research can be aimed at finding the ruling self-evidences for parties involved such as pilots, instructors, aviation companies and regulatory authorities. Re-examination of more NTSB accident reports based on the MSI-model as is carried out in Chapter 4, can be the first stage of finding more self-evidences. An illustration of a possible self-evidence is for example the FAA-rule: “duty time should not be exceeded,” that was found in the case of flight AA1420 (NTSB 2001). The Safety Board concluded that there was no evidence that the duty overtime played a role, *because* it was not discussed by the crew. However, it could also be that it was so self-evident that it was not discussed explicitly by the crew, starting a silent combat between ruling self-evidences from different origins when diverting to another airport is a point of discussion. The possible self-evidences found in the accident reports should be examined with experts. The MSI-view emphasizes that when these self-evidences can be made explicit, they can in case of necessity, serve as a subject of discussion between pilot and co-pilot and by doing so, be supportive for their NI, which enhances the chance of a safe outcome.

The possible benefits of implementation of the MSI-model in aviation have not been proven yet, only demonstrated. To be able to “prove” the benefits:

1. a specialized MSI-training program should be developed with and for a core group of instructors.
2. a specialized MSI-training program should be developed with instructors for the crews.

The training program for instructor should aim at initiating a small group of instructors into the use of the MSI-model. This includes:

- How to develop a specialized training program for the crew that aims at remedying persistent human error. This implies that the program should be developed so as to challenge the crew to its maximum and to let them err when they err.
- Recognising a change of the crew to an evident context,
- Knowing what open questions to ask when discussing crew errors so as to discover the crew's self-evidences,
- Recognising the wrong answers to the questions,
- Knowing how to explain the crew's erroneous behaviour based on the MSI-model so that a positive effect of the training is ensured.

When the instructors are initiated into the practical use of the MSI-model via a training program, the next stage is to develop a specialized training program for the crews and to counsel the instructors when backup is needed to ensure a positive training.

Added value of the MSI-model in medicine

The MSI-model is extremely appropriate for enhancing safety in domains that require high reliability but have to rely on a situational context, i.e. on a context that is uncertain. Such domains include for example surgery and anaesthaesia. In aviation the aeroplane is engineered and, therefore, contains almost no surprises, the uncertainties are mostly limited to weather conditions or other natural conditions that cannot be controlled completely by engineering. This is not the case in medicine. Handling a patient induces more uncertainties not only because a human patient is not engineered and the signals indicating the patient's state can be ambiguous, requiring a further diagnosis (Xiao *et al.* 1997), but also because a patient is a human that is malfunctioning. Handling a patient is like handling a malfunctioning aeroplane that is full of surprises. The enhanced complexity of the human patient leads to even more shifting holes than can be identified in aviation. Therefore, the role of self-evidences is expected to be of a much greater influence than is the case in high reliability domains with well-defined reference contexts.

7.5 Conclusions

The aim of our research was to develop a conceptual behaviour model with which it was able to find starting points for developing remedial measures against persistent hu-

man error, such as fixation. Fixation is an eye-catching form of erroneous behaviour, but difficult to explain. For the development of a new conceptual behaviour model that could explain fixation, we followed Schrödinger's advise: not so much to see what no one has yet seen, but to think what nobody has yet thought, about what everybody sees.

The conceptual behaviour model discussed in this thesis is based on a system theoretical approach in combination with the conceptual theory of autopoiesis introduced by Maturana and Varela (1980) in the 1970s. To be able to describe the process that is characteristic to all human behaviour, irrespective of human error or human success, the concept of autopoiesis has been elaborated so as to incorporate human social behaviour. The essential process of Maintaining Sensing Influence together with social representing was identified as the basis for all human behaviour irrespective of it being appropriate or erroneous.

The MSI-model postulates that human behaviour is not determined by the work/task situation influences as is assumed in the existing behaviour models, but by the essential process of Maintaining Sensing Influence. Although the work/task situation contributes to the human operator's behaviour, in the end it is not decisive. This does not contradict with earlier findings. Nevertheless, in many cases, the task situations seem to be decisive, this is when the existing models apply. The existing behaviour models can be viewed as situation specific cases of the MSI-model.

The MSI-model does not differ from the other models with regard to detecting human (persistent) errors. Just like the other models, it requires a norm to be able to detect a deviation from it. The essential difference with the existing behaviour models is that the MSI-model uses the MSI-process instead of the task situation to explain human error. The benefit of using the MSI-model is that for explaining human error, it is not necessary anymore to determine in advance all possible task situations.

REFERENCES

Boy GA (1998) *Cognitive Function Analysis*. London: Ablex Publishing Corporation.

Dekker S (2003) Failure to Adapt or Adaptations that Fail:Contrasting Models on Procedures and Safety. *Applied Ergonomics*, **34:** 233-238.

Dreyfus HL (1997) Intuitive, Deliberative, and Calculative Models of Expert Performance. In: Zsambok CE and Klein GA, eds. *Naturalistic Decision Making*. Mahwah, New Jersey:

Lawrence Erlbaum Associates, 17-28.

Grote G. (2004) Organizational Measures for Achieving Loose Coupling in High-Risk Systems: The Importance of Systematic Rules Management. In proceedings of: *International Conference on Systems, Man, and Cybernetics.* The Hague, The Netherlands: TU Delft. pp. 2580-2585.

Grote G, Künzler C (2000) Diagnosis of Safety Culture in Safety Management Audits. *Safety Science*, **34:** 131-150.

Hale AR, Swuste P (1998) Safety Rules: Procedureal Freedom or Action Constraints? *Safety Science*, **29:** 163-177.

Helmreich RL, Merritt AC, Wilhelm AJ (1999) The evolution of Crew Resource Management in commercial aviation. *International Journal of Aviation Psychology*, **9:** 19-32.

Hollnagel E, Woods DD, Leveson N (2006) *Resilience engineering.* New York: Ashgate Publishing.

Hutchins E, Klausen T (1996) Distributed cognition in an airline cockpit. In: Engeström Y and Middleton D, eds. *Cognition and communication at work*. New York: Cambridge University Press, 15-34.

Klein GA, Calderwood R (1991) Decision Models: Some Lessons From the Field. *IEEE Transactions on Systems, Man, and Cybernetics*, **21:** 1018-1026.

Reason J (1996) A systems approach to organizational error. *Ergonomics*, **39:** 1708-1721.

Reason J (2000) Human error: models and management. *BMJ*, **320:** 768-770.

Watzlawick P (1989) *Munchhausens haren en Wittgensteins ladder:uit de greep van de werkelijkheid*. Deventer: Van Loghum Slaterus.

Weick KE (1995) *Sensemaking in Organizations*. Thousand Oaks, California: Sage Publications.

Wickens CD, Huey BM (1993) Workload Factors. In: Wickens CD and Huey BM, eds. *Workload Transition: Implications for Indiviual and Team Performance*. Washington, D.C.: National Academy Press, 1-284.

Winograd T, Flores F (1986) *Understanding Computer and Cognition: A New Foundation for Design*. New Jersey: Ablex Publishers.

Woods DD, Roth EM (1988) Cognitive Systems Engineering. In: Helander M, ed. *Handbook of Human-Computer Interaction*. Amsterdam: Elsevier Science Publishers B.V., 3-43.

Xiao Y, Milgram P, Doyle DJ (1997) Planning behavior and its functional role in interaction with complex systems. *In: IEEE transactions on man machine and cybernetics - part A: systems and humans.* **27**:313-324

Summary

MAINTAINING SENSING INFLUENCE:
AN AUTOPOIETIC BEHAVIOUR MODEL FOR REMEDYING PERSISTENT HUMAN ERROR

"There is nothing so practical as a good theory"

Kurt Lewin (1890-1947)

In high reliability systems human errors may have serious consequences when they lead to an accident or a near accident. Therefore, much effort has been put into minimising the share of human error in (near) accidents. As a result, the accident rate has decreased to a substantially low level in the last two decades. Unfortunately, several studies into aviation accidents revealed that the share of accidents attributed to human errors remained relatively high compared with the share of accidents attributed to technical errors.

A form of human error that persisted is fixation (also called tunnel vision, or pre-occupation). Fixation can be characterised as maintaining to a course of actions despite unambiguous signals given by the system that the wrong course of actions is taken for realising the desired objective. In aviation, many measures have been developed that unambiguously show the operator(s) they took the wrong course and that urge them to change it. However, these measures are primarily based on the condition of the system operated and only secondarily on the condition of the human operator.

The difficulties that are experienced with remedying persistent human error can be ascribed to an insufficient variety of the behaviour modelled on which the developed measures are established. Ashby formulated in the 1950s the law of "Requisite variety." This law postulates that for an effective remedy of human errors, the variety of the behaviour generated by the human behaviour model should match the variety of the behaviour shown by the human operator.

The existing human behaviour models have in common that they have been induced from a quantitative input-output analysis. In the case of remedying persistent human error, this method of modelling falls short for two reasons. Firstly, humans are subject to a vast number possible inputs (input states) because in reality, situations do not neatly end at some point. To be able to model human behaviour through input-output analysis the common approach has been to reduce the vast number of possible inputs to a set of specific inputs that are considered to be relevant for the behaviour studied. This reduction is at the cost of richness in variety. Secondly, humans as creative systems are adapting systems. This implies that their present input-output relationship depends on their previous output state. In other words, their input-output relationship is not invariant.

Therefore, the input-output relationship obtained, representing human behaviour, is only valid for the defined situations within the defined period of time.

As alternative to the input-output relationship, the process “Eigen” behaviour has been taken to develop a qualitative human behaviour model. Eigen behaviour has been defined as a pattern that develops about a dynamic equilibrium and is considered to be common to all adapting operational closed systems, such as humans. The advantage of Eigen behaviour is that it can be recognised without the necessity of defining all possible input states of the system. Eigen behaviour is characterised by the system’s operational closure rather than by environmental influences. Eigen behaviour is the pattern that is general to all specific behaviour, i.e. it is the basis for the richness in variety of specific human behaviour.

For defining Eigen-behaviour, Maintaining Sensing Influence (MSI) is proposed. MSI is based on the theory of autopoiesis (or self-creation) introduced by the neurobiologists Maturana and Varela in the 1970s. MSI abstracts the richness in variety of specific human behaviour to an operational closed process that is general to all human behaviour. In our study it is assumed that all human behaviour is subordinate to MSI. Sensing Influence (SI) is defined as that what environmental influences mean to a living system, such as the human. How SI is maintained is determined by the present structure of the human, and not by the environmental influences as is assumed by the current behaviour models. Environmental influences affect but do not determine Eigen behaviour.

According to the MSI-model, behaviour only exists as the whole of acting and observing this acting. In this process the operator as well as the observer are subordinate to MSI. The way in which the observer assigns meaning to the operator’s acting (contextualizes the operator’s acting) supports his SI. Within the MSI-model, behaviour has been defined as a (by the observer) contextualizing of acting. This means that the operator’s behaviour is determined by the observer’s context. In this perspective, human operator error is a manifestation of a discrepancy between the context of the observer and the context of the operator. Interventions that use the MSI-model for remedying erroneous behaviour, particularly aim at neutralizing this discrepancy. Often the observer, as a relative outsider, is better able to realize this than an operator. When the observer becomes fixated as well, only a third party can intervene successfully.

The added value of the MSI-model has been illustrated by applying it to two aviation accident cases, flight 801 of Korean Air and flight 1420 of American Airlines, which the National Transportation and Safety Board of the United States (NTSB) both ascribed

to fixation of the crew. The diversity in the explanations for fixation in both NTSB analyses was connected through the critical process 'MSI.'

The effectiveness of the MSI-model for remedying persistent human error has been illustrated by a student counsellor who applied it to students whose study progress had come to a standstill. Although they formerly studied with success, they now showed erroneous behaviour, which persisted despite several remedial interventions. This behaviour satisfies the definition of fixation. Using the MSI-model, the student counsellor could assist the students to recognize their fixation and to reflect on the necessity of fixation for their MSI. Then, MSI-based interventions supported the students to remedy their fixation effectively.

Pilots cannot be given a personal experience of fixation and remedying fixation when they fly an aeroplane, due to the possible serious consequences. Therefore an outline has been drawn for a training program that comes as close as possible to a real life situation. The first condition for an effective training session is that the instructor allows the pilot(s) to err and to pass the boundaries of safe operation. The second condition is that the instructor remains capable to reflect on the fixation of the pilots, how undesired their behaviour may seem to him with regard to a safe outcome. Because the MSI-model contributes to realizing the second condition, it is of great importance for instructor training.

The first advantage of the MSI-model, is that the impossibility to define the vast number of possible inputs that affect human behaviour has been avoided. The second advantage is that with the MSI-model the responsibility is given back to the operators without blaming them, through which human error can become the subject of research again. The main disadvantage about Eigen behaviour and therefore the MSI-model is that the possibility is lost to predict *when* human error will occur. Although the model is based on the MSI-process that is characteristic for human behaviour in general, its interpretation remains a personal matter. The model provides an explanation for human error when it has occurred and from which interventions can be derived to establish a recovery from it. Due to its specific approach of erroneous behaviour, the MSI-model is of particular interest for developing remedial measures in case the existing remedies fall short.

Depending on the professional field, the MSI-model can serve as a starting point for developing specific interventions to enhance the safety and the performance in high-risk domains that contain uncertainties which cannot be regulated, such as surgery and anaesthesia.

Samenvatting

MAINTAINING SENSING INFLUENCE:
EEN AUTOPOIETISCH GEDRAGSMODEL VOOR HET REMEDIËREN VAN HARDNEKKIG FOUTIEF GEDRAG

"Niets is zo praktisch als een goede theorie"

Kurt Lewin (1890-1947)

In hoog-risico systemen zoals vliegtuigen kunnen menselijke fouten ernstige consequenties hebben zoals een ongeluk of een 'bijna-ongeluk'. Daarom zijn de afgelopen decennia veel inspanningen geleverd om het aandeel van menselijke fouten in (bijna) ongelukken te minimaliseren. Hierdoor is het aantal vliegtuigongelukken de afgelopen decennnia tot een zeer laag niveau gedaald. Uit verschillende onderzoeken naar vliegtuigongelukken blijkt echter dat het aandeel van de menselijke fout nog steeds hoog is vergeleken met dat van technische fouten.

Een vorm van hardnekkig foutief menselijk gedrag die nog steeds voorkomt is fixatie (ook wel tunnelvisie of pre-occupatie genoemd). Fixatie kan worden gekarakteriseerd als het vasthouden aan een ingeslagen weg, ondanks eenduidige signalen van het systeem dat voor het realiseren van het gewenste doel de verkeerde weg is ingeslagen. Er zijn in de luchtvaart vele middelen ontworpen, zoals alarmsystemen en procedures, die vliegers eenduidig erop wijzen dat zij de verkeerde weg zijn ingeslagen en hen manen de juiste weg te nemen. Echter, in eerste instantie zijn deze middelen gebaseerd op de toestand van het vliegtuig en pas in tweede instantie op de toestand van de menselijke operators.

De moeilijkheden die worden ondervonden bij het remediëren van hardnekkig foutief gedrag kunnen worden toegeschreven aan een ontoereikende variëteit van het gemodelleerde gedrag waarop de ontworpen remedies zijn gebaseerd. Ashby heeft in de jaren vijftig de wet van "Requisite Variety" geformuleerd. Deze wet stelt dat voor een effectieve remediëring van foutief gedrag, de variëteit van het gedrag dat door het gedragsmodel wordt gegenereerd minimaal gelijk dient te zijn aan de variëteit van het gedrag dat de wordt gegenereerd door menselijke operators.

De bestaande modellen van menselijk gedrag hebben als uitgangspunt dat ze geïnduceerd zijn uit een kwantitatieve ingangs-uitgangsanalyse. Voor het remediëren van hardnekkig foutief gedrag schiet de inductiemethode om twee redenen tekort. Ten eerste zijn mensen onderhevig aan een ontelbaar aantal mogelijke invloeden (ingangstoestanden) doordat in de praktijk situaties niet exact afgrensbaar zijn. De aanpak om menselijk gedrag toch te kunnen modelleren via ingangs-uitgangsanalyse, is het reduceren van het grote aantal mogelijke invloeden tot een set van specifieke invloeden die als relevant worden beschouwd voor het gedrag dat wordt onderzocht. Deze reductie

gaat echter gepaard met een vermindering van de rijkdom in variëteit.
Ten tweede zijn mensen als creatieve systemen eveneens adapterende systemen: hun huidige ingangs-uitgangsrelatie hangt samen met hun vorige uitgangstoestand. Met andere woorden, hun ingangs-uitgangsrelatie is niet invariant. Om die reden is de door de inductiemethode verkregen ingangs-uitgangsrelatie alleen geldig voor de binnen een bepaald tijdsbestek gedefinieerde invloeden.

Als alternatief voor de ingangs-uitgangsrelatie is, voor het ontwikkelen van het kwalitatieve gedragsmodel, het proces “Eigen-gedrag” als uitgangspunt genomen. Eigen-gedrag is gedefinieerd als een zich ontwikkelend gedragspatroon rondom een dynamisch equilibrium en wordt beschouwd als universeel voor alle adapterende, operationeel gesloten systemen, zoals mensen. Het voordeel van Eigen-gedrag is dat het herkend kan worden, zonder dat het noodzakelijk is eerst alle mogelijke ingangstoestanden van het systeeem te definiëren. Eigen-gedrag wordt bepaald door de operationele geslotenheid van het systeem in plaats van door omgevingsinvloeden. Eigen-gedrag is het universele patroon van menselijk gedrag, dat wil zeggen het vormt de basis voor de rijkdom in variëteit van specifiek menselijk gedrag.

Voor het definiëren van Eigen-gedrag wordt Maintaining Sensing Influence (MSI) voorgesteld. MSI bouwt voort op de theorie over autopoiese (of “zelf-creatie”) geïntroduceerd door de neurobiologen Maturana en Varela in de jaren zeventig van de vorige eeuw. MSI abstraheert de rijkdom in variëteit van specifiek menselijk gedrag tot een operationeel gesloten proces dat universeel is voor al het menselijke gedrag. In ons onderzoek wordt ervan uitgegaan dat, hoe specifiek menselijke gedrag ook is, het altijd voldoet aan het universele proces MSI. Sensing Influence (SI) is: dat wat de omgevingstoestand betekent voor een levend systeem zoals de mens. De manier waarop SI wordt behouden wordt bepaald door de momentane structuur van een mens en niet door de omgevingstoestand, zoals wordt verondersteld door de huidige gedragsmodellen. Omgevingstoestanden zijn wel van invloed, maar niet bepalend voor Eigen-gedrag.

Volgens het MSI-model is gedrag het geheel van handelen en het waarnemen van dat handelen. In dit proces zijn zowel de operator als de waarnemer onderhevig aan MSI. De manier waarop de waarnemer het handelen van de operator van betekenis voorziet (contextualiseert), ondersteunt zijn MSI. Gedrag wordt binnen het MSI-model dan ook gedefinieerd als (door de waarnemer) ‘van betekenis voorzien handelen’. Dat betekent dat operatorgedrag wordt bepaald door de context van de waarnemer. Foutief operator gedrag kan derhalve worden opgevat als een manifestatie van een discrepantie tussen

de context van de waarnemer en die van de operator. Interventies die gebruik maken van het MSI-model om foutief of afwijkend gedrag te remediëren, zijn met name gericht op het neutraliseren van deze contextdiscrepantie. Vaak zal dit de waarnemer, als relatieve buitenstaander, makkelijker afgaan dan de operator. Als de waarnemer eveneens in fixatie is beland, kan alleen een derde partij succesvol interveniëren.

De toegevoegde waarde van het MSI-model wordt geïllustreerd door het toe te passen op twee vliegtuigongelukken, vlucht 801 van Korean Air en vlucht 1420 van American Airlines, die beide door de National Transportation and Safety Board (NTSB) zijn toegeschreven aan fixatie van de bemanning. De verscheidenheid in de verklaringen voor fixatie in beide analyses van de NTSB zijn met elkaar verbonden via het kritische process van MSI.

De effectiviteit van het MSI-model voor het remediëren van hardnekkig foutief gedrag bleek uit de toepassing door een studentendecaan bij studenten die vastliepen in hun studie. Alhoewel ze voorheen met succes studeerden, vertoonden ze nu afwijkend studiegedrag, dat voortduurde ondanks allerlei op remedie gerichte interventies. Dit gedrag voldoet aan de definitie van fixatie. Met gebruikmaking van het MSI-model slaagde de studentendecaan erin de studenten te laten (h)erkennen dat dit afwijkend gedrag noodzakelijk was voor hun MSI. Vervolgens ondersteunden eveneens op het MSI-model gebaseerde interventies de studenten bij het effectief remediëren van hun fixatie.

Piloten in de praktijk fixatie en de remedie tegen fixatie te laten ervaren, zoals dat gebeurde met de studenten, is niet mogelijk. Daarom is een voorstel gedaan voor een trainingsprogramma dat de praktijk zo dicht mogelijk benadert. De eerste voorwaarde voor een effectief trainingsprogamma is dat de instructeur de vlieger(s) toestaat fouten te maken en hen de grenzen van veilig opereren te laten overschreiden. De tweede voorwaarde is dat de instructeur in staat blijft te reflecteren op het fixatiegedrag van de vliegers, hoe ongewenst hun gedrag ook mag lijken met het oog op een veilige uitkomst. Doordat het MSI-model bijdraagt aan het realiseren van de tweede voorwaarde, is het van groot belang voor het trainen van instructeurs.

Het eerste voordeel van het MSI-model, dat uitgaat van Eigen-gedrag, is dat de onmogelijkheid om de ontelbare mogelijke invloeden te definiëren wordt vermeden. Het tweede voordeel is dat met het MSI-model de verantwoordelijkheid voor het handelen is teruggelegd bij de operators zonder hen te beschuldigen, waardoor human error weer het centrale onderzoeksonderwerp kan worden. Het grootste nadeel van het MSI-model is

dat de mogelijkheid verloren is gegaan om specifiek foutief menselijk gedrag te kunnen voorspellen. Alhoewel het model is gebaseerd op het MSI-proces dat universeel is voor menselijk gedrag, blijft de invulling ervan een persoonlijke zaak. Het MSI-model levert een verklaring voor foutief menselijk gedrag nadat het is opgetreden, waaruit interventies kunnen worden afgeleid om eventueel gewenst gedrag te herstellen. Door de specifieke benadering van afwijkend gedrag is het MSI-model in het bijzonder van waarde, daar waar andere remedies niet (meer) effectief (kunnen) zijn, als laatste redmiddel.

Afhankelijk van het professionele werkterrein, kan het MSI-model als uitgangspunt dienen voor het ontwikkelen van specifieke interventies voor het bevorderen van de veiligheid en performance in hoog-risico vakgebieden zoals chirurgie en anesthesie, die onzekerheden kennen die niet kunnen worden gereguleerd.

Dankwoord

Voor alles en iedereen bedank ik Peter Wieringa voor het vertrouwen dat hij in 2000 had in mijn latente promotiecapaciteiten. Alleen daardoor heb ik de kans gekregen als eerste geldstromer aan dit proefschrift te kunnen werken. Zo goed als zeker heeft zijn vertrouwen in mijn capaciteiten de afgelopen jaren meermaals dieptepunten gekend. Evenwel, dit drukwerkje bewijst zijn gelijk van destijds. Ook mijn vertrouwen in een goede afloop heeft vele dieptepunten gekend. Wat aanvankelijk theoretische onmogelijkheden leken, bleken na veel graven en spitten opeens ongekende mogelijkheden te zijn geworden. Gelukkig heb ik dit niet in m'n eentje hoeven te klaren. Wat ik als een uitzichtloze klus beschouwde, bleek voor Jan iedere keer een nieuwe gelegenheid te zijn om zijn systeemtheoretische inzichten te kunnen toetsen. Zonder een moment van twijfel toog hij aan het werk en na een tijdje samen erover te hebben gepraat en geredeneerd, was het hem zo klaar als een klontje. En ook ik was geheel overtuigd van onze nieuwe inzichten. Probeerde ik deze inzichten vervolgens in mijn Powerbook vast te leggen, dan bleek dat vaak niet te lukken. Dan weet u gelijk wat heeft gemaakt dat ik vijfeneenhalf jaar full-time aan dit proefschrift heb gewerkt. Van de vanzelfsprekendheid dat ik een logische gedachte wel even één op één op papier kon zetten, heb ik erg veel last ondervonden. Om een voorbeeld te noemen, het begrip context staat centraal bij het betekenis verlenen aan handelen. Context was reeds door verschillende auteurs gedefinieerd en heeft als begrip bovendien een sterk 'gezond verstand' karakter. Zonder uitzondering gebruikten die definities echter vanzelfsprekendheden, zoals de aanname

dat de definitie, context is dat wat betekenis verleent, voor zich spreekt. U begrijpt dat ik daarmee niet kon leven. Maar voordat ik iets zinvols op papier kreeg, waren er toch al gauw een paar dagen en soms wel een week versteken.

Toen er dan toch een ruwe en nog onvolledige versie van dit proefschrift gereed bleek te zijn gekomen, heb ik Jenny Dankelman gevraagd of ze mij wilde ondersteunen bij het afronden van mijn proefschrift. En ik kan u in alle eerlijkheid meedelen dat er zonder haar fantastische begeleiding, haar kritische bestudering van mijn hoofdstukken en haar afgrenzing van mijn tomeloze gedachtenwegen, nog steeds geen proefschrift had gelegen.

Ook van buiten de TU Delft heb ik onmisbare medewerking gekregen. In de eerste plaats van een drietal organisaties: CAE Amsterdam, KLM en Transavia. Alle drie hebben mij de gelegenheid geboden vliegers en instructeurs in full-flight simulator trainingssessies te observeren. Een aantal sessies heb ik zelfs op video mogen vastleggen. Van deze organisaties wil ik een aantal personen nog extra bedanken voor hun belangeloze medewerking. Ronald de Wee van CAE, Ruud Wittebol van KLM en John van Limburg-Stirum van Transavia die voor mij de geschikte sessies uitzochten en de instructeurs en vliegers op de hoogte brachten van mijn komst; en niet te vergeten Martin de Koe van CAE, die voor mij een mooi snoertje heeft gefabriceerd waarmee ik de microfoon-ingang van mijn camera direct op de communicatie-kanalen van de simulator kon aansluiten zodat ik kwalitatief hoogstaande geluidsopnamen kon maken.

In de tweede plaats van Luc van den Berge uit Gent, medewerker van de Interactie-Academie te Antwerpen en doctoraal student filosofie aan de Universiteit van Leuven. Vanwege zijn autoriteit op het gebied van de algemene systeemtheorie en zijn kennis van het werk van Maturana en Varela, heb ik hem gevraagd mijn theoriehoofdstuk inhoudelijk te beoordelen. Zijn positieve waardering heeft me een enorme impuls gegeven door te gaan op de ingeslagen weg.

Wie denkt dat ik opgelucht en blij ben dat er eindelijk een proefschrift ligt en dat ik waarschijnlijk binnenkort de dr-titel mag voeren, heeft het mis. Wel ben ik heel erg blij dat ik zolang heb kunnen en mogen werken aan een onderwerp dat mij mateloos boeit. Hoewel ik soms gek werd van mijn gedachte dat het proefschrift af moest zijn binnen eerst vier en toen vijf jaar, heb ik geen moment mijn motivatie verloren. Integendeel, al

doende kreeg ik alsmaar nieuwe ideeën voor verder onderzoek. Als medewerker van NLR, ben ik nu druk doende te bezien op welke manier mijn verkregen inzichten bij kunnen dragen aan het optimaliseren van het gedrag van vliegers.

Machteld van der Vlugt
Den Haag, juli 2006

Curriculum Vitae

1969	Born on November 5, Heemskerk, the Netherlands.
	Secondary Education
1982 -1988	Atheneum, attended at College Hageveld in Heemstede.
	University Education at the TU Delft
1989 -1995	MSc study Mechanical Engineering, graduated at the Laboratory for Systems and Control.
1996 -1998	Advanced Masters study Design Engineering of Consumer Products, at the Institute for Advanced Design Engineering.
2000 -2005	PhD study at the Laboratory for Man-Machine Systems.
	Professional Experience
2006 - present	Developer of training courses at the National Aerospace Laboratory NLR, Amsterdam
2000	Research Fellow at the Laboratory for Man-Machine Systems, TU Delft
1997 - 2000	Industrial Design Engineer, junior project manager, at TNO Industrie