Cognition in Context

The effect of information and communication support on task performance of distributed professionals

Margit Biemans

Telematica Instituut

Enschede, The Netherlands, 2006

Telematica Instituut Fundamental Research Series, No. 019 (TI/FRS/019)

Cover Design: Studio Oude Vrielink, Losser and Jos Hendrix, Groningen
Book Design: Lidwien van de Wijngaert and Henri ter Hofte
Printing: Universal Press, Veenendaal, The Netherlands

Samenstelling promotiecommissie:

Voorzitter, secretaris:	prof. dr. ir. A.J. Mouthaan (Universiteit Twente)
Promotor:	prof. dr. ir. A. Nijholt (Universiteit Twente)
Assistent promotor:	dr. J. Swaak (Telematica Instituut)
	dr. E.M.A.G. van Dijk (Universiteit Twente)
Referent:	dr. H.J. van der Lugt (Telematica Instituut)
Leden: *:*	prof. dr. ing. W.B. Verwey (Universiteit Twente)
	prof. dr. ir. C.A. Vissers (Universiteit Twente)
	prof. dr. H.G. Schmidt (Erasmus Universiteit Rotterdam)
	prof. dr. P.J. Werkhoven (Universiteit Utrecht)
	prof. dr. K.A. Brookhuis (Technische Universiteit Delft, RijksUniversiteit Groningen)

ISSN 1388-1795; No. 019
ISBN-10:90-75176-43-0
ISBN-13:978-90-75176-43-8

Telematica Instituut, P.O. Box 589, 7500 AN Enschede, The Netherlands
E-mail: info@telin.nl; Internet: http://www.telin.nl
Telephone: +31 (0)53-4850485; Fax: +31 (0)53-4850400

COGNITION IN CONTEXT

THE EFFECT OF INFORMATION AND COMMUNICATION SUPPORT ON TASK PERFORMANCE OF DISTRIBUTED PROFESSIONALS

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. W.H.M. Zijm,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 29 juni 2006 om 13.15 uur

door

Margaretha Catharina Maria Biemans
geboren op 3 februari 1971
te ‘s-Hertogenbosch

Dit proefschrift is goedgekeurd door de promotor, prof. dr. ir. A. Nijholt en de assistent promotoren, dr. J. Swaak en dr. E.M.A.G. van Dijk

Acknowledgements

The mountain on the cover photo of this thesis is the Matterhorn in Switzerland. For me, this symbolises a context in which my cognition can relax and gets inspired. It also relates to pleasant feelings; beautiful views, exhausting journeys and other enjoyable activities with my family. This thesis is the result of my journey into human factors research in IT innovations at the Telematica Instituut. After having taken some deviations, I found the proper way, which resulted in this thesis. A journey I couldn't and didn't take on my own.

First and foremost, I thank Janine Swaak for her guidance, her belief in me and her down to earth approach. I liked her positive attitude and her stimulating one-liners: 'you should only run, if you know where to run to' and 'there are various ways to Rome, and you should only take one of them, not all'. I would also like to thank my Ph.D. supervisor Anton Nijholt for his support and his challenging questions on critical issues, and Betsy van Dijk for her assistance with the statistical analyses and her critical reviews.

Thanks are also due to Chris Vissers, Eric Backer and Hermen van der Lugt for their support in finalising the train simulator for my research. Furthermore, I am very grateful to Willem Zwiep, Piet Schouwerwou, Erik Witteveen, Mirjam Vos, Mat Huppertz, Simone van der Velde and all the train drivers, railway signallers, district nurses and dermatologists for making it possible to perform the experiments! Special thanks go the application engineers- Martin Wibbels, Jaap Reitsma, Hans Kruse and William van Dieten – for building the applications that are used in the experiments. I owe many thanks to Henny de Vos, Cees Glas and Peter Geurts for their second opinion on the statistical analyses.

My colleagues at the Telematica Instituut created a pleasant and stimulating work environment. Special thanks go to Mark van Setten for providing a listening ear and for his practical advice on writing the thesis, Henk de Poot for carefully proofreading the draft version of this thesis, and Robert Slagter, Erwin Fielt, Cristian Hesselman and all TAO members for our social and Ph.D. related discussions.

Family and friends thank you for showing interest and having patience with my social absence for the last months. Pepijn and Pim, finally I can say mama's book has been finished, but I have to disappoint you that I am not a 'dokter'. Frans, thank you for everything, now the time has come to bike, hike and sail: our meaning of life!

Margit Biemans
Borne, the Netherlands, May 2006

Contents

Chapter 1

Introduction

Mobile collaboration is an emerging field. Developments in technological and business areas now provide opportunities for innovative applications and new ways of working. More and more mobile applications are being developed for business opportunities. In the present research, the focus was on the cognitive aspects of distributed professionals who were working with new mobile applications that provide information and communication support. The overall research objective was to identify the added value of information and communication support for task performance and the perceived user acceptance and applicability. Two cases were studied: district nurses who deliver wound care at home, and train drivers – who are responsible for driving safely and efficiently. This chapter provides an overview of both cases and of the research approach and the research questions.

The introduction of new technologies for information provision, information sharing, communication and collaboration will change work environments and the way work is organised and perceived (Malone, 2004); (Schaffers, Brodt, Pallot, & Prinz, 2006). Ambulant professionals (e.g. maintenance engineers, patrolmen and mobile service workers) used to work with stand-alone applications at various locations and would contact their colleagues through mobile phones. Technological developments now provide opportunities to enrich this collaboration through the sharing of visual information. When context information can be shared, communication and interaction between professionals changes. Thus, professional task performance is evolving from purely stand-alone activities to activities that provide opportunities for cooperation and interaction.

The mobility of professional team members is increasing, they are more and more widely dispersed, and their working patterns alternate between working autonomously and in cooperation. This kind of loose coupling is 'a style of collaboration that occurs in groups, and that implies that workers can function in a somewhat autonomous fashion without reliance on ongoing interaction with others' (Pinelle & Gutwin, 2003). As the context of cognition changes, cognitive functioning of the individual professional will be affected, for example his or her information processing, decision-making (Klein, 1999) and information overload (De Bakker S.C., 2006).

On a technical level, third-generation mobile networks and beyond provide opportunities for streaming multimedia information (audio and video) and allow graphical rather than menu-driven user interfaces (UMTS forum, 2000). Increased bandwidth facilitates faster response times, which positively affect users' acceptation and satisfaction (Ramsay & Nielsen, 2000). Moreover, an ongoing convergence of media, telecommunication and information technology is taking place in mobile networks. As a result, technological developments offer many opportunities to design end-users' applications that have enhanced information and communication options. In addition, the information provided can be dynamically adapted to location, time and personal preferences. Examples of such applications are expert consultations during maintenance processes, location-based information systems and personalised services.

On an organisational and societal level, there is a trend towards employee or individual empowerment (Robbins, Crino, & Fredenhall, 2002). The empowerment process is best represented by focusing on elements on both the environmental and the individual level: the creation of a local work environment within a broader organisational context that will provide both the opportunity to exercise one's full range of authority and power (i.e. empowered behaviour) and the intrinsic motivation in employees to engage in that type of behaviour (i.e. psychological empowerment).

The combination of the technological and organisational trends in mobile applications and empowerment has led to the development of mobile applications that provide information and communication support to distributed professionals. In this way, distributed professionals are empowered to perform their job. Mobile applications for distributed professionals are typical examples of social technical systems. In the present research, socio-technical systems were defined as:

Definition 1 Socio-technical systems

Systems that are composed of social, psychological and technical elements.

Combining these elements of socio-technical systems creates complex systems. The nature and origin of complexity in such systems can vary substantially. Some systems are characterised primarily by cooperation between people that is mediated by technology, while others are characterised primarily by collaborative human and machine control of some technical system (Vicente, 1999). The two cases studied in the present research – home care and railway – are typical examples of the two extremes of social technical systems: home care relates to cooperation between people that is mediated by technology, while railway relates to the collaborative human and machine control of a technical system.

There is a trend towards developing mobile applications for distributed professionals in order to facilitate (1) access to context information, (2) information exchange between distributed professionals, and (3) enriched communication between distributed professionals. The focus of the present research was on information and communication support for task performance in a dynamically changing work environment, in which professionals can work both on their own and in cooperation with colleagues. These work environments are dynamic in the way that they change frequently; for example, there is dynamic information, cooperation between various people and conflicting goals. The information and communication technology in these environments should match the dynamics of task performance between people.

The following sections provide a brief and general description of the two cases covered by the present research. In each section, the general principles of that domain are described, resulting in an applied research objective. The overall objective of the research is then described and the outline of the thesis is presented.

1.1 Home care

Home care – or community health care – provides nursing care in people's homes (Ribu, Haram, & Rustøen, 2003). Home care promotes the health and welfare of society while sustaining the values of the patient and the home (Baker, 2001). In other words, home care provides medical and non-medical care in the intimacy of people's homes, that is, it provides primary health care. Patients are often referred to home health care services upon discharge from a hospital or other inpatient care facility. This obviates the need for expensive treatment in hospitals or nursing homes, that is, for secondary health care.

Home care is a complex structure of cooperation. It takes place within a complex system, organised in four main sectors, namely (1) community: patient, hospitals and GPs; (2) control: managers of hospitals and of home care; (3) cure: acute hospitals, physicians, community of physicians; and (4) care: nurses and other professionals (Bricon-Souf, Anceaux, Bennani, Dufresne, & Watbled, 2005). The coordination between all these health care professionals is essential for treating a patient in home care.

In today's health care system, patients who are sent home require a much higher level of care than was the case in the past. The increasing complexity of taking care of patients at home has led to a growing recognition that home care is a speciality practice (Baker, 2001). District nurses are often classified as generalists, as they must have multifaceted skills, be flexible, pay attention to detail (Baker, 2001) and be self-reliant (Gatley, 1992).

About 36% of patients in home care have wounds (Pieper, Templin, Dobal, & Jacox, 1999). The most prevalent wounds are surgical wounds and chronic leg ulcers. In the present research, the focus was on chronic wounds; surgical wounds were excluded. Chronic wounds include pressure ulcers, venous ulcers, arterial ulcers, diabetic ulcers and connective tissue orders (Visco et al., 2001). The majority of ulcers treated by home care nurses are venous ulcers and diabetic ulcers (Friedberg, Harrison, & Graham, 2002).

A leg ulcer has been defined as a defect or break in the skin as a result of the presence of an underlying medical or physiological disorder (Thomas, 1990, in (Roe, Luker, Cullum, Griffiths, & Kenrick, 1993). Leg ulcers are associated with venous disease in approximately 76% of patients and with arterial disease in 22-31% of patients (arterial and venous disease can coexist in a patient). The median duration of ulceration is between six and twelve months (Friedberg et al., 2002).

Leg ulcers can be treated in a hospital or nursing home or by home care. In the United Kingdom, 85% of patients with leg ulcers are mainly treated by the community health service (= home care), while in Scandinavia, 61% of such patients are treated at home (Haram, Ribu, & Rustøen, 2003). Treatment consists primarily of regularly inspecting the wounds and applying compression bandages. Patients who require wound care in the home must be referred by a physician (Ribu et al., 2003), mostly a GP or a dermatologist. These patients have increasingly complex and extensive needs (Baker, 2001).

In population studies from the United Kingdom, Australia, Sweden and Ireland published between 1986 and 2000, leg ulcer prevalence was estimated to be between 0.06% and 0.18% in the general population and up to 12.6% in persons older than 70 years (Friedberg et al., 2002). With the ageing society, the absolute number of leg ulcer prevalence will increase.

1.1.1 Telehealth

Telehealth is the delivery of health care and the exchange of health care information across distances (Oakley & Wootton, 2002), both in location and in time (Lamminen, Voipio, & Ruohonen, 2001). Telehealth involves the use of telecommunication equipment and communication networks for the purpose of transmitting health care information among health care practitioners who are located at different sites (Visco et al., 2001). In the literature, telehealth is often called telemedicine and is defined in various ways. Lamminen et al. (2001) provide an overview of various definitions. Each, however, contains elements of information and communication support to connect people over distances in both time and location. In the present research, telehealth was adopted as the umbrella definition (Oakley et al., 2002):

Definition 2 Telehealth

The delivery of health care and the exchange of health care information across distances and time.

Telemedicine is a specific subset of telehealth (May, Finch, Mair, & Mort, 2005). May et al. provide an overview of the different technologies available in telehealth. They distinguish different modes of technology mediated health care, by taking the traditional synchronous and asynchronous perspective on groupware and combining it with interactivity. This resulted in the categorisation presented in *Table 1-1* (May et al., 2005).

Tele-dermatology (telehealth of wounds) is a sub-specialty of telehealth. Typically, two types of telehealth technology are practised in tele-dermatology, namely videoconferencing and store-and-forward technology (Visco et al., 2001), both of which are categorised as 'telemedicine' by May et al. (2005). Tele-dermatology encompasses consultations between a patient with a skin disease and/or the primary health care provider and a dermatologist for diagnosis and management advice (Oakley et al., 2002). In synchronous tele-dermatology (videoconferencing), the consultation is multi-locational. In asynchronous tele-dermatology (store-and-forward), the consultation is not only multi-locational but also multi-temporal.

Table 1-1 The different modes of technology mediated health care (May et al., 2005)

Modes of technology	**Description**	**Operational contexts**
Synchronous (interactive) Closed-circuit TV/video-conferencing	Mainly telemedicine: Synchronous systems are used to transmit live sound, images and data, with or without parallel videoconferencing. Here, a range of health professionals including nurses can videoconference with each other or with patients, and by using proxies can undertake and monitor a range of clinical tests or monitor vital signs.	Psychiatry/psychology Cardiology Trauma/emergency Respiratory care
Asynchronous (store-and-forward) Email	Mainly telemedicine: Asynchronous (store-forward) system record, store and then forward images and other data for subsequent review. Put simply, clinical data is captured – either by specially modified digital cameras or other medical equipment - and then usually emailed to medical professionals who use it for diagnostic or management decision-making.	Dermatology Radiology Pathology
Synchronous (non-interactive) Remote monitoring	Mainly telecare: Synchronous systems are used to monitor specific potentials (location, vital signs, syncope, movement) using sensors that communicate with a remote base or call centre. Responses to signal may be an automated alarm or telephone call to a carer or family member, or an alert to emergency services.	Home telecare Monitoring safety of vulnerable older people. Surveillance of nursing home residents.
Asynchronous/ synchronous (interactive) Mobile (M-Health)	Mainly self-care: Localised and personalised systems mounted in mobile phones, personal digital assistants, personal computers. May be wireless or hard wired. Perform personal record keeping and calculation of test results (e.g. blood sugar, cholesterol). Can communicate with remote call centres or home base-stations. Can provide data for professional review.	Personal monitoring of chronic conditions (diabetes) cardiovascular and respiratory conditions)

1.1.2 Home care in the Netherlands

Each year in the Netherlands, about 420,000 people receive home care (Ministerie van VWS, 2005), provided by 47,590 health care professionals. On average, in 2004 a patient received 3.71 hours of care each week (Ministerie van VWS, 2005). The number of people receiving home care is increasing, because of the ageing population and because elderly people want to remain in their home for as long as possible. In addition, patients are discharged from hospital sooner than they were in the past.

The home care organisation Thuiszorg Enschede Haaksbergen (now 'Livio') participated in the present project. In general, Livio delivers cure (short-term care, e.g. palliative care and recovery after surgery), care (long-term care, e.g. chronic wounds) and special services (e.g. nutrition advice, occupational therapy) to its customers. At the start of this project (October 2003), 1,389 employees of Livio were providing health services to 1,700 patients. About 200 health care employees were qualified to provide wound care. The number of wound care patients is unknown.

Livio has care centres – or 'teams' – at five locations. Four of these centres are located geographically (i.e. in the north, south and east of the district, and in Haaksbergen), and provide long-term care (for e.g. chronic wounds). The fifth centre provides short-term care (for e.g. surgery wounds and palliative care). This short-term care and the team are called 'Cure'. Their patients are located throughout the district. The care and cure division of Livio identified five topics of specific attention; one of these is diabetes and wound care. Seven district nurses from various centres specialise in diabetes and wound care. These district nurses are called expert district nurses. In addition to performing their normal district nursing activities, they are responsible for disseminating to their team members knowledge and abilities that are related to diabetes and wound care.

District nurses deliver wound care according to the treatment plan drawn up by a physician (GP or dermatologist). When a district nurse visits a wound care patient, she removes the bandaging and cleans the wound, then dresses and re-bandages the leg. In all these activities, a district nurse must be self-reliant (Gatley, 1992). However, when a district nurse needs advice – for example regarding the condition of the wound, the dressings applied or the bandaging technique – her expert district nurse is the obvious person to contact (Hettinga et al., 2003). If there is an urgent situation, a district nurse can also contact the relevant physician. There used to be one-to-one contact between a district nurse and a patient; the same district nurse would visit the patient each time. Nowadays, however, the amount of one-to-one contact is diminishing: the tendency is for several district nurses to visit the same patient during the week. Therefore, the exchange of patient-related information is becoming more important.

1.1.3 Research objective for home care

A district nurse is one of the most mobile persons in health care. She visits patients and provides them with medical and non-medical care in the intimacy of their homes, thus obviating the need for expensive treatments in hospitals or nursing homes. As a mobile outpost of health care, a district nurse can neither quickly nor easily consult colleagues or physicians for advice or a second opinion. In addition, direct access to electronic patient information is not easily available. Therefore, the overall research objective was to establish whether mobile technology can be used to facilitate information access, communication and information exchange between district nurses and between district nurses and physicians. More specifically, the objective was to establish the effect of information and communication support on task performance.

1.2 Railway

The second case study in the present research is railway. Railway is about the transportation of trains, with passengers or freight, along dedicated infrastructure (railways). As railway infrastructure is scarce and trains have limited dynamic capabilities (e.g. long braking distance), operational traffic management is important. Operational traffic management is the system that enables vehicles to run efficiently and safely through a traffic system (Schotanus & Zigterman, 2004). In Europe alone, trains are equipped with as many as six different navigational systems for traffic management (ERTMS, 2005). This section deals only with the Dutch system. It describes the railway infrastructure, the on-board control systems and the train drivers' activities.

The railway infrastructure is divided into sections or 'blocks'. Each block is between one and one and a half kilometres long. The safety control principle is that only one train is allowed in each block at any one time. Blocks are guarded by signals. A signal is a device that provides train drivers with information about the state of the block ahead. A green light means that there is no train present in either of the next two blocks. A yellow[1] light means that there is no train in the next block, but there is one in the block ahead of that. A red light means that there is a train in the block covered by the signal.

The colour of the signals indicates the action to be taken by the train driver. A green light means that the maximum section speed is allowed, a yellow light that only 40 km/h is allowed, and a red light that the train must stop before the signal (in principle, 0 km/h is allowed). In 1962, there was a train disaster in Harmelen. The driver of a passenger missed a red light in heavy fog and collided with another passenger train. Ninety-two passengers were killed and many others were seriously injured (Raad voor de Transport Veiligheid, 2001). This accident led to the introduction of an automatic train protection (ATP) system in the Netherlands.

An ATP system provides both auditory and visual information inside the driver's cabin about the speed limit, as indicated by line-side signs and signals. In principle, if a train exceeds the maximum speed limit, an auditory signal is provided; if the train driver does not reduce speed (he

[1] In practice, the colour is orange, but in the rail sector it is called yellow.

does not respond correctly to the line-side signs and signals), the train is brought to a stop. For example, if a driver passes a yellow signal, he has to slow down to 40 km/h. At that time, a warning is sounded and a cabin light indicates that the speed limit is 40 km/h. If the train does not slow down, it is automatically stopped after a few seconds. This is called an ATP intervention. The disadvantages of the ATP system are that blocks with a speed limit below 40 km/h are not covered, nor are blocks with a temporary speed limits (Schotanus et al., 2004). This means that passing a red light at a speed below 40 km/h is not covered. Passing a red sign is called a SPAD: ‘signal passing at danger’ (Hamilton & Clarke, 2005).

The coexistence of different navigational traffic management systems in Europe is extremely costly and takes up space on board (ERTMS, 2005). A train crossing from one European country to another must switch operating standards as it crosses the border. This adds both to travel time and to operational and maintenance costs. Therefore, a European group of railway experts was set up in 1990 to develop the requirements for a new control command system (ECTS). Together with GSM-R (the new radio system for voice and data communication), it forms ERTMS (European Rail Traffic Management System): the new signalling and management system for Europe that will enable interoperability throughout the European rail network (ERTMS, 2005). In the Netherlands, only international trains will be equipped with ERTMS. All others trains will use the current ATP system, or an updated version of it (ATP Next Generation).

1.2.1 Railway infrastructure

In the Netherlands, ProRail is responsible for the rail infrastructure, railway capacity management and railway traffic control. At the end of 2004, 2,787 people were employed by ProRail (Lucas, 2005). ProRail manages 6,500 kilometres of track and 386 stations. Each day, some 5,000 passenger trains and 230 freight trains are transported. This makes the Dutch rail network one of the densest in Europe (ProRail, 2005).

In general, there are two types of railway infrastructure in the Netherlands: infrastructure with switches and that without switches. The latter is called ‘free route’ and the blocks are guarded automatically by signals. The infrastructure with switches is controlled by railway signalling. For railway signalling, a process plan is developed for all scheduled trains. According to this plan, all train routes from A to B are planned automatically (of course, taking free routes into account). This is called automatic train route creation. The railway signaller has to plan routes manually in the case of delays, disturbances or other unscheduled situations. Moreover, the representation of trains in the control room is very schematic: a train is shown as being in a certain block, but information about its exact position or speed is lacking.

1.2.2 Railway transportation

The largest passenger transport organisation in the Netherlands is the Nederlandse Spoorwegen (NS). The NS schedules over 5,000 train services a day to transport more than one million passengers (NS, 2005). On each train a ticket collector or guard is available for the passengers, and the driver is responsible for driving the train safely and efficiently from A to B. While operating the train, the driver has to respond to line-side signs and signals. Moreover, he uses his route knowledge to operate the train. Route knowledge is defined by the UK rail industry in a rather static way as 'the types of information required to be remembered in order to operate over a route' (Livingstone, Gipson, & Luther, 2005). In the present research, the emphasis was on dynamic information and cognitive aspects, so the following definition of route knowledge was applied:

Definition 3 Route knowledge

Experience-based knowledge gained from many types of information and information sources, required to operate a train safely and efficiently.

During train operations, the driver can communicate with the railway signaller. For example, when a train driver has to stop for a red signal, he must immediately contact the railway signaller, or when a level-crossing barrier has been damaged, a mandatory instruction is provided by the signaller to the train driver, upon which the crossing may be crossed.

1.2.3 Research objective for railway

A train driver operates his train and transports his passengers or freight safely and on time from A to B. The driver has exact information about the train's position and speed, but lacks information about the overall picture, for example the position and speed of other trains, and any upcoming interactions. Moreover, a train driver has only limited decision and control latitude: he can only control the speed and acceleration of the train (both positively and negatively). He can see one and sometimes two signals ahead, without getting information concerning the entire picture. Therefore, the overall research objective was to establish whether information and communication technology can be used to provide context information to train drivers, and to facilitate the interaction between train drivers and railway signallers. More specifically, the objective was to establish the effect of information and communication support on train driving task performance.

1.3 Overall research objective and questions

The overall aim of the present research was to establish the effect of information and communication support on the task performance of distributed professionals, especially as new technology must achieve a balance between the advantages and the costs (cf. (Lamminen et al., 2001). Mobile applications that provide information and communication support are designed to facilitate or improve task performance. Distributed professionals or experts typically perform both individual and team or group tasks. Therefore, the research questions were:

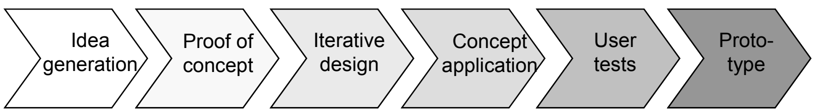
1. What is the added value of information and communication support for distributed professionals on the main task performance indicators?

2. What is the user acceptance and perceived applicability of information and communication support for distributed professionals?

1.4 Research approach

For the present research, two tailor-made mobile applications were developed in order to provide information and communication support to distributed professionals. This kind of research is known as IT innovation research, which consists of several phases and processes. According to Smits: 'Innovation processes are neither linear nor causal and are better regarded as interactive processes in which there is a large extent of co-evolution of scientific, technological and societal systems' (Smits, 2002). *Figure 1-1* is a simplified depiction of the initial phases of innovation research, that is, from idea to prototype. Of course, in practice these phases are not linear: iterations and interactions between the phases apply. For reasons of clarity, however, it is depicted as a linear figure. Moreover, innovations do not end with prototypes. The present research focused on the first phases from idea generation till user tests.

Figure 1-1 Innovation trajectory: from idea to prototype

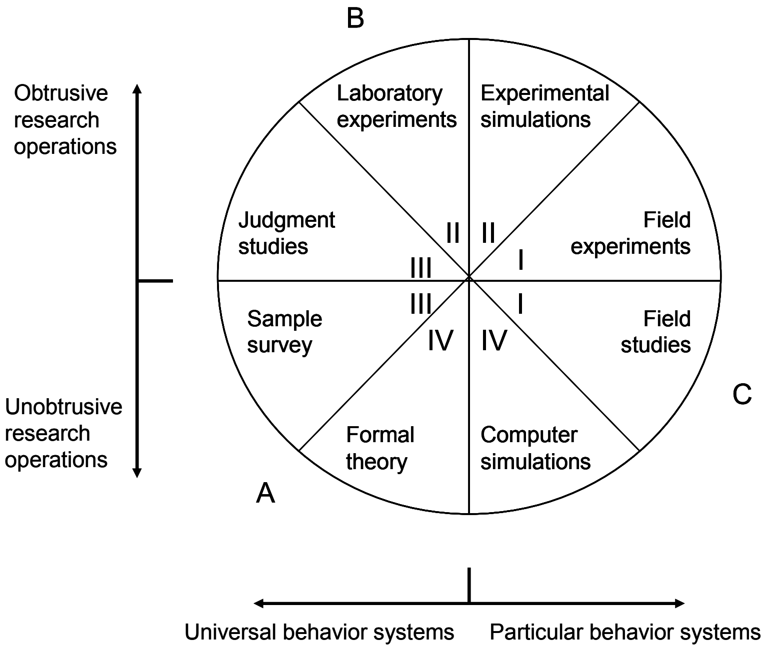


A user-centred design approach is conducted for the phases idea generation, proof of concept, iterative design and concept application. User-centred design focuses on involving users in several stages of design: from initial ideas to introduction in the workplace or market (Beaudouin-

Lafon & MacKay, 2003). In the present research, people at the workplace (i.e. district nurses and train drivers) were involved from the beginning and in all its phases. User-centred design is often an iterative process in which several designs are made and tested.

In the user test phase, research was performed to find evidence to answer the two research questions. According to McGrath (1984), three aspects are important in finding research evidence: generalisability of the evidence over the population (cf. external validity; (De Groot, 1961), precision of measurement of the behaviour (and precision of control over extraneous facets or variables that are not being studied) (cf. internal or construct validity; (Messick, 1989) and the realism of the situation or context (cf. ecological validity; (De Groot, 1961). Unfortunately, all three aspects cannot be maximised simultaneously. Increasing one of them reduces one or both of the other aspects. To gain more insight into this, McGarth developed a research strategies model (McGrath, 1984) (see *Figure 1-2*). The spatial relation between generalisability, precision of measurement and realism of context shows that all three aspects cannot be maximised in one research operation. Moreover, the model also provides insights into the kinds of research operations that can be performed.

Figure 1-2 Research strategies (McGrath, 1984)



Legend

I	Settings in naturalistic systems
II	Contrived and created settings
III	Behaviour not setting dependent
IV	No observation of behaviour required
A	Point of maximum concern with generality over actors
B	Point of maximum concern with precision measurement of behaviour
C	Point of maximum concern with system character of context

This model was chosen as a starting point for the selection of the appropriate research operations for the user tests, and thus determined the kinds of research evidence found.

As the aim of the present research was to develop tailor-made applications for dedicated settings, external validity (A) was not a primary issue. Moreover, tailor-made applications were designed for particular behaviour systems, which refer to the right-hand side of the model. The context is also relevant, so naturalistic settings were preferred. However, the user tests were to be the first tests that took place with the new applications, so their effects were still unknown. This means that testing them in practice could harm wound care patients and train passengers. Therefore, the research took place in less naturalistic settings, that is, in experimental simulations. In practice, this means that realistic settings were simulated by performing authentic tasks in experimental settings. In this way, the focus was on a combination of internal and ecological validity.

1.5 Outline

This thesis is about information and communication support that mobile applications provide to improve the task performance of distributed professionals. More precisely, the research investigated what the added value, the user acceptance and the perceived applicability are of information and communication support. The research was performed by designing and evaluating two broadband mobile applications that provide information and communication support. A user-centred design approach was conducted, and both static and dynamic user experiments were performed.

The first step was to identify the cognitive processes that underlie the task performance of distributed professionals. Observations, task analyses and interviews were conducted in order to guide the design. This resulted in design requirements and an inventory of the performance indicators of home care and railway. A conceptual task performance model was developed that focused on the cognitive processes that are affected by information and communication support. The mobile applications were evaluated with theoretically grounded measures. Central to this were the concepts of task performance, decision-making, empowerment, user acceptance and applicability to practice.

Chapter 2 describes the domain analyses and the design of the mobile applications for district nurses and train drivers. In Chapter 3, a conceptual framework for task performance is provided, on the basis of which assessment instruments were developed (Chapter 4). In Chapters 5 and 6, the user experiments are reported. Chapter 7 provides an overview of and reflections on the research.

Chapter 2

Home care and railway: domain analyses and design of mobile applications

This chapter describes the design path of the mobile applications for home care and railway. The mobile applications provide (1) access to context information, (2) information exchange between distributed professionals, and (3) enriched communication between distributed professionals. Various user-centred methods were applied to this design, namely literature review, semi-structured interviews, observations at the workplace, and workshops. During these design processes, the cognitive constructs of task performance and the domain-specific performance indicators for both home care and railway were identified. Decision-making was the most relevant cognitive construct; the general performance indicators were safety, productivity and health.

2.1 Socio-technical systems

The present research concerned information and communication support for task performance of district nurses and train drivers (distributed professionals). Both application domains are considered complex socio-technical systems. Woods (1988) provides a list of interrelated characteristics that contribute to the different types of complexity in social technical systems. Below, this list is applied to home care and railway; workers' characteristics are human or social aspects, and system characteristics are related to technical aspects. ++ means that the statement is totally applicable to this domain; + means that the statement is applicable to this domain, +/- means a neutral response, - means the statement is not applicable to this domain, -- means the statement is totally not applicable to the domain. This list is summarised in *Table 2-1*.

Table 2-1
Characteristics of complexity in social technical systems (Woods, 1988), applied to home care and railway

Complexity characteristics descriptions	Home care	Railway
Social: *Workers* have to cooperate to let the system work properly	+ +	+ +
Heterogeneous perspectives: *Workers* usually come from different backgrounds	– –	– –
Distributed: *Workers* may be located in different places; social coordination:	+ +	+ +
Hazard: *Workers'* actions can have catastrophic consequences	++	+ +
Disturbances: *Workers* are responsible for dealing with unanticipated events	+ +	+ +
Mediated interaction: *Workers* cannot observe goal-relevant system behaviour directly	+	+
Large problem spaces: *Systems* are composed of many different elements and forces	_ _	+
Dynamic: *Systems* are usually dynamic and can have long time-constants	-	+/_
Coupling: *Systems* tend to be composed of many interacting subsystems	+/_	+ +
Automation: *Systems* are highly automated; workers monitor them	_ _	+/_
Uncertainty: *Systems* tend to provide uncertain data to the workers	_	+/_

Applying these characteristics to home care and railway shows that both domains have typical aspects of complex socio-technical systems. The two domains score equally on worker-related aspects. Railway scores more on system-related aspects than does home care. This is in line with the two extremes of social technical systems: cooperation mediated by technology, and collaborative human and machine control of a technical system.

When designing socio-technical systems, it is relevant to identify the various categories of factors that shape human behaviour in such systems. Knowing these factors provides relevant input for designing useful and usable systems. In general, various system design models are available, that is, normative, descriptive and formative models. Normative models prescribe how a system should behave, descriptive models describe how a system does behave, and formative models describe requirements that must be satisfied so that a system could behave in a new, desired way (Vicente, 1999). The mobile applications that were designed for the present research aimed not at automating existing tasks, but at new ways of working for the professionals. Therefore, a formative approach was applied to guide the design of the mobile applications.

The three goals for the design phase of the mobile applications were:

1. To identify the main performance indicators for task performance supported by the mobile applications.
2. To identify the cognitive constructs for task performance supported by the mobile applications.
3. To guide the actual design of the information and communication support (i.e. mobile applications).

2.2 Performance indicators

The performance indicators safety, productivity and health should be taken into account in designing effective computer-based applications to facilitate work in complex social technical systems (Vicente, 1999).

Safety relates not only to incidents that cause large accidents, but also to those that cause system delays or situations in which task performance is interrupted. Safety often relates to how workers in complex social technical systems can deal with unanticipated events.

Productivity relates to the efficiency and effectiveness of task performance. This is relevant, especially given the company investments required to introduce information technology into the workplace. The intuitive idea is that information technology improves productivity. However, the evidence is not clear on this. Landauer found that productivity decreased shortly after the introduction of information technology in the 1970s (Landauer, 1995). The main reason he provides for this is that the first use of information technology (from 1950-1970) was to automate the most obvious human tasks. A typical example is Fitt's list – a function allocation analysis used to determine whether the human or the system is better suited to perform a certain task. After having automated all obvious tasks, the goal of information technology became to facilitate human tasks with information design. However, Landauer (1995) states that, as productivity has decreased since 1970, the challenge of supporting human capital has not yet been achieved. Here, Landauer is referring to both the usefulness and the usability or ease of use aspects. Useful is defined as 'capable of being used advantageously' (Davis, 1989), and ease of use as 'freedom from difficulty or great effort' (Davis, 1989). He predicts that using user-centred design approaches will increase the effectiveness of new applications.

Health relates not only to the absence of stress and disease, but also to job motivation and satisfaction. The present research focused on the latter aspects. Relevant here is the demand-control model of Karasek (1979), which concerns the psychological job demands and the decision latitude. Decision latitude refers to autonomy and skill discretion: the exploitation of skills on the job. People get more freedom of actions and activities. The combination of low decision latitude and high job demands is associated with job dissatisfaction (and mental strain). The implication of this for the performance indicators is that 'redesigning work processes to allow increase in decision latitude for a broad range of workers could reduce mental strain, and do so without affecting the job demands that may plausibly be associated with organizational output' (Karasek, 1979).

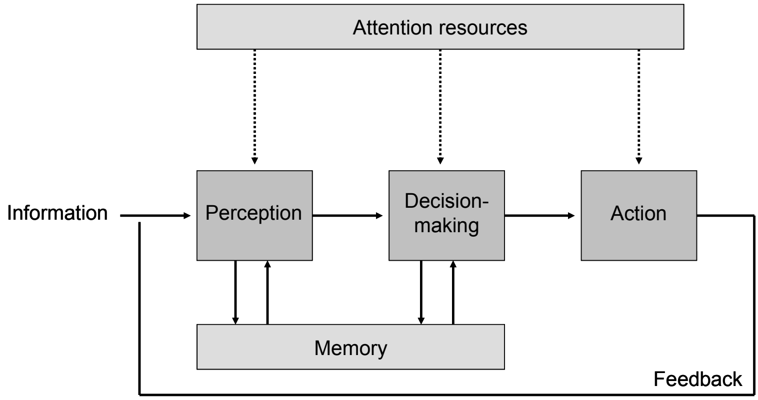
These three performance indicators of complex socio-technical systems were the starting points for the analysis of home care and railway that was carried out to identify their typical performance indicators.

2.3 Cognitive constructs for task performance

Cognitive psychology and human factors research has evolved over time. From the 1950s to the 1990s, the prominent research was related to perception and behavioural theories, for example stimulus-response research (colours and names; Stroop effect) and to information processing theories related to memory and attention, for instance selective attention from Broadbent. In the 1990s, naturalistic theories were developed. Human behaviour was no longer studied in isolation: the effect of naturalistic environments in which the behaviour took place were also included, for example naturalistic decision-making (Zsambok & Klein, 1997). It is likely that trend will continue and that the focus will move to integrated human-system environments, for example cognitive systems engineering (Hollnagel & Woods, 2005). The present research studied distributed professionals in realistic settings; thus naturalistic theories were applied. Of course, also theoretical concepts from prior psychological research (i.e. perception and information processing models) were built upon.

In the literature, various models can be found on the role of cognitive processes in task performance. One of the most well-known models is the human information processing model of Wickens (1992) (see *Figure 2-1*). In simplified form, the model states that information is perceived, decisions and responses are selected, and the response is executed. This is called the perception-decision-action cycle. The responses are again input for the perception (closed loop model). Attention resources and memory are the main human capabilities and limitations that influence this behavioural process.

Figure 2-1 Simplified model of human-information processing (Wickens, 1992)



There are various adapted versions of this model; each concerns a different domain. For example, the air traffic control performance model (Oprins & Schuver, 2003) and the situation awareness model in dynamic decision-making (Endsley, 1995b). The relevant cognitive constructs in these models for task performance are perception and decision-making.

In the present research, all the distributed professionals who performed tasks were qualified experts. Proficient task performance differs from non-proficient task performance, or as Rasmussen observed: 'The efficiency of skilled performance is due to the ability to compose the process needed for a specific task as a sequence of familiar subroutines which are useful in different contexts' (Rasmussen, 1976). Therefore, expertise is also a relevant construct for task performance.

The mobile applications that are developed should be incorporated into the daily practice of the professionals. Concepts related to the use of these applications are relevant. In the literature, user acceptance refers to both usefulness and usability (Davis, 1989).

The mobile applications can provide not only relevant context information to the professional, but also opportunities for communication. Therefore, cognitive aspects should focus not only on the individual professional, but also on the interactions between the professionals (Malone, 2004).

The previous section on performance indicators showed that decision latitude is also a relevant cognitive construct when redesigning work processes (Karasek, 1979). Taking all these constructs together, the following preliminary model was developed to identify the cognitive constructs of task performance (see *Figure 2-2*).

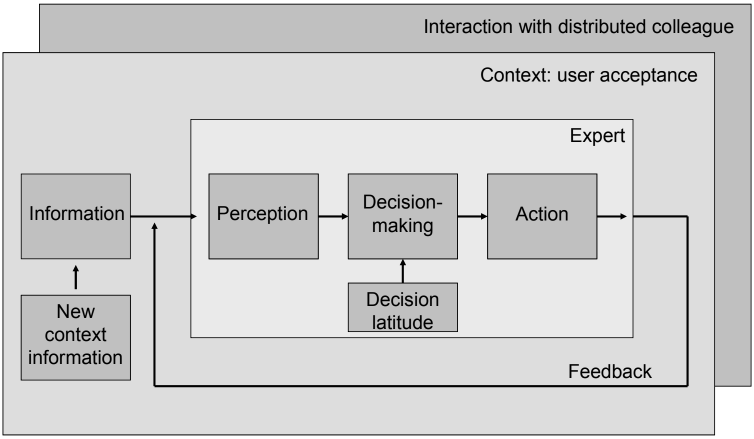


Figure 2-2 Preliminary model of cognitive constructs for task performance of distributed professionals

The model depicts the information processing constructs perception, decision-making and action. Professional task performance adds to the constructs of decision latitude, expertise and user acceptance. The mobile applications aim at adding context information, and communication or interaction between the distributed colleagues. In relation to the human information processing model of Wickens (see *Figure 2-1*), the concepts of attention resources and memory are excluded. Of course, these concepts are relevant to task performance, but were outside the scope of the present research.

The two application domains are analysed in sections 2.5 and 2.6. This task performance model was used to identify the relevance and typical existence of these cognitive constructs in each application domain.

2.4 User-centred design approach

The characteristics presented by Woods (1988) indicate that designing socio-technical systems is a complex matter. Such systems – in which activities are only meaningful in relation to their functional contexts, and to the opportunities and constraints represented by the work domain – require an extensive analysis approach (Bye et al., 2005). Such an approach should focus not only on human-machine interactions aspects, but also on an integrated analysis of the human-machine context. For the design of the mobile applications, structured interviews were conducted with potential users. The interview schemes were based on the UseIT model.

The UseIT model is dedicated to revealing user satisfaction with IT support in health care (Michel-Verkerk, Schuring, & Spil, 2003). For railway, an adjusted version of the interview scheme was developed and applied. The

UseIT model helps to identify the most urgent and relevant problems in health care, and supports decisions about which processes should be supported by IT and how. It consists of four determinants that have to be balanced in assessing the diffusion and use of information systems. These determinants are relevance, requirements, resistance and resources (Michel-Verkerk et al., 2003). Two of these determinants (relevance and resistance) stem from worker-related aspects, while the other two (requirements and resources) stem from system-related aspects.

Relevance is defined as ‘the degree to which the user expects that the IT system will solve his problems or help to realise his actually relevant goals’. Requirements are defined as ‘the degree to which the user needs are satisfied with the product quality of the innovation’. Resistance is ‘the personal attitude of all stakeholder groups towards the introduction of an information system’. Finally, resources are defined as ‘the degree to which material and immaterial goods are available to design, operate and maintain the information system’.

The following sections provide the design of the interview schemes – based on the UseIT model – and the results of the interview sessions. These results guided the design of both mobile applications. Moreover, during the interviews, observations and workshops, the performance indicators and the cognitive constructs of both domains were identified.

2.5 Home care analysis

The starting point of the home care project was that district nurses who are working on their own at a patient’s home could be supported by technology to communicate with other health care providers in order to improve her wound-care-related task performance. Therefore, a literature review, observations at the workplace, semi-structured interviews and a workshop were conducted to identify (1) what cognitive processes could be supported (focus), (2) how they could be supported (system design), and (3) what the effect will be on task performance (performance indicators). First, the semi-structured interviews based on the UseIT model are described.

2.5.1 User interviews based on the UseIT model

Semi-structured interviews were conducted with relevant stakeholders in wound care, namely dermatologists and district nurses from the diabetes and wound care section (Hettinga et al., 2003). The aim of these interviews was to identify general working practices in home care, the user- and system-related aspects of technology support in their working practices, and ideas related to technology support in their working practices. Semi-structured interviews are interviews that are based on a framework of questions that elicit open answers. The interview scheme was based on the UseIT model.

To identify the relevance, the health care professionals were asked about their current working patterns, tasks, responsibilities, problems, training, quality of work, and so on. For requirements, the health care professionals were asked for their ideas about extra information and communication that would support their work processes. The general idea of the communication and information exchange was proposed and their reactions and opinions were listed. Moreover, detailed questions about visual information exchange were posed, as this is a crucial factor for tele-dermatology (Wootton, 1999). Next, the health care professionals' attitude (resistance) towards IT in general, and towards the general description of functionalities of the mobile application in particular, were asked. Finally, to identify resources, issues were raised about the IT and IT support that was available in their organisations.

All four determinants were operationalised in the design of the interview scheme. The final scheme was composed of three phases, namely (1) the inventorying of current work practices, (2) the proposing of general ideas related to new information and communication support, and (3) the inventorying of reactions and extensions to new information and communication support opportunities.

In total, nine health care professionals were interviewed, that is, all seven expert district nurses from the diabetes and wound care section, and two dermatologists. Each interview took about 45-60 minutes. The series of interviews started with the idea of designing a mobile application to support bandaging. Bandaging affects blood circulation, which is an important aspect of wound care. Good blood circulation improves the healing of a wound. In this district, a possible problem was indicated with the quality of bandaging. After two or three interviews, it became obvious that bandaging would not be an appropriate subject for information and communication support: the pressure of the bandage on the leg cannot be conveyed properly, and bandaging takes too much time (20 minutes per patient). In the meantime, it was identified that wound care was a relevant topic for IT support.

The district nurses indicated that the most important 'relevance' results were the need for more patient- and wound-related information in the patient's home, and being interrupted by colleagues for advice during treatments. Dermatologists expected most relevance from teleconferencing, as it would reduce the hospital visits the elderly person had to pay.

The main 'resistance' results indicated that the elderly district nurses had no affinity with computer applications. Moreover, the expert district nurses expected that some patients would be reluctant to accept the use of computers in their home environment. Dermatologists asked who was going to finance the tele-consults and queried certain legal aspects, for example who would be responsible if a wrong diagnosis were made during a tele-consult?

They all expected that the helpdesk of their own organisation would provide the necessary resources and support for the application.

The main requirements identified by the district nurses were the extra information available when advice had to be provided, wound information in a standard format, store pictures and videos for education. The dermatologists also wanted to store the pictures and videos. They emphasised a standard format of delivering wound information and the importance of the quality of the colour of the picture. For example, red might indicate that the wound is healing, while a slightly different colour red might indicate that it is an angry wound. In general in health care, the privacy and security of patient-related information is relevant.

2.5.2 Home care system design

The results of the interviews were discussed in a feedback session with five of the participating expert district nurses (Hettinga et al., 2003). In a 'brown paper' session, the expert district nurses got two usage scenarios for the mobile applications (one synchronous and one asynchronous) in which they had to indicate their main requirements for the application. *Table 2-2* summarises the main requirements of the user interviews and the feedback session.

Table 2-2 Results of the user interviews and feedback session: description of the mobile application

A district nurse gets a laptop and camera to visit her patients. Using the application, the district nurse will be capable of doing the following items, all in a user-friendly way
• Seeing the presence, location and availability of medical professionals for emergency situations
• Starting a teleconference with an available expert
• Recording pictures and video material during teleconferencing
• Writing messages to experts, including pictures, videos and audio material
• Documenting the patient's wound status with text, pictures, videos, and audio material (also from teleconferencing and messages)

Based on these high-level requirements, a detailed design was made for the application. As one of the features is a kind of weblog of wounds for documenting the wounds, the application is called WoundLog.

2.5.3 Design of WoundLog

The aim of WoundLog is to allow district nurses who are at a patient's home to exchange information and communicate with a dermatologist, GP or expert district nurse. The medical teleconferencing and log service WoundLog provides a wound logbook for wound history information and communication tools for both synchronous and asynchronous consultations, that is, teleconferencing and multimedia messaging.

Teleconferencing systems are often medium- or service-oriented (e.g. netmeeting). Participants in a conference can tailor their set of media for a conference. As various district nurses visit the same patient, district nurse based (i.e. participant based) is not an adequate design solution. The interviews and workshops with district nurses revealed that the major task they perform is taking care of the patient (Hettinga et al., 2003). Therefore, the patient was chosen as the unit of design. Thus, instead of regular tailorable participant-based applications, WoundLog is a dedicated patient-based application. The most relevant patient-related information (i.e. history information and results of multimedia messaging and teleconferencing) is always available in the wound logbook, on the left-hand side of the application.

The various tasks a district nurse performs were taken as the units of design and are positioned on the right-hand side of the application, presented as task-based patches. The design solution of task-based patches is in accordance with the way people naturally search for information, namely patch based (Pirolli & Card, 1999). The tabs on the right side of the window present all relevant information and communication support related to a certain task. These tabs/tasks are 'insert new wound logbook information (reports)', 'teleconferencing', 'multimedia messaging' and 'taking pictures' (Biemans, Swaak, Hettinga & Schuurman, 2005). *Figure 2-3* depicts an overview of WoundLog.

Figure 2-3 Overview of WoundLog

The wound logbook is the central element in WoundLog. For each patient, the status information and relevant communication about the wound's recovery process is always available. The most recent information is on top. It is marked when a district nurse has read the information in the wound logbook. Each of the four task-based tabs is described in detail in the following paragraphs.

To start a teleconference, the district nurse selects the person she wishes to contact (see *Figure 2-4*). She gets an overview of all possible candidates, ordered by function, and their current presence, availability and location status. She selects the relevant expert and starts a teleconference.

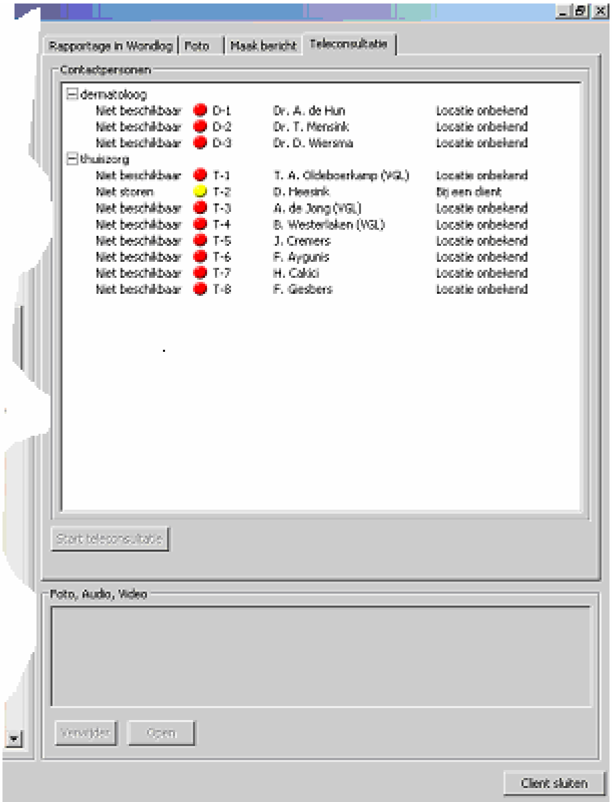


Figure 2-4 Overview of starting a teleconferencing (left half of the screen is omitted= wound logbook)

In the meantime, she focuses her camera on the wound and judges from the image on her screen whether this is the appropriate view for the expert. When the expert accepts her call, they share the patient's wound logbook, the camera view and have audio communication (see *Figure 2-5*). Each participant in the conference can make 'personal or local' photos of the video. These thumbnails are presented at the bottom of the teleconference; they are not automatically shared between the participants.

Figure 2-5 Overview of video sharing in a teleconference (left half of the screen is omitted= wound logbook)

To send a message/multimedia message, the district nurse selects the addressees from a list. She fills in the subject of the message, the report form with predefined relevant wound information and the patient's physical condition, and her question to the addressee. She can also add some wound pictures to the message. *Figure 2-6* depicts an overview of the multimedia messaging report form with pictures. A copy of the message is automatically stored in the patient's wound logbook. The answer to this message is also automatically stored in the wound logbook. Irrespective of the day of the answer, this message is stored next to the question, as it is about the same information.

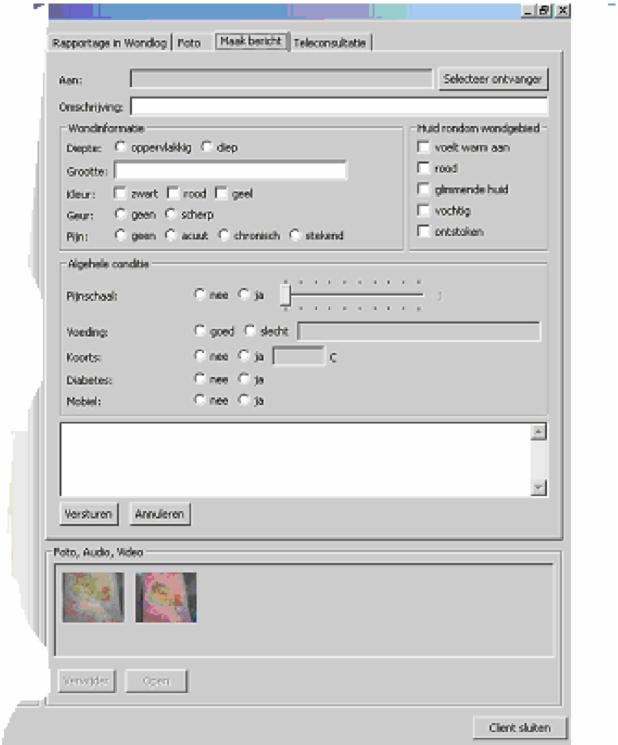


Figure 2-6 Overview of WoundLog multimedia messaging and filling in report form. The left-hand side of the screen (wound logbook) is omitted.

To report on the patient's status and the progress of the wound, the district nurse can store a wound report accompanied by relevant pictures in the wound logbook. The wound report is exactly the same form as the wound report information of the messages: predefined items about the wound and the patient's physical condition (see *Figure 2-6*).

To take a picture, the district nurse takes her camera and starts a local video. When the appropriate information is on the screen, she can push the 'take a picture' button and a thumbnail of the picture is presented at the bottom of the window. All pictures taken in this patient-based session are stored in the picture bar at the bottom and are available across all tabs. The district nurse can decide to use these pictures for all kinds of actions, for example, she can add them to a message or to the wound logbook. She can also delete pictures from this bar. *Figure 2-7* depicts the 'take a picture' task tab on the right-hand side of the WoundLog window.

Figure 2-7 Overview of WoundLog 'take a picture'. Left-hand side of the screen (wound logbook) is omitted.

For practical reasons, only one district nurse conducted a cognitive walk-through (Nielsen, 1993) with an initial design of WoundLog. She had to perform several tasks – for example 'make report in wound logbook' – and answer several questions, for instance 'What about the patient's condition last week?' and 'Did the dermatologist give an answer to your question?' While performing these tasks, she was able to navigate adequately between the four task-based tabs. As a result of this cognitive walk-through, some adjustments were made to the terminology used.

The effect of the use of history information in WoundLog is measured by the treatment decisions taken by district nurses. When a district nurse visits a patient at home, she has to assess the patient's situation in order to be able to treat him or her properly. The interviews identified the following treatment decisions and corresponding actions of the district nurse (Hettinga et al., 2003):

- normal situation → treatment according to plan
- semi-urgent situation → make note and contact expert colleague or physician later on
- urgent → immediately contact physician

Treatment decisions refer to the use of WoundLog, that is, 'semi-urgent' refers to the use of multimedia messaging and 'urgent' refers to teleconferencing. District nurses indicate that treatment decisions are based on the patient's medical background (e.g. diabetes, cause of wounds) and current situation (e.g. wound colour, fever, nutrition, pain level).

A skilful treatment decision is crucial for the quality and efficiency of the delivered health care. If the care does not have the desired effect or if the patient also suffers from something else, the sooner the treatment plan is adjusted, the sooner the patient will be cured. District nurses are not allowed to adjust the treatment plan; only dermatologist and GPs are. Therefore the decision of the district nurse to contact the physician is crucial to the patient's health care, or as Perednia says: 'it is relevant to let a dermatologist look at the right skin problem at the right time' (Perednia, 2002).

2.5.4 Home care performance indicators

The interviews and feedback session with the expert district nurses resulted in usage scenarios for using WoundLog. From the points of view of all parties involved, these scenarios implied the following benefits (Hettinga et al., 2003):

- Patients: fewer visits to hospitals, fewer volunteers necessary for transportation, increased motivation through pictures of wound healing process.
- District nurses: coaching, education and control of their wound activities, that is, bandaging technique, sterility, following wound healing process, quality of decision-making (treatment decisions).
- Dermatologists: immediate supervision of education, control and refresh courses of district nurses, and better understanding of treatment effects. Get a grip on home care situations and cost savings in wound healing products.

The main health care performance indicators for using WoundLog were deduced from these person-oriented perspectives, from the results of the interviews and feedback session, and from the literature. The starting point was the general performance indicators for effective computer-based information systems for working in complex socio-technical systems, that is, safety, productivity and health (Vicente, 1999). The indicated health care performance indicators were quality of care, efficiency of care, patient satisfaction and job satisfaction (service orientation). Relevant elements of each construct are described below:

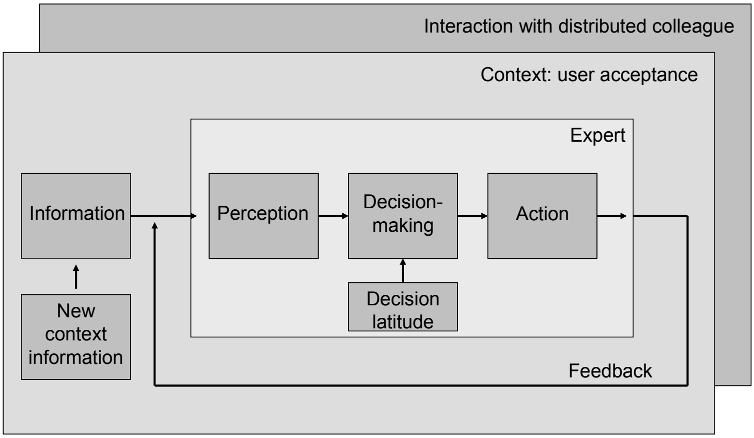
- Quality of care: quality of clinical decisions (physician), quality of treatment decisions (district nurse) and earlier adjustments of treatment plan, reduced wound healing time (Visco et al., 2001), and enhanced patient knowledge and self-care skills (Baker, 2001).
- Efficiency of care: improved efficiency and pace of treatment, improved productivity of district nurse, earlier adjustments of treatment plan, reduced wound healing time (Visco et al., 2001), decreased costs of wound healing products, decreased number of physician visits, re-hospitalisation and emergency department visits (Baker, 2001).
- Patient satisfaction[2]: better wound healing process, involvement and motivation of patient in healing process, decreased number of physician visits, re-hospitalisation and emergency department visits (Baker, 2001), and fewer volunteers necessary for transportation.
- Job satisfaction: improved decision-making (treatment), coaching, education, control and feedback on their wound activities, empowerment, fewer extra afternoon patient visits for second opinion, improved cooperation between colleagues.

2.5.5 Home care cognitive constructs

In section 2.3, the general constructs for task performance of distributed professionals were identified based on the literature. These constructs were applied to home care, and the relevance and typical existence of these constructs were described. To provide an overview of these aspects, first *Figure 2-2* is repeated (see *Figure 2-8*). All constructs of task performance are then described in the home care context.

[2] A literature study by Patistea et al. revealed that caring meanings and behaviours are context specific and they suggest that nurse clinicians should determine caring needs and actions in the context of practice: the provision of individualised care is necessary (Patistea & Siamanta, 1999).

Figure 2-8 Preliminary model of cognitive constructs for task performance of distributed professionals (repeated *Figure 2-2*)



The relevant information for district nurses in wound care is the condition of the patient (e.g. fever, appetite), wound area (e.g. angry), wound colour (e.g. black, red, yellow) and case description (e.g. diabetic, vascular problems).

The new context information is history information about the wound in the format of a wound logbook, that is, pictures and registration of the above-mentioned information on a regular basis.

Perception relates to the integration of all the information sources, based on the district nurse's knowledge (wound and healing processes) and expertise. The result of this process is a prognosis of sorts of the patient's healing situation on which her treatment decision is based.

District nurses are not allowed to make clinical decisions, that is, decisions about the kind of disease that the patient has. They are, however, allowed to make treatment decisions (i.e. whether the patient can be treated according to the treatment plan drawn up by the physician, or whether an intervention is needed). The degree of time pressure varies; for example, the toes of a diabetic or a person with vascular problems can mortify within a couple of hours or days without proper intervention. At this moment, the following treatment decision can be made, with corresponding actions: treat according to plan, immediately contact physician or (an intermediate decision) contact expert district nurse or physician later on (before next treatment).

District nurses believe that history information provides them with enhanced opportunities for decision-making; their decision latitude is increased. As an example they think that they can use the wound logbook to convince the GP to prescribe another wound dressing.

District nurses are qualified experts who possess various kinds of theoretical and practical knowledge. However, wound care experiences differ between district nurses, depending on the number of patients requiring wound care. It is likely that this experience affects treatment decision-making; certain information items trigger a district nurse to make decisions. For example, 'when a diabetic has a fever and no appetite, the condition gets worse'.

WoundLog provides opportunities for synchronous and asynchronous interactions with various health care professionals. The choice of the interaction depends of course on the urgency of the wound condition; however, it also depends on other factors, for example the availability of people and hierarchical relations.

Finally, the use of the application in practice is also a relevant cognitive construct. User acceptance depends not only on usability items, but also on the perceived usefulness of the application. User acceptance is especially important for district nurses as technology use is not widespread in this profession and many district nurses are no longer young.

These aspects form the conceptual basis for the task performance of distributed professionals supported by mobile technology. The further development of the conceptual model is described in Chapter 3.

2.6 Railway analysis

The starting point for the railway project was that train drivers could be supported by context information to improve their task performance and to facilitate interaction between them and railway signallers. A literature review, observations at the workplace, semi-structured interviews and a workshop were conducted to identify (1) what cognitive processes could be supported (focus), (2) how they could be supported (system design), and (3) what the effect will be on task performance (performance indicators). First, the semi-structured interviews based on the UseIT model are described.

2.6.1 User interviews based on the UseIT method

Whereas the home care project was aimed at designing a new application to connect individual district nurses to other health care providers, the railway project was aimed at enriching the current situation with more context information so as to improve the train driving. In the home care project, a new application was designed, while in the railway project a new display was designed to be used in conjunction with existing information in the driver's cabin.

To identify requirements for the new display, an adjusted version of the UseIT method was used, as the UseIT model was originally developed for health care. The four main determinants were applied to railway. For relevance, railway personnel were asked about their current working practices, communication patterns, division of roles, tasks, responsibilities, problems and quality of work. To identify the requirements, train drivers were asked for their ideas concerning extra information and communication to support their work processes. Next, the general idea of providing train drivers with context information was proposed and their reactions and opinions were listed. Moreover, detailed questions were asked about the effect of context information on the railway performance indicators. The attitude of the railway personnel towards context information support in general, and their attitude towards the general description of functionalities of the context information in particular, were asked in order to identify the resistance. Finally, for resources the focus was on the availability of support for the technical system.

All four determinants were operationalised in the design of the interview scheme. The final interview scheme was composed of three phases, namely (1) the inventorying of current work practices, (2) the proposing of general ideas about new information and communication support, and (3) the inventorying of reactions and extensions to new information and communication support opportunities.

A total of eight railway employees were interviewed, namely six train drivers and two signallers. The length of the interviews varied considerably, due to the amount of time each interviewee could spare (the interviews had to be conducted while they were at work).

The train drivers and railway signallers indicated that the most important results were related to communication. Since the 1990s, when the European Commission forced the Dutch railway company to split up into two separate organisations (NS and ProRail), there have been problems with communication between people from the two companies. In practice, too many people need to communicate to get something done, and these people come from different companies, work at different locations and have different interests. In particular, communication between train drivers and railway signallers is not optimal, because they tend to interpret procedures differently.

At the moment, railway signallers have access to general information about the location and speed of trains. They would like to have more detailed information when a train stands still (requirements). The requirements of the train drivers are related to specific conditions: route knowledge is sufficient in normal conditions, while in thick fog, signal information is required. A general requirement was that the extra display should not distract them; they have to look outside for the signals, and are not allowed to rely on signal information on a display.

The train drivers' responses to the idea of additional information in their cabin varied. Some felt threatened (resistance); the context information would replace their knowledge, experience and route knowledge. However, others reacted positively and thought that the information would be useful for driving smoothly and in thick fog. Railway signallers had fewer concerns; they mainly said that train drivers should follow the safety procedures.

Nobody expected any problems with support for the technical systems (resources).

2.6.2 Railway system display design

The results of the interviews and the preliminary ideas about the context information were discussed in several follow-up workshops with train drivers, railway signallers, e-learning specialists, trainers and researchers (Van der Velde, Huppertz, & Biemans, 2004). In general, three kinds of railway scenarios were categorised: normal, abnormal and emergency.

- Normal scenario: normal timetable is applied; almost all trains are on time (within 4-5 minutes delay), exceptions of a few trains are allowed.
- Abnormal scenario: adapted timetable is applied; most trains are behind schedule.
- Emergency: an accident has occurred, several tracks are out of service and new timetables have to be made.

Train drivers indicated that context information was most useful in normal and abnormal scenarios. In cases of emergencies, railway signallers follow procedures to apply new timetables. Thus, the focus of the present research was on normal and abnormal scenarios, and excluded emergency cases. The relevant context information elements for both normal and abnormal routes were identified and ordered. Train drivers distinguished between relevant context information about their own train and that about other trains. *Table 2-3* lists the identified context information elements for train drivers.

Table 2-3 Results of the user interviews and workshops: context information elements for train drivers.

Context information elements	
Information about own train	Information about other trains
Upcoming signals	Upcoming signals
Path predictions[3]	Path predictions
Punctuality	Punctuality
	Speed
	Order of trains
	Type of train (train number)

These high-level descriptions formed the basis for the detailed design of the context information display in both experiments. Depending on the specific research questions and their corresponding information conditions, the dedicated context information was tailor made.

2.6.3 Railway performance indicators

The starting point for the identification of the railway performance indicators was the general performance indicators for effective computer-based information systems for working in complex social technical systems, that is, safety, productivity and health (Vicente, 1999). The railway performance indicators that were likely to be affected by context information were identified in a workshop with stakeholders (Aris & Van den Dikkenberg, 2003). During some brown paper sessions, the stakeholders identified the possible effects of information support. The effects were also discussed during the user interviews with the train drivers and railway signallers. By combining these results with the general performance indicators, the following railway indicators were identified: safety, punctuality, energy costs, job satisfaction and service.

- Safety: safety on track, as a result of train driving (external factors like car accidents are excluded)
- Punctuality: arrival and departure punctuality
- Energy costs: accelerating and braking trains
- Job satisfaction: better utilising knowledge and capabilities of train drivers (cf. empowerment)
- Service: combination of punctuality, informing passengers and driving comfort.

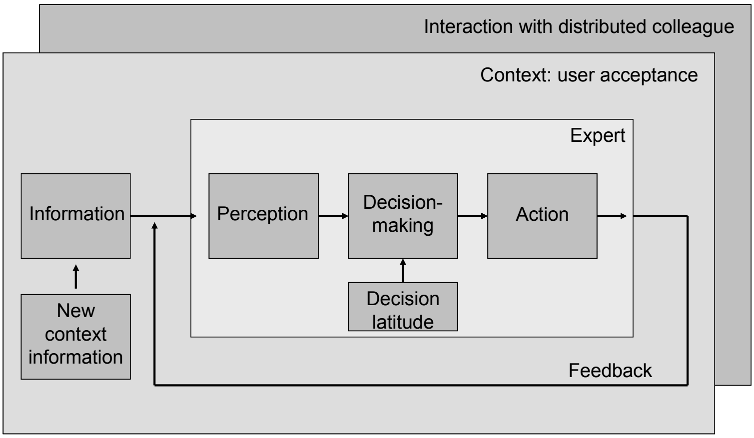
[3] In practice, route knowledge covers path predictions of their own train. As in the experiments route knowledge was not available, these information elements have to be provided to cover the lack of route knowledge.

It was also expected that the communication between the train drivers and the railway signallers would change when context information support was provided. Train drivers would be able to see what is happening, so they would not have to ask the signaller for information (or they would be able to contact the signaller to provide suggestions).

2.6.4 Railway cognitive constructs

In section 2.3, the general constructs for task performance of distributed professionals were identified based on the literature. These constructs were applied to railway, and the relevance and typical existence of these constructs were described. To get an overview of these aspects, first *Figure 2-2* is repeated (see *Figure 2-9*). All constructs of task performance are then described in the railway context.

Figure 2-9 Preliminary model of cognitive constructs for task performance of distributed professionals (repeated *Figure 2-2*)



The relevant information for train drivers comes from various sources, namely outside information (signs and signals), on-board systems (ATP, speed) and information from railway signallers (temporary orders). The information is dynamic and can change rapidly. Therefore, timing is relevant and time pressure exists.

The new context information provides an overview of the current and future general context of the train; track and train information in the driver's near context, for example position, speed and punctuality of trains, path predictions and order of trains.

Perception relates to the integration of all these various information sources, combined with route knowledge, timetable, procedures and expertise of the train driver. The result of this process is a prognosis of sorts of the future situation on the track, on which the driver's driving decision is based.

The context of the train driver is a dynamic situation, in which time pressure is high, multiple competing goals exist (safety, punctuality, energy costs, service orientation) and decisions have to be made in split seconds. Such naturalistic settings constrain typical decision-making activities.

Train drivers vary in their ideas that context information provides them with enhanced opportunities for decision-making. Some believe that the context information will force them to work in a certain way. Others think that the context information will be useless because such information can be derived from their signals combined with their route knowledge. Yet others think that it will increase their decision latitude.

The actions a train driver can take are rather limited. In general, he can only control the train's speed; its direction is determined by infrastructure and planning. The train driver can also communicate with the signaller.

Train drivers are qualified experts, often with a lot of experience. It is likely that train driving benefits from that; certain information items trigger the drivers to make certain decisions.

Context information provides opportunities for train drivers to see other trains and their planned movements. This might introduce interactions with railway signallers for planning suggestions (e.g. 'Take the freight train off my track').

Finally, the use of the context display in practice is also a relevant cognitive construct. User acceptance depends not only on usability items, but also on the perceived usefulness of the application. Perceived usefulness is especially important as train drivers have varying ideas about the usefulness of the display before it has been installed and the display information has to be integrated with other information that is already available in the cabin.

These aspects form the conceptual basis for task performance of distributed professionals, supported by mobile technology. The further development of the conceptual model is described in Chapter 3.

2.7 Cross case analysis

In the previous sections, the two application domains – home care and railway – were analysed. The focus was on the design of the mobile applications, the identification of performance indicators, and cognitive constructs for the task performance of distributed professionals. In this section, the different analyses are compared in order to identify the relevant differences between the application domains. The comparison is of 1) the design of the mobile applications, 2) the performance indicators, and 3) the cognitive constructs.

2.7.1 Design of the mobile applications

The main difference between the two designs is the role of the mobile application in the socio-technical system. In home care, a mobile application was developed that facilitates the collaboration between professionals, while for railway, a display was developed that enriches the information situation. This is in line with the different types of socio-technical systems, namely technology mediated cooperation, and collaborative human and machine control of a technical system. In relation to the information, for home care the current and the history information are relevant, while for railway the current and the future information are relevant.

For home care, the hypotheses were that the history information supports the individual district nurse in treatment decision-making, and that videoconferencing supports the cooperation between the health care professionals, especially in combination with the history information. For railway, the hypotheses were that the context information supports the individual train driver in making driving decisions, and that the interaction between the train driver and railway signaller is improved, as the driver generally has the same information as the signaller.

2.7.2 Performance indicators

The starting point for the identification of the performance indicators were the general performance indicators for socio-technical systems (safety, productivity and health). The domain-specific performance indicators were inferred from them. The two domains of the present research differ from other socio-technical systems in their customer orientation: home care concerns patients and railway concerns passengers.[4] This makes it necessary to extend the general performance indicators of Vicente with a fourth category, namely service orientation. Service orientation is related to the effect of the mobile application on the customers. *Table 2-4* provides an overview of the relation between the various performance indicators. Notice that railway has two different (and sometimes contrasting) performance indicators that are related to productivity.

[4] Freight trains were excluded from the experiments.

Table 2-4 Performance indicators.

Performance indicators.		
	Home care	Railway
Safety (Vicente ,1999)	Quality of care	Safety
Productivity (Vicente ,1999)	Efficiency of care	Punctuality
	---	Energy costs
Health (Vicente ,1999)	Job satisfaction	Job satisfaction
Service orientation (the present research)	Patient satisfaction	Passenger satisfaction

2.7.3 Cognitive constructs

The general model of cognitive constructs for the task performance of distributed professionals (see *Figure 2-2*) was taken as the starting point for the identification of the typical cognitive constructs for the two domains. *Table 2-5* provides an overview of all typical elements identified in the two domains.

Table 2-5 Overview of the cognitive constructs.

Cognitive constructs		
	Home care	Railway
Information	• Various sources of information • Rather static information	• Various sources of information • Dynamic information
Context information	• Current information • History information	• Current context • Future context
Perception	• Integration of various information sources • Wound knowledge • Healing experiences • Treatment plan • Prognosis	• Integration of various information sources • Route knowledge • Procedural knowledge • Timetable • Prognosis
Decision-making	• Treatment decisions • Static information • Time pressure	• Driving decisions • Dynamic information • Time pressure • Competing goals • Timing
Decision latitude	• Positive feelings	• Mixed feelings
Action	• Treatment according to plan • Immediately contact physician • Intermediate: keep close eye on	• Limited degrees of freedom: speed control • Communication with railway signaller

Cognitive constructs		
	Home care	Railway
Expertise	• Most district nurses have many years of general experience • Wound care experience varies	• Most drivers have many years of experience
User acceptance	• Rather unfamiliar with technology at work → usability	• Context information has to be integrated in existing information → usefulness
Interaction with colleagues	• Asynchronous: documentation of wound and patient information (wound logbook) • Synchronous: videoconferencing with physician or colleague district nurse	• Mainly with railway signallers (different views)

This table shows that the main differences in the cognitive constructs between home care and railway are related to the information provided, decision-making, familiarity with technology, and the interactions with colleagues.

2.8 Summary

Two domain analyses were conducted to guide the design of the mobile applications to provide information and communication support for home care and railway. The methods used in the analyses were semi-structured user interviews (based on the UseIT model) and workshops. For home care, a medical teleconferencing application was developed that provides both synchronous and asynchronous ways of communication and information transfer. Moreover, the application includes a wound logbook. Therefore, the application is called WoundLog. For railway, context information elements were identified to enrich the information context of the train driver's cabin. The communication means for train drivers and railway signallers remained the same. The general assumption was that the designed information and communication support would improve the task performance in home care and railway.

To make this assumption testable, the main domain-specific task performance indicators were identified. The starting point was the general performance indicators in socio-technical systems according to Vicente (1999), namely safety, productivity and health. Service orientation was added as a domain-specific indicator for both cases.

To gain insights into how information and communication support can improve task performance, the relevant cognitive constructs for task performance in home care and railway were identified. The starting point was a general information processing model. The main cognitive constructs are decision-making, perception and decision latitude.

Two domain analyses were applied to guide the design of information and communication support to improve task performance of distributed professionals in home care and railway. To test this assumption, domain-specific performance indicators and the main cognitive constructs were identified. The assumptions were tested by applying the designs to several user experiments. Before the experiments could be conducted, a conceptual cognitive model of task performance had to be developed (Chapter 3). Moreover, measurement instruments needed to be selected, adapted and designed to assess task performance (Chapter 4).

Chapter 3

Conceptual framework for task performance

In this chapter a conceptual framework is presented for task performance with mobile applications. First, a general description of task performance is given, with an emphasis on single- and multi-user tasks, the process and the outcome of task performance, and task complexity. Decision-making is the most relevant construct for the task performance of distributed professionals. The focus is on decision-making strategies, structures and process-oriented theories. General theories are presented for proficient decision-making in real-life settings, extended with concepts from dedicated home care and railway studies. Empowerment is also a relevant construct in relation to information support for task performance. Information support empowers professionals to better use their knowledge and abilities for task performance. Finally, the conceptual decision-making framework for task performance that guided the user experiments for home care and railway is presented.

3.1 Task performance

The present research concentrated on task performance in real-life settings. Wood provides a general description of tasks: ‘all tasks contain three essential components; products, (required) acts and information cues’ (Wood, 1986). Information and acts are input components and productivity is the measurable result of acts (output component). Zigurs and Buckland give a comparable definition for group tasks, except they extend it with goal-based behaviour. The present research focused on expert task behaviour (district nurses and train drivers) that is goal-directed. In the present research, the behavioural perspective of Zigurs and Buckland (1998, p.316) was adopted on the general construct of both single- and multi-user tasks, which resulted in the following definition of 'task':

Definition 4 Task

The behaviour requirements for accomplishing stated goals, via some process, using given information..

This leads to the following relevant aspects for both individual and group tasks. (1) Tasks are about explicit goals, (2) tasks are based on provided information, (3) tasks are related to human aspects (behavioural requirements), (4) a process is necessary to perform tasks, and (5) tasks relate to measurable output

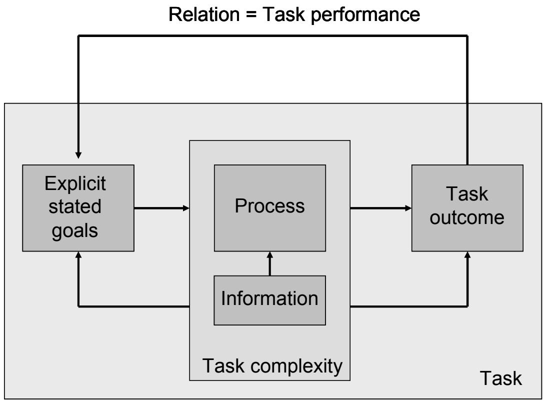
In real-life settings, tasks are carried out under circumstances of incomplete information and knowledge, resources, time pressure and conflicting goals. This makes real-life tasks complex. Wood (1986) defined task complexity as 'the relationships between task inputs' (i.e. human capabilities and limitations) and the provided information that is necessary for task performance. This is in accordance with Campbell, who states that task complexity can be defined objectively as being related to the cognitive demands put on the task performer (Campbell, 1988). Wood (1986) distinguishes three types of task complexity: (1) component complexity (the number of distinct acts and distinct information cues involved in the task), (2) coordinative complexity (the relations between the task inputs – information cues and acts – and the task output), and (3) dynamic complexity (the stability of the relationships between task inputs and output). The definition of task complexity used in the present research is:

Definition 5 Task complexity

Task complexity is the connection between task input and output in relation to human capabilities and limitation.

A further distinction can be made between task and task performance. Tasks relate to the input components and measurable output of acts of people achieving goals. Goals are the external reference criteria of tasks. The relation between the measurable task output and its reference criterion is called task performance. Often, both task and task output are called task performance. *Figure 3-1* depicts the model on the relation between the various task elements.

Figure 3-1 Model of tasks, task complexity and task performance



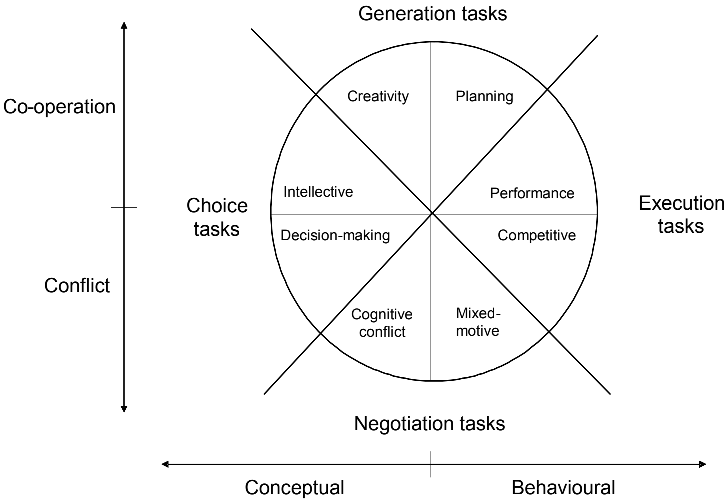
In professional working contexts with interchanging single- and multi-user tasks, a task starts by stating explicit goals (or even competing goals). Accordingly, users perform tasks through a certain process, supported by information. The interaction between the behavioural aspects of the task process, the information, goals and outputs creates complexity. Information is provided for that process, and for example the quality of the provided information is related to the complexity of the task: insufficient, incomplete, too much or complex information increases task complexity. At the end of that process the task is accomplished, which creates a measurable output. The relation between the accomplished task and the stated goals is called task performance.

Group tasks

Group tasks are also the behaviour requirements for accomplishing stated goals, via some process, using given information (see *Definition 4 Task*). The group task circumplex model (McGrath, 1984) is a conceptual framework for task classification. It proposes that all group tasks can be categorised into four major types, distinguished by two components: conceptual vs. behavioural oriented, and cooperation vs. conflict emphasis (see *Figure 3-2*). The four task types are:

- Generation tasks: primarily cooperative tasks, including creativity (e.g. idea generation and brainstorming), planning and scheduling.
- Choice tasks: primarily conceptual tasks, including the solving of structured problems as well as problems that require consensus.
- Negotiation tasks: primarily conflict resolution tasks, including conflicts of viewpoint and motivational conflicts.
- Execution tasks: primarily behavioural tasks, including excelling in performance, where there is some objective standard and competing for victory.

Figure 3-2 Group task circumplex model (McGrath, 1984)



In the present research, the circumplex model was applied in order to classify the group tasks of home care and railway. Wound care can generally be classified as a conceptual task, and train driving as a behavioural task. Wound care is a typical choice task in which a structured problem has to be solved: treatment according to plan or advice from other people is necessary. Cooperation and decision-making are important characteristics of this task. Train driving is a typical execution task, balancing between cooperation and conflict. The amount of infrastructure is limited, so conflicts might appear in the case of who goes first (freight or passenger train); on the other hand, cooperation seems a fruitful strategy to improve the general performance of the system (if a train is behind schedule, the train in front of it can ride at top speed, thus helping the delayed train).

Home care and railway are two different tasks in relation to group task performance. Home care is a conceptual task, related to treatment decision-making and cooperation, while railway is a behavioural task, related to execution and performance (driving decision-making). Decision-making is the main construct for both single-user and multiple-user task performance.

The cognitive constructs for a preliminary task performance model were identified in Chapter 2. This section on task performance shows that decision-making is the most important cognitive construct for information and communication support for the task performance of distributed professionals. When information and communication support are provided, the decision latitude of the distributed professionals changes (Karasek, 1979). The concept of decision latitude in the context of task performance refers to empowerment, namely a process of increasing personal, interpersonal or political power so that individuals can take action to improve their life or work situation (Speer & Peterson, 2006). Decision-making and empowerment are part of the larger conceptual model on cognitive aspects for task performance. Decision-making is the heart of the model. Therefore, decision-making is described first, followed by empowerment.

3.2 Decision-making

Decision-making is a broad construct with a long research history that relates to many aspects, for example problem-solving, judgement and choice. The present research concentrated on proficient decision-making in real-life settings. Decision-making is the cognitive process of deciding on a particular course of action, based on the available information. Applying this to home care and railway provided the following definition for the present research:

Definition 6 Decision-making

The cognitive processes that use context information to make decisions, which relate to measurable outputs (cf. task performance).

A distinction is made between decision-making structure, strategies and processes. The decision-making structure refers to what needs to be done, independent of how or by whom. Strategies look at the way decisions are made, and processes relate to the cognitive processes underlying decision-making. In the following sections, decision-making is studied in three ways: structure, strategy and processes, with a focus on home care and railway.

3.2.1 Decision-making structure

The decision-making structure identifies the constraints of what needs to be done, regardless of who is going to do it (Vicente, 1999).

Definition 7 Decision-making structure

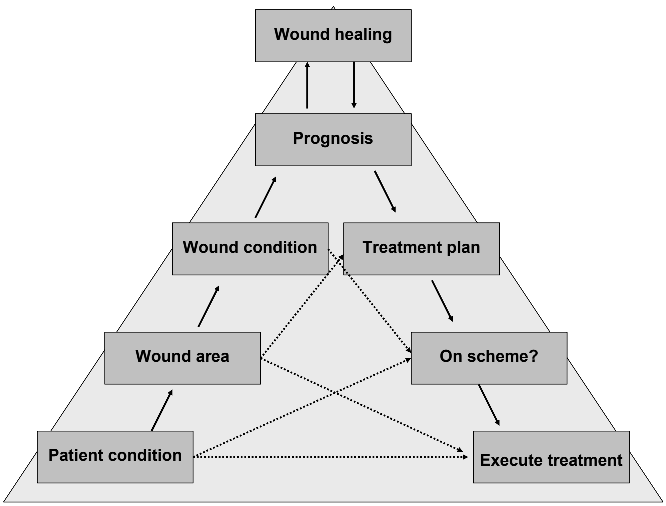
The decomposition of decision tasks, which can be processed in various orders, regardless of who is going to perform them.

The decision-making structure can be identified by applying a control task analysis that describes the decision tasks in information processing terms. This method is called decision ladder (Rasmussen, Pejtersen, & Schmidt, 1990). In principle, decision ladders make a distinction in data-processing activities and their results, which are states of knowledge. A decision ladder does not prescribe the different steps to take in a decision-making process, but focuses on the structure of the decision-making process. In this way, it is easier to represent the heuristic short cuts that experts take in the decision-making process. The decision ladder method was applied to home care and railway to illustrate the kinds of heuristics possible. The aim was not to identify all possible heuristics, but to establish the effect of context information on task performance. The success or failure of providing context information to professionals can be explained in relation to decision-making.

Home care

An example of a control task analysis for wound care in home care is provided below. The three main information elements are the patient condition, wound area and wound condition. *Figure 3-3* depicts the results of the control task analysis for home care, that is, the various states of knowledge. Starting from the bottom left ('patient condition') and going up the triangle (to 'wound healing'), and then down the right-hand side of the triangle (to 'execute treatment') the decision-making structure of wound care is shown. The dotted arrows represent possible short cuts that can be taken by experts. Note: in control task analysis, it is not decided who is going to perform which action.

Figure 3-3 Control task analysis of wound care in home care



The left-hand side of the triangle relates to assessing various information sources, the top relates to comparing the information with the explicit goal (wound healing), and the right-hand side relates to comparing the goals with the expectations in such situations. The dotted lines indicate that limited information sources already provide enough information for experts to make decisions.

Railway

An example of a control task analysis for train driving is provided below. The three main information elements are railway infrastructure and line-side information, route situation and the planned route. *Figure 3-4* depicts the results of the control tasks analysis for railway. Starting from the bottom left ('rail infrastructure and line side') and going up the triangle (to 'safe and reliable transport'), and then down the right-hand side of the triangle (to 'train handling') the decision-making structure of train driving is shown. The dotted arrows represent possible short cuts that can be taken by experts. Again note that in control task analysis, it is not decided who is going to perform which action.

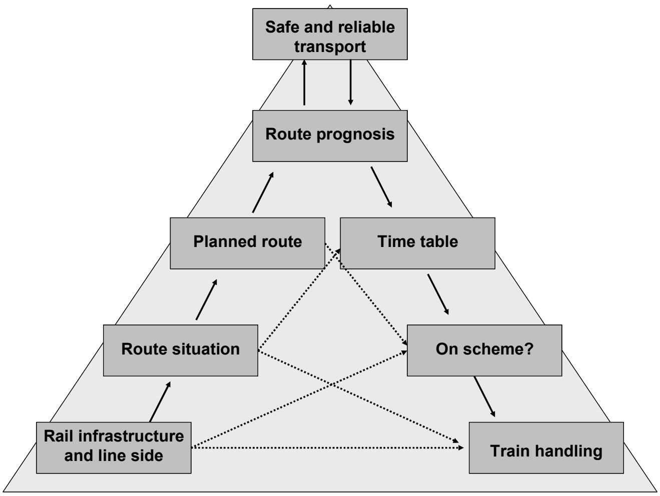


Figure 3-4 Control task analysis of train driving in railway

Again, the left-hand side of the triangle relates to assessing various information sources, the top relates to comparing the information with the explicit goal (safe and reliable transport), and the right-hand side relates to comparing the goals with the expectations of such situations. The dotted lines indicate that limited information sources already provide enough information to experts to make decisions.

3.2.2 Decision-making strategy

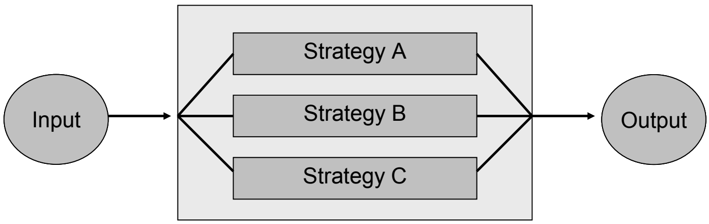
Decision makers use various information processing strategies to make decisions. Strategies focus on how work is done. They provide insights into the different ways in which the activities can be accomplished. The definition of decision-making strategy used in the present research is:

Definition 8 Decision-making strategy

Long-term plan of action designed to achieve a particular goal. It defines the flow of information.

Different decision strategies are typically used by the same decision maker depending on the properties of the task. Moreover, the strategy used is contingent upon task complexity; with low task complexity, decision makers often use full processing strategies, while an increase in task complexity typically results in the use of reduced processing strategies to reduce cognitive effort (Paquette & Kida, 1988). The most familiar decision-making strategy is defined by Simon (1957), and is called satisficing, that is, selecting the first option that works. Dedicated strategies are found in the literature. Some of these are cognitively complex, requiring the decision-maker to consider large amounts of data combined in a complex, typically compensatory fashion. Others are reduced processing strategies, which require a limited information search and simpler evaluation processes (Paquette et al., 1988), for example additive compensatory. Strategies are essential in helping the human to maintain performance in situations of stress, very high or very low workload, or when subject to other influences (McLeod, Walker, & Moray, 2005). The way a strategy determines the information flows is depicted in *Figure 3-5*.

Figure 3-5 Decision-making strategy



The choice of a decision strategy depends on the task complexity (Paquette et al., 1988), the decision-maker's perception of accuracy (benefit) and effort (cost) of the strategy (Beach & Mitchell, 1978) (Chu & Spires, 2003), and his or her own subjective task formulation (Vicente, 1999). The use of strategies in understanding and explaining human behaviour and performance in real-world situations is rare (McLeod et al., 2005). As strategies provide ways to reduce the information for real-life decision-making, it is a relevant construct for the present research. Therefore, typical treatment decision-making and train driving strategies found in the literature are described in the following sections.

Home care

It is known from the literature that various strategies are applied to treatment decision-making. Differences in treatment strategies originate from various sources, namely (1) hospital or GP, (2) home care organisation, or (3) individual home care provider.

In relation to (1), an observational evaluation study in three European foot-care centres revealed significant differences between these centres with regard to patient and wound characteristics, as well as with regard to management and outcomes (Eneroth et al., 2004). Besides definition differences, the variation in treatment strategies is probably a consequence of the shortage of scientific evidence on diabetic foot ulcerations treatment (De & Scarpello, 1999) as well as differences in health care systems, with regard to availability of resources and the system for remuneration (financial compensation) (Eneroth et al., 2004).

To illustrate the second variation in treatment strategies: during the interviews, several district nurses wanted to show their home care organisation procedure for wound care, but no one could find it either in their Intranet system or in their paper file.

Lauri and Salaneträ (2002) studied nursing decision-making in various countries (including North America, Scandinavia and Switzerland) and within various nursing disciplines (e.g. intensive care and psychiatric care). Although these countries and disciplines are not directly comparable to home care in the Netherlands, the study shows that, in general, four kinds of decision-making strategies are applied, varying between disciplines and countries. These four strategies are: analytical, analytical-intuitive, intuitive-analytical and intuitive (Lauri & Salanterä, 2002).

Analytical decision-making proceeds according to a certain systematic process, and the decision can be reached by analysis of the situation (Lauri et al., 2002). The typical characteristics of the intuitive process include rapid information processing, simultaneous cue use, pattern recognition, the evaluation of the cues at a perceptual level, and the principle of weighted-average organisation (Lauri et al., 2002). As the results show, analytical and intuitive processes are not contradictory, but they are the poles of a continuum (cf. (Hammond, 1996).

It is likely that providing history information to district nurses will mainly contribute to analytical decision-making. In the literature, there is no evidence regarding which decision-making strategy is more effective in the case of treatment decision-making.

Railway

A Swedish railway project identified two main driving styles for train drivers: a more feedback-related driving style, acting upon automatic train protection (ATP) system indications, and a more feed-forward driving style, acting ahead of indications (Kecklund et al., 2001). The same two driving styles came up during interviews with train drivers. There it was found that freight train drivers typically drove with a feed-forward driving style, and some passenger train drivers drove from signal to signal (Van der Velde et al., 2004). The current user interface of the on-board systems supports a feedback-related driving style. There is no conclusive evidence which driving style is most effective from a safety point of view (Kecklund et al., 2001). The nature of the driving strategy and especially how they vary, depends to a large extend on the driver's training, experience and confidence (McLeod et al., 2005), p 676. Confidence relates to the trust of a train driver in the system (is a red light *always* preceded by a yellow light?). At a more strategic level, drivers can also apply various driving styles. During the interviews, railway signallers indicated that not every train driver always aims at making up his delay: 70% of the train drivers drive at the maximum speed to reduce the delay, while 30% stick to the timetable speed.

3.2.3 Decision-making process

The cognitive processes that take place in decision-making can typically be described in information processing terms. Information processing theory relates to studies of human problem-solving that rely on the prior knowledge an individual has gained about the issues and areas concerned. This refers to intuitive processes like rapid information processing, simultaneous cue use, pattern recognition and the evaluation of cues at a perceptual level (Lauri et al., 2002). Information processing theory is a useful construct in understanding how distributed professionals make decisions in real-life contexts that are meaningful and familiar to them. Examples of characteristics of real-life settings are dynamic information, time constraints and competing goals. The typical decision-making theory that focuses on proficient decision-making in real-life situations is naturalistic decision-making.

Naturalistic decision-making

Naturalistic decision-making (NDM) theory is about the way people use their experience to make decisions in field settings. It is a descriptive approach that addresses proficient decision-making with constraints of limited time, high stress and incomplete knowledge, in other words, complex real-world settings. It is defined as (Zsambok et al., 1997):

Definition 9 Naturalistic decision-making

The study of NDM asks how experienced people, working as individuals or groups in dynamic, uncertain and often fast-paced environments, identify and assess their situation, make decisions and take actions whose consequences are meaningful to them and to the larger organisation in which they operate.

Four markers can be distinguished in this long and descriptive definition of NDM (Zsambok et al., 1997);

1. Task and setting factors
2. Research participants
3. Purpose of the research
4. Locus of interest in the decision episode.

The task and setting factors are the key contextual factors that affect real-life decision-making, in contrast to their counterparts in the traditional decision research paradigm. **Error! Reference source not found.** lists the factors in both research paradigms.

Table 3-1
Characteristics of naturalistic and traditional decision-making research (Randel & Pugh, 1996)

Naturalistic decision-making	Traditional decision-making
Ill structured problems	Well defined problems
Uncertain, dynamic environments	Relatively stable environments
Shifting ill-defined, or competing goals	Single, well understood goals
Action;/feedback loops	Single decision event
Time constraints	Decision deliberated
Multiple players	Single decision maker
Organisational goals and norms	Individual preference

The research participants were experienced decision makers, not naive subjects. The purpose of the research was to discover how experienced people actually make decisions in context-rich environments, not how they ought to make decisions in approximation to a rational standard. The locus of interest within the decision episode refers to interest not just in the option selection process, but also in situation awareness. To identify the applicability of NDM to the present research, the four markers were compared with the two cases.

In principle, the task and settings factors were applicable to both cases: the information was ill structured (and often competing), the information was rather dynamic (patient and wound conditions can change rapidly, with amputations as a consequence), goals were competing (punctuality versus energy costs), multiple players were involved, and organisational goals and norms applied. The more dynamic aspects (dynamic environments and time constraints) seem to be more applicable to railway than to home care. However, Bryans and McIntosh characterised district nurses' decision-making as (Bryans & McIntosh, 1996):

- Problems are uncertain and ill structured
- Dynamic and interactive nature of practice
- Temporal unfolding of information
- Contextual influence (home situation)
- Time constraints.

Qualified district nurses and train drivers took part in the experiments, thus the research participants were experienced and qualified decision makers.

The objective of the present research was to study the effect of extra information on task performances. This means that the purpose was not to discover how experienced people make decisions in context-rich environments, but to use the fact that experienced people make decisions in a certain way to test whether extra information can improve this decision-making.

In both cases, information processing models were applied to identify task performance. The heart of both performance models consists of the three elements of situation awareness, namely perception, comprehension and projection. Thus, the locus of interest is not on the option selection process.

The three basic principles on how decisions are made according to NDM are:

1. Decisions are made by holistic evaluations of potential courses of action (COA), rather than by feature-by-feature comparison of alternatives (Lipshitz, Klein, Orasuma, & Salas, 2001).
2. Decisions are recognition based; decisions makers rely on recognition of the situation and pattern matching of COAs rather than an exhaustive generation and comparison of alternatives (Klein & Calderwood, 1991).
3. Decision makers adopt a satisficing criterion rather than search for an optimal solution (Klein et al., 1991).

These three principles converge in the recognition primed decision (RPD) model, which is often applied in NDM (Klein, 1989). The RPD model was developed for firefighter research and claims that, in general, experienced decision makers assess a situation and judge its familiarity (principle 2), and imagination takes place to identify a course of action (principle 1) based on satisficing strategy (Klein, 1999) (principle 3). It was designed to better understand how experienced firefighters could handle time pressure and uncertainty. In these cases, people were not comparing decision options (principle 1), but relied on the first option they considered, because they mentally simulated this course of action to see if it would work, and to look for unintended consequences that might be unacceptable (Lipshitz et al., 2001). Four features of the RPD model are relevant (Klein, 1989) that match the basic principles of NDM:

1. Recognising cases as typical
2. Situational understanding
3. Serial evaluation
4. Mental simulation.

To further identify the applicability of NDM to the application domains, the factors that influence decision-making according to NDM (Currey & Botti, 20063) were applied to home care and railway. NDM emphasises factors associated with:

- The task; level of complexity
- The decision maker; primarily knowledge and experiences
- The environment; external information.

The following sections focus on decision-making in home care and railway, and emphasise the literature on the three NDM factors (task, decision maker and environment).

3.2.4 Decision-making in home care

Decision-making is a key element of district nurses' assessment practice (Bryans et al., 1996). Nurses working in the community often face decision-making challenges that are different from those encountered by their hospital colleagues. Nursing care provided in the patients' homes is even more idiosyncratic than hospital care because of the sheer diversity to be found in people's home environments, lifestyles, family and neighbourhood support systems (Bryans et al., 1996). Nursing studies indicated that the decision-making process is task dependent; different tasks serve as stimuli to use different cognitive modes, ranging from intuition to analysis (Lauri et al., 2002).

In the present research, the wound care task of district nurses was reduced to treatment decision-making in order to identify the effect of context information on task performance. When a nurse comes to the home of a patient, she first has to decide whether she can treat the patient according to the treatment plan or whether the situation requires additional help from the physician. Treatment decision-making is defined as:

Definition 10 Treatment decision-making

The decision by a district nurse whether a patient can be treated according to the treatment plan or whether additional help from the physician is required.

For treatment decision-making, district nurses use different types of information about the patient to arrive at a decision concerning the patient's current health status (Dowding & Thompson, 2003). During their education, district nurses consult general nursing textbooks that contain important practical nursing knowledge on all kinds of subjects, for example infectious diseases, chronic illness and wound care. However, the transition from a controlled hospital environment to the home requires the appreciation of a totally different context in which the nursing care is carried out (McIntosh, 1996). In relation to this transition from theory to practice, various types of knowledge for treatment decision-making can be distinguished: the technical rational type of knowledge (Schön, 1991), evidence-based knowledge (Buss, Halfsen, Huyer Abu-Saad, & Kok, 2004) and intuitive knowledge (Gatley, 1992).

Technical knowledge has been described as the application of scientific theory and technique to the instrumental problem of practice. Evidence-based knowledge relates to experimental results from science. Intuitive knowledge relates to patterns of personal knowledge that were found credible through reflections and actualisation by individual nurses (experiences).

Buss et al. (2004) found that nurses in nursing homes act and think based on traditional knowledge, that is, their basic education and the information they receive from their colleagues. Protocols used in nursing homes on pressure ulcer prevention were based on traditional knowledge instead of recent, evidence-based knowledge. This matches the nurses' work practices habits (protocols were developed by nurses), but do not match national guidelines.

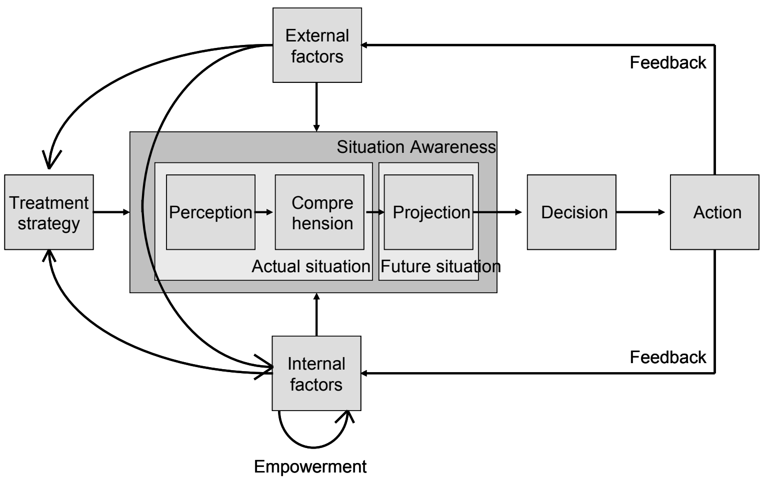
Structured interviews with 102 Norwegian district nurses about their own knowledge of wound treatment revealed that the main sources of nurses' knowledge are their own nursing practice and their colleagues (Haram et al., 2003). Working experience and the number of years of employment are important factors in building up knowledge (Haram et al., 2003). Combining these factors may undermine best practice in wound treatment (Haram et al., 2003).

A nurse's skill acquisition and decision-making path towards becoming an expert goes through five stages: novice, advanced beginner, competent, proficient and expert (Lauri et al., 2002). During these stages, the various types of knowledge have to be acquired, integrated and made ready for application. Roe and colleagues commented that nurses who have been qualified for the longest are the least likely to have updated their knowledge or received continuing education (Roe et al., 1993; Ribu et al., 2003). This matches with results of the user interviews for the present research (Hettinga et al., 2003). District nurses may lack the necessary education, training and knowledge of wound care programmes and guidelines, which may result in suboptimal care (Ribu et al., 2003).

On the basis of clinical experience, the treatment of patients with leg and/or foot ulcers has mainly been left to district nurses, with insufficient support from GPs (Ribu et al., 2003). An interview-based study in Norway revealed that district nurses possess a high degree of uncertainty about to wound treatment, wound assessment, wound healing products and wound causes (Haram et al., 2003). Moreover, 60% of the district nurses in this study believed that their knowledge of wound treatment was insufficient, 27% believed that they had enough knowledge, and 13% said that they were uncertain whether their knowledge was sufficient (Haram et al., 2003). Most of the nurses (47%) sometimes felt uncertain in wound treatment situations, 18% seldom felt uncertain and 7% often felt uncertain. The uncertainty was mainly related to the treatment, the assessment of problems arising in connection with the ulcers, wound-healing products and the causes of wounds (Haram et al., 2003).

For the present research, a wound care performance model was developed, based on the various cognitive constructs described earlier. The starting point was the task performance model described in Chapter 2. The general information processing model perception-decision-action is at the heart of the model. As history information should improve projection, the entire construct of SA was selected to describe the perception part of the model. The treatment strategy applied was also relevant as it determines the information flow and the way goals and performances are achieved in real-life settings. The various kinds of knowledge of a district nurse are part of her internal factors. Patient-related information is part of the external factors. The communication with physicians or colleagues is about patients. Therefore, in this model, communication is part of the external factors. The combination of internal and external factors determines the treatment strategy. The results of the actions are feedback to the internal and external factors. *Figure 3-6* depicts the wound care performance model

Figure 3-6 Wound care performance model



As the wound care performance model was the starting point for the user experiments described in *Chapter 5*, relevant model components are given in *Table 3-2*.

Table 3-2 Components of wound care performance model

Component	Elements
Treatment decision-making strategy	• Analytical decision-making • Analytical-intuitive decision-making • Intuitive-analytical decision-making • Intuitive decision-making
External factors	• Patient condition • Wound area • Wound information • Communication with experts
Internal factors	• Intuitive knowledge • Technical knowledge • Evidence-based knowledge • Mental workload
Decisions	• Recognition of the situation • Deviation from treatment plan • Communication
Actions	• Treatment according to plan • Call physician or experienced colleague later on • Call physician immediately

3.2.5 Decision-making in railway

In contrast to the large amount of literature on wound care and treatment decision-making, very little literature exists on railway driving and decision-making. According to Wilson and Norris, one of the reasons might be that 'For many years, railway around the world had a low profile regarding research and innovation' (Wilson & Norris, 2005b). Available rail human factors research concentrates on train driver vigilance, and perception and their recognition of and acting on signs and signals, including SPADs (signal passed at danger) (Wilson & Norris, 2005a).

Train driving is a dynamic control and decision-making task (Kecklund et al., 2001). It is a highly complex skill that requires the continual integration of interdependent perceptual, motor and cognitive processes (Salvucci & Macuga, 2002). A train driver has to use and integrate information from various sources, for example line-side signals and signs and ATP information. As stated, in general there are two driving styles: the feedback and the feed-forward driving style. The current system presents very little information about the task of planning ahead; the feed-forward driving style is not supported. Drivers use experience and expectations to compensate for the lack of information for planning the driving task ahead (Kecklund et al., 2001).

As train driving is a complex task, there is no simple dedicated train driving decision-making description. However, three decision-making models have recently been developed. One is intended to model driver performance (Hamilton et al., 2005), while the other two focus on train driver situation awareness (McLeod et al., 2005) (Livingstone et al., 2005). These three models are compared in this section, and specific attention is paid to context information to support train driving.

Hamilton and Clarke provide a general model of skilled train driver performance by focusing on the interaction of the usability of the line-side and cab interfaces and the driver's fundamental information processing capabilities (Hamilton et al., 2005). Their goal is to derive a general model of drivability assessment, to be used for safety analysis in design development. The inputs of this model are the time of key events when driving a proposed line layout, and the outputs are data on driver performance times, workload and error proneness. The model is an interaction between human capabilities and the recognise-act cycle. The human capabilities are based on memory, knowledge and remembering. The recognise-act cycle is about goal-driven behaviour, and the general information processing chain (i.e. perception-decision-action) is extended with recognition. Information flows between almost all elements, and outside information is a result of action, and is input for goal-driven perception.

McLeod et al. (2005) provide a model of train driver performance to understand its interaction with the automatic warning system (AWS) (i.e. automatic train protection system; ATP). They want to understand train driver performance in terms of the context and the situation at the time a signal is intended to influence driver behaviour. The focus of their situational model is on the time domain, because situational factors are inherently time dependent or time limited. Their starting point is the simple information processing model of perception-decision-action. The situational aspects of the model focus on establishing key aspects of the state of cognition at a specific time, in a specific situation and a specific context ('now state'). This cognitive now state consists of the immediate priority (reduce speed) and expectations concerning the next few moments (speed limit sign within a kilometre). Aspects of the cognitive states are continually changing and updating as the train drives along the route. These immediate and expectation aspects also affect the allocation of resources of the driver, and his chosen driving strategy. This cognitive now state is built on the knowledge, skills and experiences of the driver and on his current world model. Examples of such knowledge, skills and experiences are knowledge of procedures, recent route history and experiences with false alarm rates and incidents in the past. His current world model is a mental representation as it affects his current driving activity, for example current speed limit, state of the track, and relevant route knowledge about elements of the next few kilometres.

Livingstone et al. (2005) focus on training and assessment requirements for train driver route knowledge and situation awareness. Endsley's model of situation awareness in dynamic decision-making (Endsley, 1995b) was chosen as the starting point. To make this model more comprehensible to an audience that is not composed of psychologists and more specific to their research purpose, a literature study and subject matter expert interviews were conducted. The result is a simplified version of the Endsley model. Again, perception-decision-action is the main cognitive process of the model. However, perception was extended to situation awareness, consisting of perception, comprehension and projection. All elements were influenced by external factors and individual factors. External factors are extrinsic to the individual; examples are the external environment and cues, on-board systems and workplace organisation. Internal factors are intrinsic to the individual; examples are age, experience, workload, attention and mental models. Also an information flow goes from external factors to internal factors.

Comparison of train driver performance models

All three models use the general information processing chain (perception-decision-action) as the starting point. The Hamilton and Clarke (2005) model is the only one that explicitly uses the concept of recognition. Recognition is positioned between perception and decision. From the NDM theory it is known that proficient decision-making is mainly based on recognition; recognition primed decision model. The position of recognition in a model is questionable: is it part of perception or decision, or is it between those states? The other elements of the Hamilton and Clarke model are comparable to the other two models.

Livingstone et al. (2005) use Endsley's perception-decision-action model in which perception is part of situation awareness. The definition of Endsley on situation awareness (SA) is the one most frequently used in the literature, and was applied in the present research (Endsley, 1995b):

Definition 11 Situation awareness

"The perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.

Perception and comprehension refer to the current status of the situation, and projection to the future situation. This seems to be an integrated view of the immediate priorities and expectations of the model of McLeod and colleagues. This extended version of perception in relation to situation awareness is applicable to the present research, mainly as future system states are an integrated part of task performance. This relates to elements of the designed context information display of train drivers; path predictions and order of trains.

Both McLeod et al (2005) and Livingstone et al. (2005) state that external and internal factors influence perception-decision-action. The latter provide the names 'external' and 'internal', whereas the former call external factors 'the current world model' and internal factors 'knowledge and experience'. However, the difference between both driver performance models is that in the model of McLeod et al. internal and external factors are integrated into the cognitive now state (immediate priorities and expectations) that influences the perception-action cycle, whereas that integration of Livingstone et al. takes place in the extended perception part of the model (situation awareness). Another difference is that McLeod et al. add driving strategies to their model, as part of the cognitive now state. Driving strategies relate to the allocation of attention and influence the way external and internal factors are used. Allocation of attention is relevant to the present research as it relates to the use of context information; immediate priorities and expectations. In Livingstone et al. (2005), attention is one of the fourteen elements of internal factors, although its specific impact is not explicated. Driving strategy is not part of the model of Livingstone et al.

For the present research, a train driver performance model was developed, based on the integration and selection of the three train driver performance models analysed before. The perception-decision-action chain is the heart of the model. As projection is relevant to path predictions and order of trains, perception is constructed as situation awareness. The construct of driving strategies (in relation to the allocation of attention) is also a relevant part of the model as it determines the flow of information and the way goals and performances are achieved in real-life settings. The choice of a driving strategy depends on the task complexity, the train driver's perception of accuracy (benefit) and effort (cost) of the strategy and his own subjective task formulation (see *Section 3.2.2*). Driving strategy is the result of the integration of external and internal factors, and it is input for perception. Recognition is part of the decision. The results of the actions are feedback to the internal and external factors. *Figure 3-7* depicts the train driver performance model.

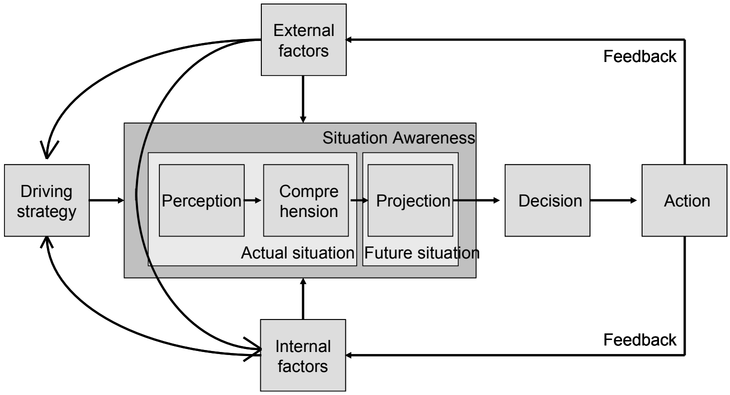


Figure 3-7 Train driver performance model

As this train driver performance model was the starting point for the user experiments described in *Chapter 6*, the relevant model components are given in *Table 3-3*.

Table 3-3 Components of train driver performance model

Component	Elements
Driving strategy	• Feedback driving strategy • Feed-forward driving strategy
External factors	• Cabin information • Line side and signals • Temporary procedures from railway signaller
Internal factors	• Knowledge and experiences, e.g. route knowledge, procedures • Mental workload
Decisions	• Recognition of the situation • Control of speed • Communication
Actions	• Apply/remove power • Apply/remove brake • Operate voice communication

3.3 Empowerment

Chapter 2 identified decision latitude as an important construct for the task performance of professionals. According to Karasek (1979), decision latitude and skill discretion both refer to the autonomy professionals have in performing their tasks. Decision latitude is also an important construct for the success of designing new information to support task performance. Its success depends on the user's perception of how he can exploit his knowledge, abilities and skills on the job. This relates to the construct of empowerment. Gutierrez (Speer et al., 2006) defined empowerment as 'a process of increasing personal, interpersonal, or political power so that individuals can take action to improve their life situations'. In relation to work situations, the construct of psychological empowerment at the workplace is defined by Thomas and Velthouse (1990) as 'the increased intrinsic task motivation, manifested in four cognitions; impact, competence, meaning and self-determination'. Impact relates to the degree to which an individual can influence strategic, administrative or operating outcomes at work. Competence or self-efficacy is an individual's belief in his or her capabilities to perform activities with skill. Meaning is the value of a work goal or purpose judged in relation to an individual's own ideals or standards. Self-determination is an individual's sense of being able to initiate and regulate actions (Spreitzer, 1995).

Empowerment is a constellation of experienced psychological states or cognitions (Spreitzer, 1995) (Thomas & Velthouse, 1990). This more complex perspective on empowerment focuses on the individual experience of empowerment (Thomas et al., 1990), that is, what individuals need to experience or feel in order for such interventions to be effective rather than the specific management practices intended to 'empower' employees (Spreitzer, 1995). The concept of empowerment in the present research concerns the increase in expert task performance by feelings of better using those experiences. Empowerment is defined as:

Definition 12
Empowerment

The increased intrinsic motivation for task performance by providing opportunities to better use knowledge and abilities at the workplace.

In relation to the task performance model, the hypothesis is that context information (the external factor) will empower the professionals. Empowerment will have an indirect effect on task performance as it is a way to increase the intrinsic motivation of professionals (the internal factor) by providing them with opportunities to better use their knowledge and abilities at the workplace.

3.4 Summary

A conceptual decision-making framework for context information to support task performance of distributed professionals was developed. Decision-making theory was evaluated on strategies, structures and theories of real-life task performance. As a result, a task performance model was developed and related to empowerment.

The home care performance model was applied to the experiments described in Chapter 5. The goals of the experiments were to identify the effect of context information on home care performances and the acceptability and applicability of WoundLog in practice. As district nurses are not used to working with technology, the experiment was divided into two parts. First, the concept of information support was applied without the use of technology; district nurses were asked to make treatment decisions in a paper and pencil task, with history-based wound information. Second, the concept of information support was tested using the mobile application WoundLog. WoundLog provides history-based information and teleconferencing facilities. A comparison of the two wound care experiments identified the effect of technology use.

In Chapter 6, the train driver performance model is applied to the experiments. The aim of the experiments was to identify the effect of context information on railway performances and its acceptability and applicability in practice. Train driving is a complex task, mainly due to characteristics of the task and the environment, for example dynamic information, competing goals, and time pressure. Therefore, the experiment was divided into two parts. First a static simulation was performed in which train drivers received authentic driving situations, enriched with context information. In such a static situation, without dynamic information and time pressure, the train drivers were asked to make driving decisions. The goal of this static experiment was to identify the effect of context information on decision-making, regardless of the task environment. In the second experiment, a dynamic multi-user simulator was used in which train drivers performed authentic tasks, in a more real-life situation enriched with context information.

In the following chapter measurement methods are identified to assess the theoretical constructs of the task performance model in the experiments.

Chapter 4

Assessment of task performance

This chapter focuses on the constructs that were measured in the user experiments. The research questions, the performance indicators and the task performance models were taken as starting points to identify the relevant constructs to measure. The selected constructs were task performance, task support, task processes, user acceptance and applicability. The selection of measurement instruments was guided by principles of unobtrusiveness with main task performance and validity.

4.1 Identification of constructs to be measured

As stated in section 3.1, the present research focused on two general research questions, namely:

1. What is the added value of information and communication support for distributed professionals on the main task performance indicators?
2. What is the user acceptance and perceived applicability of information and communication support for distributed professionals?

In practice, added value aims at the improvement of task performance. Task performance is operationalised in performance indicators. The starting point for the identification of the domain specific performance indicators was the generic performance indicators of social technical systems (i.e. safety, productivity and health) (Vicente, 1999). For the present research, service orientation was added as a general performance indicator. These general performance indicators were applied to both cases, and the case-specific performance indicators were given in *Table 2-4*. *Table 4-1* provides this table again. Thus, added value is operationalised in several case-specific performance indicators.

Table 4-1 Performance indicators (copy of *Table 2-4*).

Performance indicators.		
	Home care	Railway
Safety (Vicente ,1999)	Quality of care	Safety
Productivity (Vicente ,1999)	Efficiency of care	Punctuality
	---	Energy costs
Health (Vicente ,1999)	Job satisfaction	Job satisfaction
Service orientation (the present research)	Patient satisfaction	Passenger satisfaction

The identification of the case-based performance indicators provides opportunities for case-specific conclusions. To provide generic conclusions on the added value of information and communication support for distributed professionals, the kinds of tasks supported by the case-specific information need to be identified.

To gain insights into how the task performance of distributed professionals is supported by the extra information, the process of task performance needs to be established. Although insights into these processes were not the main focus of the research, they are useful in generalising the results of the two cases. It was already identified that decision-making is the most relevant cognitive construct for task performance. Naturalistic decision-making (NDM) is the decision-making theory that bests describes how experts make decisions in real-life situations. Zsambok and Klein state that in NDM, situation awareness is relevant for identifying how decisions are made, that is, decision episode (Zsambok et al., 1997). As stated, situation awareness refers to the perception of elements in the environment, the comprehension of their meaning and the projection of their status in the near future (Endsley, 1995b).

In human factors research, it is relevant to know not only how decisions are made, but also how much effort is needed to perform those tasks. This relates to the construct of mental workload: what is the effort needed to use the information and communication support in task processes?

Thus, in relation to the first research question, task performance, task support and task processes were relevant constructs to measure.

The second research question is about the user acceptance and perceived usability of the information and communication support provided by the mobile applications. In general, the aim was to identify what the professionals perceive of such support: are they willing to work with it, and what are their expectations of using it in practice?

Thus, in relation to the second research question, it was relevant to measure user acceptance and applicability of the concept in practice.

The general principle of the user experiments was that qualified district nurses and train drivers would perform authentic tasks facilitated by mobile applications that provide them with information and communication support. In order to keep these professionals focused on performing their tasks in their usual way, the measurement instruments had to be such that they would not obstruct their main task performance and preferably could be validated. Especially the construct validity is of relevance. Construct validity refers to establishing correct operational measures for the concepts being studied (Yin, 1994).

In the following sections, the measurement instruments for the constructs task performance, task support, task processes, and user acceptance and applicability are identified.

4.2 Task performance

Task performance is operationalised in domain-specific performance indicators. Proficient decision-making is often described in terms of satisficing (Simon, 1957): experts take decisions not by weighing each cue against one other, but by applying the first course of action that is satisficing. This means that proficient decision-making aims not at optimising task performance, but at satisficing task performance. In the following sections, the measurement of the specific performance indicators for each case is described.

4.2.1 Treatment decision-making performance

In this stage of the development of the WoundLog application, the focus was on the district nurses: what can they do with WoundLog and how do they perceive using it? The results of the user experiments were used as input for the further development of WoundLog. The task performance of district nurses in home care relates to the performance indicators quality and efficiency of care and patient and job satisfaction (see section 2.5.4). *Table 4-2* shows the home care performance indicators in relation to the wound care activities of district nurses. These descriptions formed the basis for the identification of measurements for task performance.

Table 4-2 Descriptions of home care performance indicators

Home care performance indicator	Description
Quality of care	The quality of treatment decision-making.
Efficiency of care	Improved efficiency and productivity of district nurses; reduced wound healing time and decreased number of physician visits, re-hospitalisation and emergency department visits.
Patient friendliness	Faster wound healing processes, involvement and motivation of patient in healing process, decreased number of physician visits, re-hospitalisation and emergency department visits.
Job satisfaction	Improved treatment decision-making, coaching and education; feedback on their wound activities.

Treatment decision-making affects all the performance indicators of district nurses (reduced wound healing times are a result of improved treatment decision-making). It is the most obvious construct for quality of care. To avoid a situation in which all performance indicators boil down to the same construct, treatment decision-making was dedicated to the quality of care.

The experiments took place in a semi-controlled environment, with paper case descriptions instead of real patients. In such circumstances, it is impossible to measure efficiency of care and patient friendliness. The experiment took three hours per district nurse, and therefore the effect of treatment decision-making on the efficiency of care could not be established. There were no real patients involved, and therefore the effect of treatment decision-making on patient friendliness could not be established. The quality of care is related to treatment decision-making: if the delivered care is not effective, or the patient also suffers from something else, the sooner the treatment plan is adjusted, the sooner the patient will be cured. Job satisfaction relates to educating and coaching district nurses to exploit their knowledge and abilities. In this way, it empowers district nurses in their work activities. Job satisfaction is covered by the construct of empowerment.

Thus, in general, quality of care can be identified by assessing the construct of treatment decision-making, and job satisfaction by measuring the construct of empowerment. Assessment of the construct empowerment is described in *Section 4.2.3*. Patient friendliness and efficiency of care cannot be measured directly. However, as district nurses worked with the application in the experiment, they could state their perceptions of possible effects on each performance indicator (including efficiency of care and patient friendliness).

Treatment decision-making is a relevant sub-task and of crucial importance for the well-being of the patient and the quality of the delivered care. The general hypothesis is that history information improves treatment decision-making. The communication facilities of WoundLog are used in cases of negative treatment decisions: help from a colleague or a physician is necessary. The quality of the treatment decisions is measured by establishing a standard for good and bad treatment decisions. This standard is called the frame of reference.

An expert district nurse (a member of the Diabetic and Wound Care section of the Livio home care organisation) composed eighteen sets of patient material for the experiments. The material originated from the Internet, pharmaceutical companies and patients in her district (informed written consents were acquired). The authentic patient material is composed of a description of the patient's current situation as a district nurse visits him at home; case description, a colour picture of the wound and a report on the patient's condition and wound information (*Figure 4-1*). The available wound history information consisted of history reports on the patient's condition and wound information.

Name of the patient P 6 Mr. Smith
Description: burn
Date: current situation

Mr. Smith is 49 years old and works as fire fighter.

On his day off, he and his wife were sitting in the garden, enjoying the beautiful weather. The neighbour was preparing the barbecue with charcoal. But, it didn't work, so he put some alcohol on it. The barbecue burst into flames and dropped down. Unfortunately, a piece of charcoal came down on Mr. Smith's leg.

Wound information		***General condition*:**	
Size:	5 cm	**Pain scale**:	6
Depth:	superficial	**Nutrition:**	good
Colour:	rood	**Fever:**	no
Scent:	no scent	**Diabetes:**	no
Pain:	sharp pain	**Mobile:**	yes
Wound area:	warm		

Figure 4-1 Example of patient material, case description, status report and wound picture

The expert district nurse qualified the patient material as six 'normal treatment', six 'semi-urgent treatment' and six 'urgent treatment' cases. To validate the ratings of the expert district nurse, two experienced district nurses together also rated all patients' material on treatment decisions, that is, normal, semi-urgent and urgent. Interrater reliability (Cohen's kappa; K) is a measure for the amount of agreement between the ratings, corrected for the amount of agreement expected by chance. The measured Cohen's kappa was 0.25. In literature, no real benchmark is available on what poor and good scores are for interrater reliability. Often applied is the classification of Landis and Koch (1977) on agreement measures for categorical data. They state that $K>0.21$ and $K<0.40$ is a fair strength of agreement (Landis & Koch, 1977). *Table 4-3* shows the cross table of the ratings.

Table 4-3 Cross table of ratings of expert and experienced district nurses.

		Expert District Nurse I			
		Normal	Semi-urgent	Urgent	Total
Experienced District Nurse II	Normal	5	3	1	9
	Semi-urgent	1	0	1	2
	Urgent	0	3	4	7
	Total	6	6	6	18

The fair strength of agreement between the district nurses, as reflected in a relative low Cohen's kappa, is a source of concern. It might indicate that treatment decisions are difficult to make and that they vary a lot between experts. The main difference between the two ratings stems from the semi-urgent category. The two experienced district nurses rated only two patients as semi-urgent, and these had been rated by the first nurse as normal and urgent. Three of the patients who the expert district nurse rated as semi-urgent were rated normal by the two experienced district nurses, and the other three were rated urgent. The semi-urgent category refers to patient situations in which the district nurse makes a note and contacts an experienced colleague or physician later on. This is a situation in which assistance from another person is necessary (and WoundLog can be used).

During the experiment, it turned out that in one district, the expert district nurse worked only four hours a week. Contacting her for extra patient visits is not part of the normal working practice of the district nurses in that district. As 55% of the participants in the experiment originated from that district, the semi-urgent category was meaningless to them. In the urgent category also assistance or intervention from another person is necessary (and WoundLog can be used). Therefore, it was decided to merge the categories 'semi-urgent' and 'urgent'. This led to the category 'intervention necessary'. Thus, two frames of references were identified: the normal, semi-urgent and urgent categories (NSU), and the normal and intervention-necessary categories (NI). The two frames of references relate to the treatment decisions as identified by the expert district nurse.

Conclusions: frame of reference

In general, treatment decisions concerning patients are difficult to quantify: what is a good decision, and what is a bad decision? Second opinions are normal in health care, especially as there is no such thing as a 'gold standard' in medicine. For the identification of a frame of reference in the present research, the measured strength of the interrater's agreement is fair. Qualified district nurses provided varied treatment decisions on the same case material. Such personal aspects as trust in the situation and proactivity probably play a role, but also knowledge, education and experience might be relevant. In the interviews, one of the dermatologists already indicated that 'an urgent situation from the one district nurse differs from an urgent situation from another district nurse' (Hettinga et al., 2003). Therefore, he wanted to know the district nurse, so he could 'judge' her decision.

Two frames of references were developed to compare treatment decisions. The first is based on the ratings of an expert district nurse: normal, semi-urgent, and urgent treatment decisions (NSU). These three categories match the results from the user interviews and consequently the design of the WoundLog application. Two qualified district nurses also rated the patient material. Their results were hardly comparable with the ones from the expert district nurse. The main differences were in the semi-urgent category. Due to organisational circumstances, this category was applied less in practice. The semi-urgent and urgent categories were merged into the 'intervention necessary' category. This resulted in the second frame of reference: normal and intervention-necessary categorisation (NI).

Relevant in the identification of the frames of reference is that the measurement error is high. This influenced the statistical analyses of the results of the treatment decision-making experiment in a negative way.

4.2.2 Train driving performance

The identified railway performance indicators are safety, punctuality, energy costs, job satisfaction and passenger satisfaction (see *Section 2.7.2*). *Table 4-4* lists the descriptions of the railway performance indicators.

Table 4-4 Descriptions of railway performance indicators

Railway performance indicator	Description
Safety	Safety on track, as a result of train driving (external factors like car accidents are excluded).
Punctuality	The arrival and departure punctuality (in relation to timetable).
Energy costs	Accelerations of trains (power on).
Job satisfaction	Better utilising knowledge and capabilities of train drivers (cf. empowerment).
Passenger satisfaction	A combination of punctuality, informing passengers, and driving comfort.

The performance indicators are slightly interrelated (punctuality is also part of passenger satisfaction), but some are competing indicators. The relevance of each performance indicator is situation dependent: when a train is behind schedule, punctuality is more important than energy costs. Train driving decision-making is a continuous control task. Identifying each train driving decision as good or bad in relation to a frame of reference is not possible. Therefore, not each train driving decision was measured, although the results of the train driving decisions (i.e. the actions of trains) were measured. Log data consists of time, position and speed.

In general, two types of comparisons are possible in order to quantify the task performance of train drivers. First, two performance indicators are related to an external standard: safety and punctuality. Second, the performance indicator energy costs cannot be related to an external standard, but needs to be compared within the experiment. The frame of reference for each of these three performance indicators is described in detail. The fourth performance indicator – passenger satisfaction – cannot be measured directly, as passengers were not part of the experiment. The fifth performance indicator – job satisfaction – is operationalised as empowerment (cf. home care) and described in a separate section (*Section 4.2.3*).

Safety

Safety relates to the safety on track as a result of train driving; external factors like car accidents were excluded. The standards for safety are the speed limits as defined by line-side signals and signs. The first direct measure is signal passing at danger (SPAD); a danger signal (red signal) is passed. Most other safety violations are identified by the railway safety system (ATP interventions). These interventions are the second direct measure.

Moreover, two indirect measures of safety were identified. 'Indirect' means that they do not directly assess safety, but provide an indication of safety. The identified indirect measures are mental workload and how long it takes to react to the dead-man's handle. The dead-man's handle is an on-board safety system. A train driver has to respond to the dead-man's handle every 60 seconds, to indicate that he is still able to drive the train. If he does not respond within seven seconds, the ATP system stops the train. At first, the dead-man's handle signal is visual (a blue light); if the train driver does not respond within two seconds, a beep sounds. If the train driver does not respond to the beep within five seconds, the ATP system stops the train. Thus, in total, a train driver gets seven seconds to respond to the dead-man's handle. The driver's reaction time is an indication of mental effort: the more he is mentally loaded, the slower his response to the dead-man's handle. In relation to mental workload assessment instruments, this relates to secondary task performance (see section 4.4.2). Another indirect measure of safety is mental workload. It relates to the amount of resources demanded by a task. If a task (train driving supported by extra context information) requires too many resources, safety might become at risk. Assessing mental workload is described in more detail in the section about task processes (4.4.2).

In the present research, safety was measured in four ways: two direct measures – namely signal passed at danger (SPAD; passing a red signal) and automatic train protection system interventions (exceeding speed limits, except for SPADs and dead-man's handle) – and two indirect measures, namely mental workload and how long it takes to react to the dead-man's handle.

Punctuality

The standard for punctuality is the timetable. A train driver has to drive according to the timetable. The timetable establishes the arrival and departure times at stations. Departure punctuality depends on the interaction between the train driver and the train guard and their passengers. As they were not part of the simulation, the departure punctuality was excluded from the analysis. Thus, punctuality was computed as the difference in seconds between the measured arrival time and the arrival time stated on the timetable.

Energy costs

For energy costs, an objective frame of reference was hard to establish (what is a 'cost-effective' ride?). Therefore, a relative frame of reference was developed. Energy costs were analysed by comparing the mean positive accelerations per route and information condition. To make comparable the various trains (Intercity and local trains) and the various instantiations of the normal and abnormal routes, the mean acceleration per instantiated route per train was measured (reference route), regardless of the information condition. Each individual ride was then compared to this reference route. The deviation from this reference route is the computed energy costs per ride.

Passenger friendliness

Passenger friendliness consists of a combination of punctuality, information provision and driving (or braking) comfort. Punctuality was measured as part of the performance indicator punctuality. The other two aspects were not taken into account in this study, so passenger friendliness is not assessed directly.

Job satisfaction

Similar to the case with home care, job satisfaction in railway relates to train drivers who can better exploit their knowledge and abilities in their job. In this way, it empowers train drivers in their work activities. In this way, job satisfaction is covered by the construct of empowerment (*Section 4.2.3*).

4.2.3 Empowerment

Health is defined as one of the three general performance indicators of computer-based applications in work environments (see *Section 2.2*). The degree of freedom that professionals have to perform their job is a relevant part of the health construct. In relation to that, Karasek (1979) stated that decision latitude and skill discretion both refer to the autonomy professionals have in performing their tasks. Autonomy in professional task performance relates to the way professionals can apply their knowledge and abilities in performing their jobs. This autonomy or decision latitude forms the basis for professional job satisfaction. Of course, salary also adds to job satisfaction, but that was outside the scope of the present research. Moreover, the constructs of decision latitude, autonomy, and exploiting knowledge and abilities relate to the construct of psychological empowerment. Therefore, in the present research, job satisfaction was measured by empowerment. The construct of empowerment conducted in the present research relates to psychological empowerment at the workplace as used by Spreitzer (1995). Spreitzer adopted the definition of Thomas and Velthouse (1990): 'Thomas and Velthouse argued that empowerment is multifaceted and that its essence cannot be captured by a single concept. They defined empowerment more broadly as increased task motivation manifested in a set of four cognitions reflecting an individual's orientation to his or her work role; meaning, competence, self-determination and impact'. The four cognitions are described as (see also *3.3):*

- Impact, which is related to the degree an individual can influence strategic, administrative or operating outcomes at work.
- Competence or self-efficacy is an individual's belief in his or her capabilities to perform activities with skill.
- Meaning is the value of a work goal or purpose judged in relation to an individual's own ideals or standards.
- Self-determination is an individual's sense of having in initiating and regulating actions

The definition of empowerment is consistent with the assumption of the present research that information and communication support provides opportunities to professionals to better exploit their knowledge and capabilities.

Spreitzer (1995) developed an empowerment questionnaire and performed a first validation of a multidimensional measure of psychological empowerment in a work context. Her final questionnaire is composed of three statements relating to each of the four cognitions (i.e. meaning, competence, self-determination, impact). To reduce application time, one statement was selected for each cognition to be incorporated in the questionnaire. The questions are given in **Error! Reference source not found.**.

Table 4-5 Selected questions on empowerment

Cognition	Question
Meaning	The work I do is very important to me
Competence	I have mastered the skills necessary for my job
Self-determination	I can decide on my own how to go about doing my work activities
Impact	I have a great deal of control over what happens in my department[5]

The level of agreement with each question had to be indicated on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). The same questions (in Dutch) were applied pre-test and post-test to all experiments. Empowerment was measured as the difference between the pre-test answers (empowerment in the current situation) and the post-test answers (empowerment when information and communication support are provided

4.3 Task support

Home care and railway are two different cases. In the present research they were aligned by the concept of mobile information and communication support for distributed professionals. To make a comparison of the measured effects on task performance, the types of tasks supported by the information and communication were identified.

In *Section 3.1* on Task Performance, group tasks are classified according to the group task circumplex model of McGrath (1984). This model distinguishes four major types of tasks, namely generation tasks, choice tasks, negotiation tasks and execution tasks. Wilson and Morrison (2000) took the group task circumplex model as the starting point for the development of an instrument to measure perceived effectiveness for task type (PETT). The PETT measure was developed for use in research specific to computer-mediated communication (Wilson & Morrison, 2000).

[5] Instead of 'department', a case-specific word was chosen in the final questionnaire.

In the PETT questionnaire, each task type is operationalised through four activity items. Research participants are asked to rate the effectiveness of the information and communication support of the mobile application for each task activity, based on their experiences during the experiment (Wilson & Connolly, 2001). The questionnaire needs to be completed post-test. The 16 answers are provided in relation the perceived effectiveness of the application for the four types of tasks. The responses are indicated on a 5-point Likert scale.

As it is based on the circumplex model of McGrath (1984), the perceived effectiveness of task type seems a useful measurement instrument to identify the group tasks that are supported in home care and railway. However, it uses 16 questions to identify supported tasks, namely 4 questions per task type (Wilson et al., 2001). As task support is one of the constructs measured in the present research, the number of questions had to be reduced. Therefore, the four questions per task type were adjusted to one question, dedicated to the present research. The questions are given in *Table 4-6*.

Table 4-6 Selected questions on Perceived Effectiveness of Task Type

Task type	Question
Planning	The information and communication support provided help me to better plan the tasks that need to be performed.
Complex decision-making	The information and communication support provided supports me in complex decision-making
Negotiation	The information and communication support provided helps to negotiate who will be responsible for something
Group performance	The information and communication support provided help to improve group performances

The level of agreement with each question had to be indicated on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). Post-test, participants were asked to answer the questions. This provides an indication what kinds of tasks the provided information and communication support could be effective. The questions were translated into Dutch and applied after all experiments.

4.4 Task processes

As stated in the introduction to this chapter, the relevant constructs to be measured for task processes were situation awareness and mental workload.

4.4.1 Situation awareness

The most commonly used and widely cited definition of SA is Endsley's (1995a, p.88), that is: 'The perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future' (see also *Definition 11*). SA is most often used in typical human factors research areas, such as military research, air traffic control, driving and aviation (Stanton, Salmon, Walker, Baber, & Jenkins, 2005). However, it has also been applied to the health care domain, which is also characterised by complex dynamics and safety critical context (Blandford & Wong, 2004); (Gaba & Howard, 1995). 'Situation awareness (SA) is an understanding of the state of the environment (including relevant parameters of the system). It provides the primary basis for subsequent decision-making and performance in the operation of complex, dynamic systems.' (Endsley, 1995a)

Decision-making is the most relevant construct of the present research. The extra information and communication provided change the context and work environment in which the decisions have to be made. The assumption of the present research was that the extra information and communication provided would lead to a better projection of the status of the relevant subjects (i.e. wound care patients and trains). A better projection of their status in the near future should improve decision-making and thus improve task performance. SA was a relevant construct in the present research as it provides insights into what professionals do with the extra information and communication support, and what happens in the preceding processes of decision-making.

In general, SA can be measured in many different ways, for example by physiological techniques, performance or task measures, subjective techniques, or questionnaires (Endsley, 1995a), and can be applied during the whole design cycle (Stanton et al., 2005). However, the main issue in measuring SA is that no criteria exist that establish the level of SA required for successful performance. Does an operator need to have an SA that is 100% perfect (in both completeness and accuracy), or is some lesser amount sufficient for good performance? (Endsley, 1995a) Endsley answers this question herself by stating that '… likely, SA can be seen as a factor that increases the probability of good performance, but that it does not necessary guarantee it, as other factors also come into effect, for example decision-making, mental workload, performance execution, system capabilities, and SA of other people' (Endsley, 1995a).

The present research did not aim at establishing an objective level of SA, but at comparing SA between the various information and communication conditions of the applications. Moreover, its results were to be used to explain and understand the results on task performance; the effect of information and communication support on perception, comprehension and projection.

In the present research, SA is an additional factor that was measured. The selection criteria for the user experiments were that SA assessment should not obstruct the main task performance and that, preferably, validated instruments are used. Stanton et al. (2005) provide an overview of SA assessment methods, categorised by for example type of method, training time and application time. Two methods were found that do not obstruct main task performance: situation awareness rating scales (SARS) and situation awareness rating technique (SART) (Stanton et al., 2005). These two methods are self-rating techniques that require limited training and application time. However, both methods are domain specific (aviation), which means that all categories measured concern aviation-specific items.

Thus, as no SA measurement instrument is available to measure SA in home care and railway properly, it was decided to develop a tailor-made SA instrument. This instrument is closely related to the three levels of SA, namely perception, comprehension and projection. This questionnaire is a self-rating technique that should be applied post-test and uses pen and paper (as most district nurses are not used to working with computers). To avoid making it a rehearsal task, the answer possibilities are not categories or predefined descriptions; instead, open-ended questions are posed.

The advantage of developing a SA measurement instrument is that is dedicated to the two application domains. The disadvantage is the construct validity (does it really measure what it claims to measure?). However, construct validity is also not known for SARS or SART (Stanton et al., 2005).

Dedicated SA questions were developed for home care and railway. The goals were to gain an insight into the decision-making process of the participants in the experiments, and to identify whether information and communication support influences this process. Therefore, SA results were not related to an external standard, but were compared between the various information conditions.

For home care, SA was measured after each communication task; two district nurses talked together about a patient to make a proper treatment decision. The communication was facilitated by mobile phones, or by the WoundLog application. When WoundLog was used, also information about the patient was available for both district nurses. With mobile phones, only one district nurse had the visual information. The main task for them was to describe the case they had just discussed, in relation to the wound (perception), the case information (comprehension) and the prognosis (projection).

For railway, SA was measured after each journey. The assumption was that context information is most useful when 'something' happens, for example a traffic jam, stopping for a red signal, or a speed reduction because of a freight train in front. Therefore, post-experiment the train drivers where first asked to mark on a map where the most remarkable situation of the journey took place. They were then asked to describe that situation, also in relation to other trains. Next, they had to indicate how the situation was solved, and whether a better solution could have been found.

Table 4-7 Description of contents of measurement instrument of situation awareness

SA levels	Home care	Railway
Perception	Describe the most relevant characteristics of the wound	Describe or sign the most remarkable situation of the last ride, when necessary include other trains.
Comprehension	Which information from the case description was most relevant for you?	How did this situation solve or evolve?
Projection	What do you think the prognosis is for this patient?	Was a better solution possible?

As stated, no criteria exist to establish that establish the level of SA required for successful performance (Endsley, 1995a). Therefore, the results of the SA measurement were used in a relative way. The results of the various levels of information and communication support applied were compared with each other. In this way, the aspects of the information and communication support that influence perception, comprehension and projection could be identified. The questionnaires used in the experiment are given in *Appendix A*.

4.4.2 Mental workload

A strict definition of mental workload is scarce in the literature; the construct is hard to define. It is generally accepted that mental workload is about the relationship between task demands and human resources. Therefore, the description of Stanton et al. was applied (Stanton et al., 2005):

Definition 13 Mental workload

Represents the proportion of resources demanded by a task or a set of tasks.

Mental workload is a construct that is typically used in human factors research; it provides indications of whether the task environment is adapted to human resources. Comparable to SA, an optimal level of mental workload to perform tasks cannot be established. The general rule is that situations of very low mental workload and situations of very high mental workload should be avoided. Mental workload has a long research tradition;

various types of mental workload assessment techniques have been applied. On a high level, mental workload can be assessed in three ways:

1. Secondary task performances measures
2. Physiological measures
3. Subjective rating techniques.

Secondary task performance measures the ability to perform a secondary task in addition to the primary task. This is based on the assumption that as operator workload increases, the ability to perform the secondary task diminishes due to a reduction in spare capacity, and so secondary task performance will suffer (Stanton et al., 2005). Secondary task measures can be obtrusive in measurement.

Physiological measures involve the measurement of those psychological aspects that might be affected by increased or decreased levels of workload, for example heart rate and heart rate variability. Psychological measures can be obtrusive, and there are doubts about the construct validity and sensitivity of the techniques (Stanton et al., 2005).

In subjective rating assessment techniques, participants have to provide ratings regarding their perceived mental workload during task performance. These subjective ratings are administered either during or after the task performance. These techniques are unobtrusive to primary task-performance and are easy to apply and analyse (Stanton et al., 2005). However, the disadvantage of this technique is that participants are prone to forgetting certain parts of the task where variations in their mental workload might have occurred.

As subjective rating assessment techniques are less obtrusive, they were selected to measure mental workload. Home care or treatment decision-making, and railway or train driving are two different tasks. In the present research, the aim was not to identify mental workload, but to establish how mental workload affects task performance by providing information and communication support. Verwey and Veltman (1996) call this adding a task to a task environment. They state that 'However, little is known about the diagnosticity of the various workload assessment techniques, that is, to what extent the various techniques are sensitive to one or another type of workload' (Verwey & Veltman, 1996). Diagnosticity refers to whether the technique discriminates between types of workload, and sensitivity refers to whether the technique discriminates between the levels of workload (O'Donnel & Eggemeier, 1986).

What is obvious in the treatment decision-making and train driving tasks is that when extra information is added, the tasks relate to visual workload. Verwey and Veltman (1996) found that both the SWAT (Subjective

Workload Assessment Technique) and RSME (Rating Scale for Mental Effort) are sensitive for visual workload, but that they were not diagnostic with respect to mental and visual workload. As a validated Dutch version of the RSME is available, that subjective workload assessment technique was selected. The Dutch version is called BSMI (Bepalings Schaal Mentale Inspanning) (Zijlstra & van Doorn, 1995). The BSMI is a one dimensional (vertical) scale, labelled with statements at various points on the scale, for example no effort at all required, a little effort required, till very much effort required. The scale is divided into centimetres, each labelled as 10. The BSMI scale ranges from 0 (no effort at all) to, in principle, infinity, although the last figure is 150 (112 is the last statement labelled as 'very much effort'). Post-task, participants have to indicate the effort required to perform the task by putting a mark on the scale. The applied BSMI is depicted in *Appendix A*.

4.5 User acceptance and applicability

In general, professionals spend a significant part of their time performing their job. The way they can perform their everyday work is relevant. Introducing new technology at the workplace influences this strongly. Therefore, user acceptance of technology is a relevant construct. Research on technology acceptance has a long tradition, from several perspectives. In 2003, Venkatesh, Morris, Davis and Davis compared eight prominent models and their extensions and integrated elements across those models to develop a unified model; the Unified Theory of Acceptance and Use of Technology (UTAUT). They also empirically validated this model. UTAUT measures the success of a new technology by assessing the intention of users to use the new technology after its introduction. Knowing these factors helps to understand the drivers of acceptance, which can guide the design interventions for an introductory path. The advantage of this method is that the intention to use is measured, rather than real usage. This matches the phase of the experiments of the present research: applications are in an experimental test phase, and not ready for introduction in practice.

In the UTAUT model, four constructs play a significant role as direct determinants of user acceptance and usage behaviour. *Figure 4-2* depicts this model.

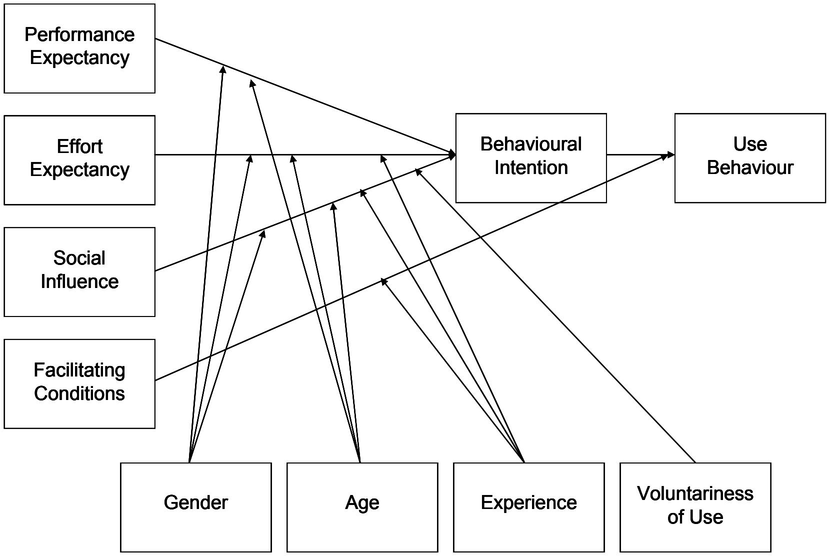


Figure 4-2 Unified Theory of Acceptance and Use of Technology model (Venkatesh, Morris, Davis, & Davis, 2003)

These four core determinant of intention and usage of new technology are:

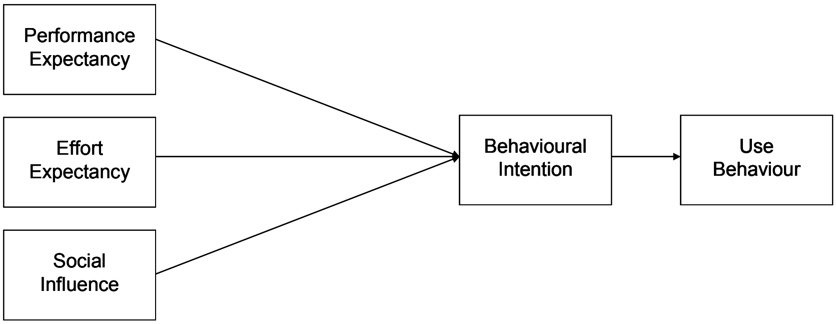
1. Performance expectancy: 'the degree to which an individual believes that using the system will help him or her attain gains in job performance' (Venkatesh et al., 2003).
2. Effort expectancy: 'the degree of ease associated with the use of the system' (Venkatesh et al., 2003).
3. Social influence: 'the degree to which an individual perceives that important others believe he or she should use the new system'. (Venkatesh et al., 2003).
4. Facilitating conditions: 'the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the system' (Venkatesh et al., 2003).

As Venkatesh et al., (2003, p. 454) empirically established that 'facilitating conditions will not have a significant influence on behavioural intention' but 'do have a direct influence on usage', facilitating conditions were not taken into account in the present research. The aim of the experiments was to identify the intention to use the information and communication support, not the actual use of it in everyday practice.

The relationship between the four core determinants and the intention to use is influenced by four key moderators: gender, age, voluntariness of use, and experience. In the experimental phase of the research, the participants were asked to perform authentic tasks with the information and communication support. The voluntariness of use was not an issue in this case, and thus was not measured. Gender was also not really an issue in the two domains of the present research. Almost all district nurses are female and almost all train drivers are male, so analyses on gender are not relevant. Therefore, gender was also omitted as a key moderator. Moreover, as the user experiments are part of the first phases of user-centred design, detailed analyses on the moderated effect of age and experience are outside the current scope. The aim is to identify performance and effort expectations and social influence of the concept applications. In later phases, the effect of age and experiences became relevant.

User acceptance of the information and communication support was measured by the behavioural intentions to use the new technology in work environments. To be able to understand and explain the reasons behind the participants' intention to use, the performance expectancy, the effort expectancy and social influence were measured. *Figure 4-3* depicts the simplified UTAUT model for the present research.

Figure 4-3 Simplified UTAUT research model



UTAUT identified four statements to measure performance expectancy and four statements to measure effort expectancy (Venkatesh et al., 2003). UTAUT also identified four statements for social influence. The fourth statement relates to whether the organisation has supported the participant in using the system. As that was not relevant to the user experiments, this statement was omitted from the questionnaire. The general concept of behavioural intentions to use technology in the UTAUT statements was replaced in each experiment with a dedicated term, respectively use of WoundLog, and radar use. In total, eleven statements concerning user acceptance were posed.

UTAUT uses a 7-point Likert scale to indicate the level of agreement with the statements. To maintain consistency over the various questionnaires in the experiments, the agreement with the statements was provided on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). The statements that were used in the dynamic train simulation experiment (radar use) are given in *Appendix A*.

Applicability

Professionals who have experienced information and communication support for their task performance can give their perceptions of the applicability of the concept to practice. Applicability relates to domain-specific questions about the designs tested. As the designs of the two cases were totally different (WoundLog application versus context information display), dedicated questionnaires for each experiment were developed.

For home care, information preferences of the history information, and the expected positive and negative effects of using it in practice were asked for. For railway, perceived usefulness of each information elements and its predicted use in practice were asked for.

4.6 Summary

This chapter identified the measurement instruments for assessing task performance in both experimental tasks. The relevant constructs identified were task performance, task support, task processes, and user acceptance and applicability. For each of these constructs, several measurement instruments were developed. *Table 4-8* provides an overview.

Table 4-8 Overview measurement instruments for task performance

Construct	Home care	Railway
Task performance	**Quality of care**: treatment decision-making **Efficiency of care**: perception questionnaire **Patient friendliness**: perception questionnaire **Job satisfaction**: empowerment adapted version of (Spreitzer, 1995)	**Safety**: SPAD, ATP interventions, reaction time dead man's handle, mental workload (BSMI: (Zijlstra et al., 1995) **Punctuality**: log data **Energy costs**: log data **Passenger satisfaction**: perception questionnaire **Job satisfaction**: empowerment adapted version of (Spreitzer, 1995)
Task support	Adapted version of PETT (Wilson et al., 2001)	Adapted version of PETT (Wilson et al., 2001)
Task processes	**SA**: dedicated perception, comprehension, projection questionnaire **Mental Workload**: BSMI: (Zijlstra et al., 1995)	**SA**: dedicated perception, comprehension, projection questionnaire **Mental Workload**: BSMI: (Zijlstra et al., 1995)
User acceptance	Adapted version of UTAUT (Venkatesh et al., 2003)	Adapted version of UTAUT (Venkatesh et al., 2003)
Applicability	Case-specific questions	Case-specific questions

Chapter 5

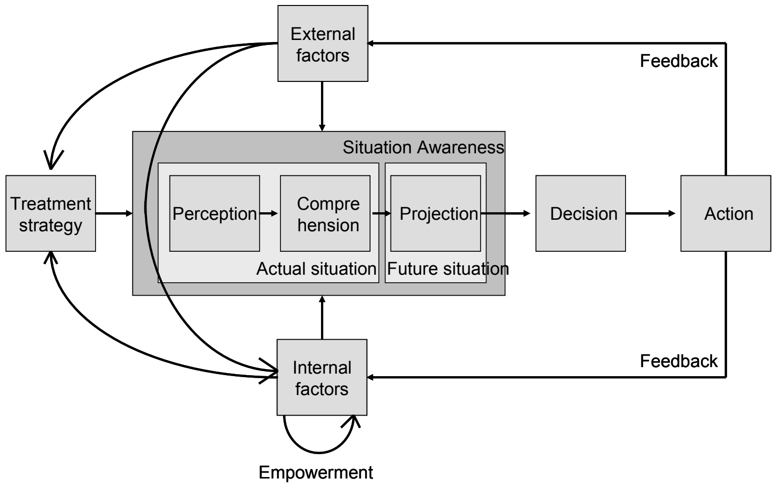
The effect of information and communication support on task performance in home care

This chapter describes the two experiments used to test the hypotheses about the information and communication support to improve the treatment decision-making by district nurses. Treatment decision-making is a relevant task for wound care. As technology is not part of the everyday practice of most district nurses, the first experiment was a paper survey in which the district nurses had to make treatment decisions. The two experimental conditions were with and without history information. The second experiment was a user experiment, in which the district nurses had to make treatment decisions, conditionally supported by technology. A 2x2 within-subjects experimental design was applied: two information conditions (with and without information) and two communication conditions (with and without communication). The results indicate that history information does not improve treatment decision-making, although the district nurses expect that it will have a positive effect on the quality, efficiency and patient-friendliness of wound care. Moreover, the district nurses felt empowered by the history information, which might serve as a stimulus to use the history information in practice.

5.1 Treatment decision-making

Treatment decision-making is relevant to wound care. Possible treatment decisions are treatment according to plan, physician needs to be consulted later on today, and physician needs to be consulted immediately. If the delivered care is not successful, or the patient also suffers from something else, the sooner the physician can adjust the treatment plan, the sooner the patient will be cured. To make appropriate treatment decisions, the district nurse has to use and integrate information from various sources. *Figure 5-1* depicts the wound care performance model.

Figure 5-1 Wound care task performance model



The various elements of the wound care performance model are given in *Table 5-1* (see also *Table 3-3*).

Table 5-1 Components of wound care performance model

Component	**Elements**
Treatment decision-making strategy	• Analytical decision-making • Analytical-intuitive decision-making • Intuitive-analytical decision-making • Intuitive decision-making
External factors	• Patient condition • Wound area • Wound information
Internal factors	• Intuitive knowledge • Technical knowledge • Evidence-based knowledge • Mental workload
Decisions	• Recognition of the situation • Deviation from treatment plan • Communication
Actions	• Treatment according to plan • Call physician or experienced colleague later on • Call physician immediately

The general research question is whether history information can improve home care performances. History information extends the external factors. Applying this question to the wound care task performance model results in the following assumptions:

- History information will change perception, and also comprehension and projection.
- Improved projection relates to better anticipation, and will improve decision-making.
- Improved decision-making will improve actions.
- History information will improve the utilisation of knowledge and abilities (psychological empowerment = internal factor).
- Increased empowerment is the increased intrinsic task motivation to perform a task.

Technology is not part of the everyday practice of most district nurses. Using WoundLog might interfere with using the history information WoundLog provides. Therefore, the first experiment was a paper survey, in which district nurses had to make treatment decisions with and without history information. The hypothesis tested in the survey was:

Hypothesis 1 *History-based wound information improves treatment decision-making.*

In the second experiment, WoundLog was used to provide history information, and to support communication between two district nurses. First the same hypothesis was tested as in the survey, but now with the history information provided by technology. Second, communication support was provided. The hypotheses tested in the user experiment were:

Hypothesis 1 *History-based wound information improves treatment decision-making.*

Hypothesis 2 *The combination of history-based wound information and communication support improves treatment decision-making.*

For practical reasons, all participating subjects took part in both experiments. To get acquainted with the use of history information without being distracted by the technology, all subjects first completed the survey, and then took part in the experiment. Comparing the results of the survey and the experiment tested hypothesis 3:

Hypothesis 3 *Technology is an interfering factor for treatment decisions of district nurses.*

5.2 Survey

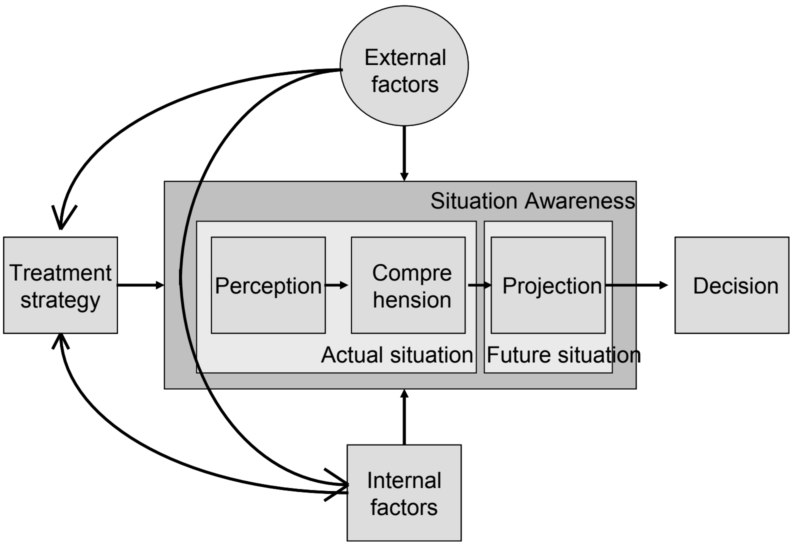
Technology is not part of the everyday working practice of most district nurses. Testing the effect of history information on treatment decision-making with WoundLog might have been disturbed by this inexperience. Therefore, the concept of history information on treatment decision-making was first tested without technology. In the survey, district nurses had to indicate treatment decisions based on authentic patient material, conditionally supported by history information. In relation to the task performance model of wound care, providing history information added to the external factors. The assumptions were that perception (and thus comprehension) and projection change, which results in better decisions. The second assumption was that history information empowers the district nurses, which results in improved task performance. The goal of the survey was to test the hypothesis:

Hypothesis 1

History-based wound information improves treatment decision-making.

In the survey, district nurses had to decide about 'paper patients'. Therefore the general task performance model was adjusted: actions and feedback were omitted (*Figure 5-2*), and history information (external factor) was added. To emphasise these changes in external factors, the square was changed into a circle. The district nurses also had to complete pre- and post-survey questionnaires about the main constructs of the research. The survey was in paper format, as it was used to test the information concept of WoundLog (wound logbook) without using technology.

Figure 5-2 Wound care performance model



5.2.1 Method

Experimental design

To test the effect of the use of history information, treatment decision-making without history information was compared with treatment decision-making supported by history information. The information condition was the independent variable and had two conditions: with history information and without history information (current situation). For practical reasons, half of the patient's material (n=9) consisted of history information in textual format, and the other half (n=9) of history information in textual format, accompanied by wound pictures. The patients' material was delivered in a stratified way: four normal treatments, and four intervention-needed treatments (two semi-urgent and two urgent).

Participants

The home care organisation Livio distinguishes five levels of district nurse qualification (levels 1 to 5). The higher the level, the more qualified medical nursing activities are allowed. WoundLog refers to nursing activities on level 3, 4 or 5. Therefore, district nurses from those levels participated in the experiment. Livio selected the district nurses for the experiment. Computer experience was not necessary for participation. The municipality of Enschede was the customer of the project and paid for the participation of the district nurses. The municipality was interested in the results of using WoundLog, as part of its Care and Technology programme. The results of the experiment were used to improve WoundLog and to facilitate a pilot study.

Twenty district nurses (19 female and 1 male) participated in the survey. One was a level-3 district nurse, ten were at level 4 and nine were at level 5. Their mean age was 44.1 (standard deviation (SD) =11.7). On average they had 10.5 years of experience as a district nurse (SD=8.1), and their paid working hours were on average 25.9 hours a week (SD=6.9). Thirty per cent of them hardly ever used a personal computer (PC), 40% used a PC weekly and 30% used a PC on a daily basis. PC use was mostly related to Internet and e-mail (85%), word-processing (65%) and games (60%). Instant messaging and digital photos were hardly ever used (5-15%). Only one district nurse was familiar with video conferencing (5%).

Tasks

First, the district nurses had to complete questionnaires about such personal matters as age, district nursing level and experience, and computer experiences. They also had to complete pre-survey questionnaires about the main constructs of the research, that is, health care performance indicators, empowerment and task support. They then had to provide treatment decisions for eight cases: two normal treatments and two intervention-needed treatments (two semi-urgent, two urgent). The patient material was provided in randomised order, supported by the different history information conditions. Finally, post-survey questionnaires were conducted on the same constructs as the pre-survey questionnaires, and information preferences were asked for.

5.2.2 Results

Treatment decisions were analysed on case level, and questionnaires on district nurse level.

Treatment decisions

The treatment decisions of the district nurses in all conditions were compared to the NSU (normal, semi-urgent and urgent) and NI (normal and intervention) frames of references. The results of the treatment decisions first had to be coded according to each framework. If the decision matched the frame of reference, 1 was scored; if there was no match, 0 was scored. Treatment decisions were analysed on case level, so the coded scores of all district nurses per case were added up and then divided by the total number of treatment decisions for that case. Paired samples t-tests were conducted to compare the experimental conditions of this parametric data. The mean and the standard deviation of the treatment decisions per condition are given (*Table 5-2*), as are the results of the paired samples t-tests. As the NI framework had only two scoring possibilities, the mean is higher than for the NSU framework with three scoring possibilities.

Table 5-2 Mean and SD of correct treatment decision per condition per frame of reference

Correct treatment decisions		
	NSU Mean (SD)	NI Mean (SD)
Without information	0.52 (.33)	0.78 (.29)
History information	0.58 (.40)	0.87 (.29)

Paired samples t-tests with the results of the NSU frame of reference showed no significant differences between the information conditions, namely with or without history information (p=0.601). Paired samples t-tests with the results of the NI frame of reference also showed no significant differences between the information conditions, namely with or without history information (p=0.191).

These results allow the rejection of hypothesis 1, which means that history-based wound information does not affect treatment decision-making, regardless of the frame of reference taken.

Questionnaires

Questionnaires were completed before and after the survey. The measured constructs were: contribution and status of the health care performance indicators; decision-making and empowerment; task-support; and information preferences. The results of the questionnaires are given below.

Health care performance indicators

The contribution and status of the three performance indicators were tested pre- and post-survey. The district nurses had to indicate their contribution to each performance indicator on a 5-point Likert scale (1= no contribution, 2= little contribution, 3= average contribution, 4= much contribution and 5= very personal contribution). For the status of the three constructs, report marks between 0 and 10 were asked for (0= very bad, 10= excellent). Pre- and post-survey questionnaires were compared to identify the perceived effect of history information on the status and contribution of the three performance indicators. First, the mean and standard deviation of the contribution and status of the pre- and post-survey of each performance indicator are given (*Table 5-3*). Second, these non-parametric data were statistically compared with the Wilcoxon signed-ranks test.

Table 5-3 Pre- and post-survey results of the perceived contribution and status of the health care performance indicators

Health care performance indicators						
	Quality of health care		Efficiency of health care		Patient-friendliness of health care	
	Pre-survey mean (SD)	Post-survey mean (SD)	Pre-survey mean (SD)	Post-survey mean (SD)	Pre-survey mean (SD)	Post-survey mean (SD)
Contribution (1-5)	3.80 (.41)	4.00 (.33)	3.45 (.51)	3.95 (.23)	3.85 (.37)	3.05 (.23)
Status (0-10)	7.15 (.81)	8.28 (.57)	6.66 (.88)	8.00 (.49)	6.95 (.76)	8.00 (.34)

For quality of health care, the Wilcoxon signed-ranks test showed a significant difference between the pre- and post-survey results for the status (Z=-3.407, N=19, p=0.001).

For efficiency of health care, the Wilcoxon signed-ranks test showed significant differences between the pre- and post-survey results for both the personal contribution and the status, respectively (Z=-2.714, N=20, p=0.007), and (Z=-3.896, N=19, p=0.000).

For patient-friendliness of health care, the Wilcoxon signed-ranks test showed significant differences between the pre- and post-survey results for both the personal contribution and the status, respectively (Z=-2.000, N=20, p=0.046), and (Z=-3.542, N=19, p=0.000).

The results show that district nurses expect that using history information will have a positive effect on the status of the quality, efficiency and patient-friendliness of health care. They also expect that their personal contribution will have a positive effect on the efficiency of care, and a negative effect on the patient-friendliness of care. This expected negative effect might be caused by the district nurses' expectations that their patient visits will be more time-consuming if history information is involved, and that the patients will have to pay for these prolonged visits.

Empowerment

Empowerment relates to the meaning, competence, self-determination and impact of the working activities (Spreitzer, 1995). The district nurses had to indicate their agreement with four statements on the constructs on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). To compare the pre- and post-test results, the post-test results were coded as: -2, -1, 0, 1 or 2. The sum of the pre-test figures and the coded post-test are listed as the post-test results. First, the mean and standard deviation of the constructs were listed (*Table 5-4*). Second, statistical analyses were performed.

Table 5-4 Pre- and post-survey results of empowerment

Empowerment		
	Pre-survey Mean (SD)	Post-survey Mean (SD)
Meaning	4.65 (.59)	5.65 (1.14)
Competence	4.42 (.84)	5.68 (0.89)
Self-determination	4.05 (.92)	5.20 (1.15)
Impact	3.30 (.92)	4.30 (1.34)

The Friedman test allowed the rejection of the null hypothesis that all tasks were equally supported by the history information. Thus, history information and communication affected task support ($N=19$, $\chi=73.446$, $df=7$, $p=0.000$).

Detailed statistical analyses (Wilcoxon signed-ranks test) showed significant differences between the pre- and post-survey results for meaning ($Z=-2.954$, $p=0.003$), competence ($Z=-3.619$, $p=0.000$), self-determination ($Z=-3.581$, $p=0.000$) and impact ($Z=-3.397$, $p=0.001$).

The results indicate that district nurses perceive that history information will increase the meaning, their competences, their self-determination and the impact of their work.

Task support

A post-survey questionnaire was conducted to identify what kinds of tasks are supported by the history information. Tasks that are generally supported by history information are planning, team performance, complex decision-making and negotiation (Wilson et al., 2001). The district nurses had to indicate their agreement with four statements on task support on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). The scores are given (*Table 5-5*) and the results of the statistical analyses are described.

Table 5-5 Post-survey results of task support

Task support post-survey		
	Mean	SD
Planning	4.00	0.92
Team performance	4.00	1.08
Complex decision-making	4.60	0.60
Negotiation	4.89	0.32

The Friedman test allows rejection of the null hypothesis that all tasks were equally supported by the history information. Thus, history information and communication affected task support (N=19, χ=16.008, df=3, p=0.001).

Detailed analyses (Wilcoxon signed-ranks test) showed significant differences between negotiation and planning (Z=-3.002, p=0.003), complex decision-making and team performance (Z=-2.072, p=0.038), and negotiation and team performance (Z=-2.859, p=0.004).

The results show that district nurses regard the history information as more suitable for negotiation than for planning or team performance, and more suitable for complex decision-making than for team performance.

Information preferences

In the post-survey questionnaire, the district nurses were asked about their perceptions of the kinds of information they received for treatment decision-making. All twenty participating district nurses preferred the most enriched information condition: history information in the form of both a textual report and visual (picture) information. Nineteen stated that they could make the best treatment decision with all information provided, while one district nurse stated that she could make the best treatment decision with only textual information. They all least liked to work with current information condition: without history information.

5.2.3 Summary of the results

The goal of the survey was to test the hypothesis that history-based wound information improves the treatment decision-making by district nurses, without using technology. The secondary goal was to identify the district nurses' perceptions of this concept. Treatment decisions were scored according to two different frames of references: one based on three treatment decision categories (NSU), and one based on two treatment categories (NI). Hypothesis 1 was therefore rejected: history information does not affect the treatment decision-making by district nurses, tested with both frames of references.

Treatment decisions were not affected by history information, although the questionnaire data showed that district nurses expected that the history information would improve the quality, efficiency and patient-friendliness of health care. The history information would also improve their personal contribution to the efficiency of care, but would have a negative effect on the patient-friendliness of health care (due to prolonged visits, paid for by patients).

The district nurses also expected to be empowered by the history information: the meaning, their competences, their self-determination and the impact of their work would increase when using history information. In relation to task support, they expected that the history information would be more suitable for negotiation and complex decision-making than for planning or team performance.

5.3 User experiment

A user experiment with WoundLog was conducted, in which district nurses had to indicate treatment decisions of authentic patient material, conditionally supported by history information and communication. In relation to the wound care task performance model (see *Figure 5-3*), providing history information adds to the external factors, as does providing communication support.

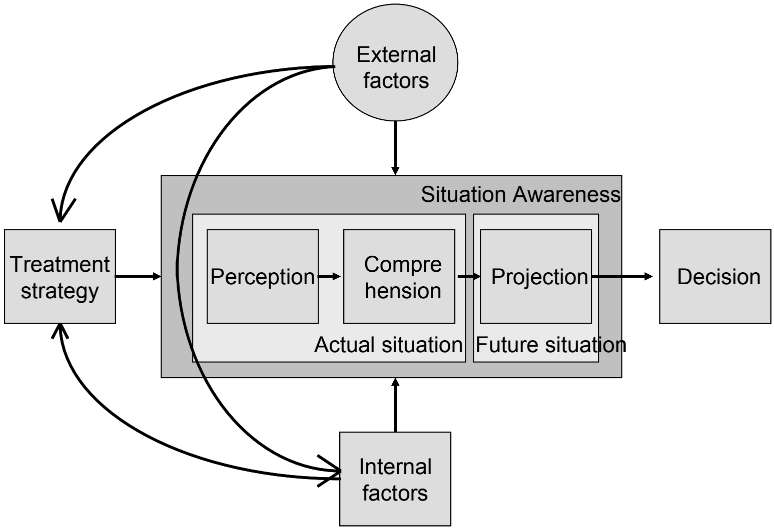


Figure 5-3 Wound care performance model

The assumptions were that perception (and thus comprehension) and projection change, which results in better decisions. The second assumption was that history information empowers the district nurses, which results in improved task performance. As only treatment decisions were measured in the present research, actions, and feedback were omitted from the model. To emphasise the changes in the external factors, the square figure was replaced by a circle. The goals of the user experiment were to test the following hypotheses:

Hypothesis 1 *History-based wound information improves treatment decision-making.*

Hypothesis 2 *The combination of history-based wound information and communication support improves treatment decision-making.*

Hypothesis 3 *Technology is an interfering factor for the treatment decision-making by district nurses.*

5.3.1 Method

Experimental design

A 2x2 within subjects design was applied. The information conditions (2) were with and without history information. The communication conditions (2) were with and without communication. *Table 5-6* depicts this experimental design. The method applied to each condition of information and communication support is also presented in the table. The number of patient cases per method is given between brackets.

Table 5-6 2x2 experimental design

		Information support	
		No information	Information
Communication support	No communication	Paper survey (4)	Wound logbook (3)
	Communication	Mobile phone (4) and Paper survey (4)	WoundLog (3)

'Paper survey' refers to the results of the previous sections, which were also used in the experiment. 'Mobile phone' refers to the mobile phones district nurses use to communicate with each other about patients. 'Wound logbook' refers to the use of the history information in WoundLog. 'WoundLog' refers to using both the history information and the communication properties of the WoundLog application.

The patient material was delivered in randomised order and in a stratified way to each information and communication condition. The ratings of the expert district nurse were used for the stratification of the patient material. For the paper and wound logbook tasks, normal and intervention-needed patient material were provided in randomised order. For the mobile phone and WoundLog tasks, only intervention-needed patient materials were provided, as only those patients were discussed in practice. Three kinds of technology support were provided to accomplish the experimental tasks: mobile phones, WoundLog and wound logbook.

Participants

The participants in the experiment were the same as in the survey.

Tasks

Ten sessions with two district nurses each were conducted. During a three-hour session, they performed treatment decision tasks with authentic patient material, under various information and communication conditions. Tasks without information and communication support were conducted by paper and pencil. In the survey, four patient material situations (= cases) without history information were conducted. The results of these four cases were used as the no information and no communication task (= paper survey task) for the experiment. The four cases with history information in the survey were used to identify the effect of testing the concept with technology (hypothesis 3). The technology-supported tasks in the experiment are the wound logbook (information support), mobile phone (communication support) and WoundLog (information and communication support). For all tasks, treatment decisions (normal, semi-urgent or urgent) had to be made.

After the survey, the same group of district nurses performed the experiment. Therefore, the pre-survey questionnaires also served as pre-experiment questionnaires. The questionnaires were about personal aspects and the main constructs of the research, namely health care performance indicators, empowerment and task support.

The district nurses received 30 minutes of training in the use of WoundLog. The focus of the training was on using the wound logbook, and on the set-up, acceptance, process and camera handling in teleconferencing. After that, they had to accomplish three different tasks in randomised order: wound logbook, mobile phone and WoundLog. Finally, post-experiment questionnaires were conducted on the same constructs as the pre-survey questionnaires.

During the multi-user tasks (mobile phone and teleconferencing), both district nurses communicated about the patient material, and after that, they each completed their own questionnaire about treatment decision and case complexity. Furthermore, each district nurse completed questions about their perceived mental workload and their situation awareness.

After having performed all the tasks, they had to complete a set of post-experiment questionnaires on their perceptions of the health care performance indicators, empowerment, user acceptance and applicability of WoundLog.

5.3.2 Results

The same district nurses participated in both the survey and the user experiment. Results about their age, experience, use of computers can be found in *Section 5.2.1*. The patients' material was delivered in a stratified way to the district nurses. Treatment decisions were analysed on case level, and the questionnaires on district nurse level.

Treatment decisions

Treatment decisions of all four experimental information and communication support conditions were compared with both frames of references (NSU and NI). For these analyses, the results were coded according to the two frameworks. These data were ordered on case level. To resemble authentic tasks, the cases in the multi-user sessions were about semi-urgent and urgent patients. Therefore, the results of the experiment cannot be related to general information and communication conditions and their interactions. Thus, only paired samples t-tests can be performed between two corresponding conditions. The means and standard deviations of the treatment decisions per condition for each frame of reference are given in *Table 5-7*.

Table 5-7 Correct treatment decisions per experimental condition per frame of reference

Correct treatment decisions		
	NSU frame of reference Mean (SD)	NI frame of reference Mean (SD)
Paper survey without information	0.52 (.33)	0.78 (.29)
Paper survey with information	0.58 (.40)	0.87 (.18)
Wound logbook	0.46 (.30)	0.76 (.21)
Mobile phone	0.52 (.43)	0.96 (.09)
WoundLog	0.52 (.41)	0.93 (.15)

To test the effect of history information on treatment decision-making, the results of conditions with and without information were compared, regardless of the technical and communication condition. The paper survey with and without information had already been analysed. The paired samples t-tests did not reveal statistical differences, respectively NSU (p=0.601), NI (p=0.191). A paired samples t-test compared the differences between the paper surveys without information and the wound logbook results (with information). No statistical differences were found, respectively NSU (p=0.691) and NI (p=1.000). Paired samples t-tests between the results of mobile phone conferences (without information) and WoundLog conferences (with information) also revealed no significant differences, respectively NSU (p=1.000), and NI (p=0.520).

To test the effect of history information and communication support on treatment decision-making, the results between mobile phone conferences (without information) and WoundLog conferences (with information) were compared. As stated, paired samples t-tests did not reveal significant differences, respectively for NSU (p=1.000), and NI (p=0.520).

The results indicate that no effect of information and communication support can be identified on treatment decision-making. These results led to the rejection of the two hypotheses. History information does not affect treatment decision-making. History information and communication support do not affect treatment decision-making. It is likely that the main reason for the lack of results on treatment decision-making is related to the measurement error in identifying standards for treatment decision-making (see *Section 4.2.1 on the frame of references*).

Cooperative tasks

Each pair of district nurses had to discuss six patients: three by mobile phone and three by WoundLog. The patients were all from the category 'intervention needed' (semi-urgent and urgent), according to the two frames of references. For each patient, one of the district nurses had the task to initiate a conference with the other district nurse to get advice about the treatment decision related to a certain patient. After each patient-based conference (six in total), both district nurses had to complete a questionnaire. The questions were about their treatment decision, their perceived mental workload, situation awareness and decision-making in cooperation.

Treatment decisions made by mobile phone and WoundLog had already been compared: the paired samples t-tests showed no significant differences, respectively for NSU (p=1.000), and NI (p=0.520).

The treatment decisions of each pair of district nurses were compared. As they had discussed a patient to identify the best treatment, their treatment decisions should have been the same. For mobile phones (N=8 pair), 83% of the treatment decisions were identical; for WoundLog (N=10 pairs), 67% of the treatment decisions were identical.

These results indicate that history information has no effect on treatment decision-making in cooperative tasks (communication provided). However, the amount of overlap in treatment decisions is higher for mobile phones than for WoundLog conferences (not statistically tested).

Situation awareness

Situation awareness (SA) was operationalised according to Endsley (Endsley, 1995b). She divides SA into perception, comprehension and projection. After each cooperative task, the district nurses had to indicate these aspects in relation to the case under discussion: the relevant elements of the patient (perception), the relevant elements of the case (comprehension) and the prognosis of the patient (projection). The idea was to identify the effect of history information on the three different aspects of SA. It could also be helpful in identifying in which stage history information was incorporated correctly, and where it possibly failed. This means that it should provide insights into the decision-making process. To avoid making it a rehearsal task, the questions were not posed as multiple-choice but as open-ended questions.

The results showed that these answer possibilities are perhaps too open-ended. The answers of the two district nurses that discussed a patient case varied a lot on the aspects mentioned. *Table 5-8* provides the results of district nurses 6 and 7, who discussed patient case 10, supported by mobile phone.

Table 5-8 Prototypical examples of the results of SA measurements

Nurse 6, Case 10, Mobile phone			Nurse 5, Case 10, Mobile phone		
Wound information	**Patient condition**	**Wound healing process**	**Wound information**	**Patient condition**	**Wound healing process**
Size:	Pain scale:	I do not know	Size: bigger	Pain scale:	When treatment is appropriate, wound will heal, but will remain vulnerable
Depth:	Nutrition:		Depth:	Nutrition: bad	
Colour:	Fever: yes		Colour: red and yellow coating	Fever: yes	
Smell: penetrating smell	Diabetes:		Smell:	Diabetes:	
Pain: acute pain	Mobility: Not mobile		Pain:	Mobility:	
Wound area: red around wound			Wound area: red around toe		

These results show that comparisons, both in a quantitative and qualitative way, were hard to make. In relation to the aspects measured, an illustrative observation was made during the experiment. After having received a mobile phone call for information on a patient, the district nurse hung up and said out loud: 'Did the patient suffer from diabetes or not?' This indicated exactly the kinds of aspects that the SA questionnaire was intended to measure, namely the kinds of information on which the district nurses based their decisions. The variety of results might indicate various aspects. District nurses find it hard to verbalise pre-phases of their decision-making. Post-task measurement was inadequate; information processing steps are hard to remember. Intuitive treatment decision-making strategies might have been applied, which makes it hard to verbalise. The answer categories did not match the pre-phases of decision-making. Or, each district nurse makes decisions on her own way, which makes those processes hard to compare.

However, the number of times history information was mentioned in the answers was counted, for example 'the pain is getting worse' or 'the size is getting smaller'. In the mobile phone tasks (47 answers on patient cases), history information was mentioned six times. As no history reports or pictures were available during the mobile phone tasks, it is likely that they were part of the patient description. In WoundLog tasks (58 answers on patient cases), history information was mentioned 29 times. This indicates that the only result from the SA analyses is that there is an indication that some district nurses really read the patient's history information. This does not indicate that they also used the history information for their treatment decision-making.

Mental workload

Mental workload was measured after each cooperative task. After each patient-discussion, each district nurse had to indicate her mental workload according to the BSMI (Zijlstra et al., 1995). In both conference situations, one district nurse was the initiator of the conference and the other district nurse the receiver. The initiator and receiver tasks were randomised between the district nurses: each district nurse was initiator three times and receiver three times. The initiation task (both mobile phone and WoundLog) required more actions and communication. The receiver task required more listening to and understanding the information. The initiator and receiver task might indicate differences in mental workload. *Table 5-9* lists the results according to the medium (mobile phone and WoundLog) and the task type (initiator or receiver).

Table 5-9 Post-task results of mental workload

Mental workload		
	Mobile phone Mean (SD)	WoundLog Mean (SD)
General	51.50 (31.83)	45.61 (18.70)
Initiator	62.13 (41.13)	49.59 (20.36)
Receiver	40.88 (19.15)	47.55 (23.87)

Paired samples t-tests did not reveal significant differences between mental workload for mobile phone and WoundLog conferences ($p=0.375$). Paired samples t-tests between the initiator and receiver of the conference differed significantly for mobile phone initiator and receiver ($t=2.289$, $df=23$, $p=0.032$), but not for WoundLog ($p=0.672$).

These results indicate that the perceived mental workload of district nurses is the same whether they use a mobile phone or WoundLog. When looking at the initiator and receiver of the conference task, the mobile phone initiator perceives more mental workload than the receiver of the mobile phone conference. This seems to be task dependent, as the initiator of the mobile phone conference has to give the receiver all information about the wound; the receiver only has to listen. In WoundLog conferences, both initiator and receiver of the conference call can see and actively explore the wound information.

Questionnaires

Health care performance indicators

The status of and their contribution to the three performance indicators (quality, efficiency and patient-friendliness of health care) were tested pre-survey and post-experiment. The district nurses had to indicate their relevance and contribution to each construct (3) on a 5-point Likert scale (1= no contribution, 2= little contribution, 3= average contribution, 4= large contribution and 5= very personal contribution). For the status of the three constructs, report marks between 0 and 10 were asked for (0= very bad, 10= excellent). Pre- and post-survey questionnaires were compared to identify the effect of extra information on the relevance and the contribution of the three constructs. The mean and standard deviation of the pre-survey and post-experiment results of each performance indicator are given (*Table 5-10*), as are the results of the Wilcoxon signed-ranks test for non-parametric data.

Table 5-10 Pre-survey and post-experiment results of the perceived contribution to and status of the of health care performance indicators

Health care performance indicators						
	Quality of health care		Efficiency of health care		Patient-friendliness of health care	
	Pre-survey mean (SD)	Post-experiment mean (SD)	Pre-survey mean (SD)	Post-experiment mean (SD)	Pre-survey mean (SD)	Post-experiment mean (SD)
Contribution (1-5)	3.80 (.41)	4.00 (.00)	3.45 (.51)	4.00 (.00)	3.85 (.37)	4.00 (.46)
Status (0-10)	7.15 (.81)	8.30 (.57)	6.66 (.88)	7.95 (.69)	6.95 (.80)	8.35 (1.04)

For the quality of health care, the Wilcoxon signed-ranks test revealed significant differences between the pre-survey and the post-experiment results for both contribution and status, respectively ($Z=-2.000$, $N=20$, $p=0.046$), and ($Z=-3.360$, $N=20$, $p=0.001$).

For the efficiency of health care, the Wilcoxon signed-ranks test revealed significant differences between the pre-survey and the post-experiment results for both the contribution and the status, respectively ($Z=-3.317$, $N=20$, $p=0.001$), and ($Z=-3.778$, $N=20$, $p=0.000$).

For the patient-friendliness of health care, the Wilcoxon signed-ranks test revealed a significant difference between the pre-survey and the post-experiment results of the status ($Z=-3.450$, $N=20$, $p=0.001$).

These results show that district nurses expect that the use of WoundLog will have positive effects on the quality, efficiency and patient-friendliness of health care. They also expect a positive effect on their personal contribution to the quality and efficiency of health care.

Empowerment

Empowerment relates to the meaning, competence, self-determination and impact of the working activities (Spreitzer, 1995). The district nurses had to indicate their agreement with four statements on the constructs on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). Again, pre- and post-survey questionnaires were conducted to identify the effect of history information on the various constructs. To compare the pre- and post-test results, the post-test results were coded as -2, -1, 0, 1 or 2. The sum of the pre-test figures and the coded post-test are listed as the post-test results. *Table 5-11* lists the mean and standard deviation of the constructs; the statistical analyses are then described.

Table 5-11 Pre-survey - and post-experiment results of empowerment

Empowerment		
	Pre-survey Mean (SD)	Post-experiment Mean (SD)
Meaning	4.65 (.59)	5.40 (1.1)
Competence	4.42 (.84)	5.00 (1.62)
Self-determination	4.05 (.92)	4.55 (1.00)
Impact	3.30 (.92)	3.75 (1.41)

The Friedman test led to the rejection of the null hypothesis that all cognitions were equally supported by the history information and communication ($N=18$, $\chi=62.448$, $df=7$, $p=0.000$).

The Wilcoxon signed-ranks test on the individual constructs revealed significant differences between the pre-survey and the post-experiment results for meaning ($Z=-2.830$, $p=0.005$), competence ($Z=-3.087$, $p=0.002$), self-determination ($Z=-2.500$, $p=0.012$), and impact ($Z=-3.019$, $p=0.003$).

The results indicate that the district nurses think that the meaning, their competences, their self-determination and the impact on their work will increase when using WoundLog in their everyday practice.

Task support

A post-experiment questionnaire was conducted to identify what kinds of tasks are supported by the use of WoundLog. Typical group tasks for information and communication support are planning, team performance, complex decision-making and negotiation (Wilson et al., 2001). The district nurses had to indicate their agreement with four statements on task support on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). The scores are given in *Table 5-12*, and the statistical analyses are described.

Table 5-12 Post-experiment results of the Perceived Effectives of Task Type support for the information and communication support

Perceived Effectiveness of Task Type		
	Mean	SD
Planning	3.25	1.12
Team performance	3.80	1.15
Complex decision-making	4.40	0.88
Negotiation	4.60	0.60

The Friedman test led to the rejection of the null hypothesis that all group tasks were equally supported by the history information. History information and communication support affected group task support ($N=20$, $\chi=21.130$, $df=3$, $p=0.000$).

Detailed analyses (Wilcoxon signed-ranks test) showed significant differences between team performance and planning ($Z=-2.041$, $N=20$, $p=0.041$), complex decision-making and planning ($Z=-2.810$, $N=20$, $p=0.005$), negotiation and planning ($Z=-3.361$, $N=20$, $p=0.001$), and negotiation and team performance ($Z=-2.437$, $N=20$, $p=0.015$).

The results show that the district nurses perceive using WoundLog as less suitable for planning, and more suitable for negotiation than for team performances.

User acceptance

User acceptance was measured by an adapted version of the UTAUT questionnaire (Venkatesh et al., 2003). UTAUT measures the acceptance of new technology by identifying the behavioural intentions to use the technology. The relevant constructs for the intention to use in this experiment were performance expectancy (4 questions), effort expectancy (3 questions) and social influence (3 questions). The UTAUT questionnaire consists of 10 statements to which the district nurses had to respond on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). *Table 5-13* lists the results of the various constructs of UTAUT.

Table 5-13 Post-experiment results of UTAUT

Unified Theory of Acceptance and Use of Technology		
	Mean	SD
Performance expectancy	3.54	0.59
Effort expectancy	3.68	0.81
Social influence	3.08	1.06

The Friedman test led to the rejection of the null hypothesis that all constructs are equally important for the intention to use. Intentions to use differed in their importance (N=20, χ=6.427, df=2, p=0.040).

A Wilcoxon signed-ranks test revealed significant differences between social influence and performance expectancy (Z=-2.096, p=0.036), and between social influence and effort expectancy (Z=-2.093, p=0.036).

These results indicate that the behavioural intentions of district nurses to using WoundLog will be more based on the performance and effort expectancies than on social influences. To increase the intention to use, social influences should be involved in training.

Information preferences

In the post-experiment questionnaire, the district nurses were asked about their perceptions of the kinds of information they received for treatment decision-making. Nineteen of the nurses preferred the most enriched information condition, namely history information provided both textually (report) and visually (pictures). The other nurse preferred the history information without pictures. All twenty said that they could make the best treatment decision if all the information is provided. Eighteen least liked to work with the current information condition (without history information), while two least preferred the most enriched information condition. As this latter figure is opposite to the most preferred information condition, likely reading mistakes took place.

Applicability of WoundLog

In the post-experiment questionnaire, the district nurses had to indicate their expectations about the positive and negative effects of using WoundLog in practice. The main results reveal that 80% of the district nurses expected a positive effect of WoundLog on the quality and efficiency of health care, 70% a positive effect on patient-friendliness, and 50% a positive effect on education and knowledge development of the nurses.

Considering the negative effects of and hindrances caused by using WoundLog in practice, 50% of the district nurses expected problems with patient acceptation and technical issues (mainly camera handling) and carrying and taking care of the mobile device (65%). No district nurse expected a negative effect of using WoundLog on the quality of health care.

These results indicate that the main expected positive effects of using WoundLog are the quality and efficiency of health care. Client acceptation, camera handling and taking care of the apparatus were the main expected negative effects.

5.3.3 Summary of the results

The goals of this experiment were to test the effect of history information on the task performance of both single- and multi-user tasks. Task performance was operationalised as treatment decision-making. Moreover, a secondary goal was to identify the district nurses' perceptions of using WoundLog. Treatment decisions were scored according to two frames of references: one based on three treatment categories (NSU), the other based on two treatment categories (NI).

The two information conditions (with and without history information) did not affect treatment decisions, tested with both frames of references. These results led to the rejection of hypothesis 1.

The combination of history information and communication support also did not affect treatment decision-making. These results led to the rejection of hypothesis 2. The treatment decisions of two district nurses (who discussed a patient) differed more when they used WoundLog than the mobile phone. This might indicate that using WoundLog distracted the nurses from their main task, namely discussing a patient. The mental workload figures do not support this explanation, however: no differences between mobile phone and WoundLog. Another explanation is that district nurses perceive a wound picture differently (see also results frame of reference). Explaining a wound to a colleague by phone means that the mental picture or perception of the wound is conveyed (= subjective picture), rather than an objective picture.

Treatment decisions were not affected by history information or by the combination of history information and communication. However, the questionnaire data showed that district nurses expected that the history information would improve the quality, efficiency and patient-friendliness of health care. The history information would also improve their contribution to the quality and efficiency of care.

The district nurses also expected to be empowered by the history information: the meaning, their competences, their self-determination and the impact of their work would increase when using history information, which results in increased intrinsic motivation for task performance. In relation to task support, they expected that the history information would be less suitable for planning and more suitable for negotiation than for team performances.

When applying WoundLog in practice, some district nurses expect problems with patient acceptation, camera handling and taking care of the apparatus.

5.4 Summary

The research objective of these experiments was to establish whether using WoundLog improves health care performances. The health care performances identified are quality and efficiency of care, patient and job satisfaction. First, the results of both experiments are related to these performance indicators, and then the detailed hypotheses for the experiments are discussed.

Performance indicators

Quality of care refers to the quality of the treatment decisions (see *Section 2.5.4*). The results of the treatment decisions of both experiments show that history information and the combination of history information and communication do not affect treatment decisions. The district nurses perceived that the history information and communication would increase the status of quality of care, and in the experiment district nurses also perceived that their personal contribution to the quality of care would increase.

Efficiency of care refers to the efficiency of the wound healing process and the efficiency of the activities of several people (fewer extra visits by expert district nurse or physician, fewer exhausting visits by the patient to hospital). The activities of people involved in wound care were not measured in the experiment. The efficiency of the wound healing process relates to the quality of care characteristics of treatment decision-making (see above). Thus, the only results measured for efficiency of care are the perceptions of the district nurses. In both experiments, the district nurses expected that history information and the combination of history information and communication would have a positive effect on efficiency of care, and on their personal contribution to the efficiency of care.

Patient friendliness again refers to the wound healing process and to the visits by several people. Thus, the results of the treatment decisions are again applicable. The only specifically measured results for patient friendliness are the perceptions of the district nurses. In the two experiments, district nurses expected that the history information and the combination of history information and communication would have a positive effect on the patient friendliness of care. Whereas in the survey district nurses expected that the history information would decrease the patient friendliness, in the user experiment, they expected no effect.

Job satisfaction refers (again) to treatment decision-making, but also to coaching, feedback and learning on the job. The empowerment results of both experiments revealed that district nurses feel empowered and thus intrinsically motivated to work with the history information and the combination of history information and communication.

Hypotheses testing

The overall objective – namely to establish whether using WoundLog improves health care performances – was refined in three hypotheses:

Hypothesis 1 *History-based wound information improves treatment decision-making.*

Hypothesis 2 *The combination of history-based wound information and communication support improves treatment decision-making.*

Hypothesis 3 *Technology is an interfering factor for treatment decisions of district nurses.*

These hypotheses were tested in three methodological ways: survey, user experiment and a comparison of the two. Two kinds of data were collected: treatment decisions and questionnaire data.

To test the effect of using technology on treatment decision-making (hypothesis 3), the results of the paper surveys with history information were compared with the wound logbook results. The paired samples t-test did not reveal significant differences, respectively for NSU ($p=0.291$), and for NI ($p=0.165$). Again, the results of the mobile phone conferences (no new technology) and the WoundLog (new technology) conferences were applicable to this hypothesis. As stated, the paired samples t-test did not show significant differences, respectively for NSU ($p=1.000$), and NI ($p=0.520$).

Analyses of treatment decisions led to the rejection of all three hypotheses. The main reason for this is probably related to the measurement error in identifying standards for treatment decision-making (see *Section 4.2.1 on the frame of references*).

Analyses of questionnaire data revealed that the district nurses have high expectations of WoundLog. They expect it to improve the quality, efficiency and patient-friendliness of health care. Their personal contribution to these indicators might also improve the quality and efficiency of care. Considerations are put on patient acceptance and technology handling.

WoundLog mainly supports negotiation and complex decision-making, and is less suitable for planning and team performance. The district nurses expect to be empowered by WoundLog. They intend to use WoundLog because of their efficiency and productivity expectations, and because they find it easy to use and they have skilful interactions with it. The social support for using WoundLog can be increased.

Thus, in general the experiments did not prove that WoundLog improves health care performances. District nurses expect it to improve the main health care performance indicators. Possible reasons for the positive feelings are that it is easy to use, they have efficiency and productivity expectations, and they are intrinsically motivated to use it, which might serve as a stimulus.

Chapter 6

The effect of information and communication support on task performance in railway

This chapter describes the two experiments that were used to test the hypotheses that context information improves the task performance of train drivers and the communication between train drivers and railway signallers. To make the train driving task less complex, the first experiment was a static simulation of train driving decision-making: dynamic aspects and time pressure were excluded. The three experimental conditions were actual situation, context information about the actual situation, and context information about the future situation. For practical reasons, these conditions had to be provided in a fixed order. The second experiment was a PC-based multi-user simulation in which train drivers had to ride various routes, conditionally supported by context information. A 3x2 within-subjects experimental design was applied: three context information conditions (actual situation, context information about own train and context information about neighbouring trains) and two route conditions (normal and abnormal routes). The results indicate that context information does not affect safety, punctuality or energy costs, although train drivers expect that it will have a positive effect on the status of punctuality, energy costs and passenger friendliness. They perceived that their own contribution to safety will decrease. Moreover, they felt empowered by the context information, which might serve as a stimulus to use context information in practice.

6.1 Train driving

Train driving is a dynamic control and decision-making task (Kecklund et al., 2001). To perform this task, the driver has to use and integrate information from several sources, both internal and external. *Figure 6-1* depicts the task performance model of train driving.

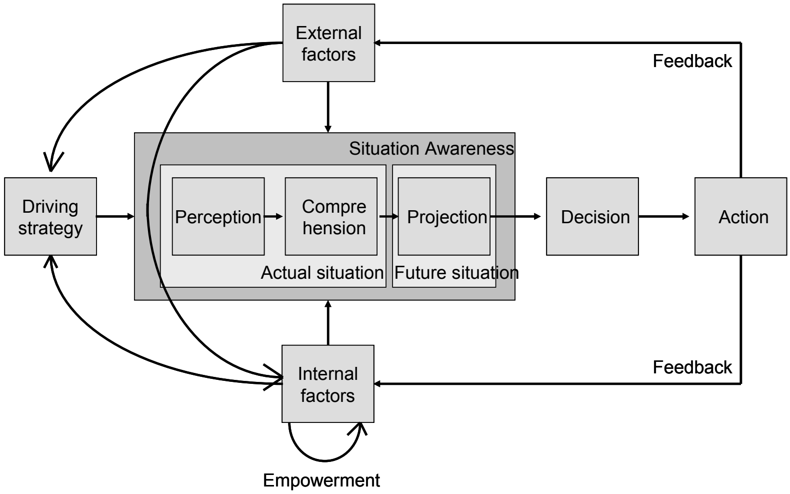


Figure 6-1 Task performance model of train driving

Table 6-1 lists the various elements of the task performance model (see also *Table 3-3*).

Table 6-1 Components of train driver performance model

Component	Elements
Driving strategy	• Feedback driving strategy • Feed-forward driving strategy
External factors	• Cabin information • Line side and signals • Temporary procedures from railway signaller
Internal factors	• Knowledge and experiences, e.g. route knowledge, procedures • Mental workload
Decisions	• Control of speed • Communication
Actions	• Apply/remove power and brakes • Operate voice communication

The main research question was whether context information can improve railway performances. Context information is a form of additional external information. Applying this question to the task performance model (see *Figure 6-1*) results in the following assumptions:

- External information will change perception and consequently comprehension and projection.
- Improved projection will improve decision-making.
- Improved decision-making will improve actions.
- External information will improve the utilisation of knowledge and abilities (psychological empowerment = internal factor).
- Increased empowerment is the increased intrinsic task motivation to perform a task.

Prototypical situations of normal and abnormal scenarios were sketched, and several experts (i.e. train drivers, railway signallers, trainers and e-learning specialists) were asked what context information was necessary to improve the railway performances. Two types of information were relevant: information about the train driver's own train and information about other trains. From their own train, it became clear that in practice a lot of information is available as route knowledge, for example punctuality and path predictions. Therefore, the train drivers indicated that they wanted to have information about upcoming signals, and to a lesser extent about punctuality. Unfortunately, route knowledge was not available in the experiments. Therefore, path predictions and punctuality had to be provided to overcome this. Regarding other trains, train drivers wanted to have information about their upcoming signals, path predictions, punctuality, speed, type of train (i.e. train number) and order of trains.

Identifying the effect of context information on railway performances is complicated. A lot of interrelated factors are involved in this complex socio-technological system, for example information is dynamic, time pressure is high and there are competing goals. Therefore, in the first experiment, the effect of context information on railway performances was identified, while dynamic information aspects and time constraints were excluded; in other words, it was a static simulation experiment. The hypothesis tested in this experiment was:

Hypothesis 1 *Context information improves train driving decision-making.*

The second experiment concerned real-life situations of train driving; dynamic information and time pressure were added. A user experiment was performed with a multi-user PC-based train simulator. The hypotheses applied to this experiment were:

Hypothesis 1 *Context information improves train driving decision-making.*

Hypothesis 2 *Context information improves communication between train drivers and railway signaller.*

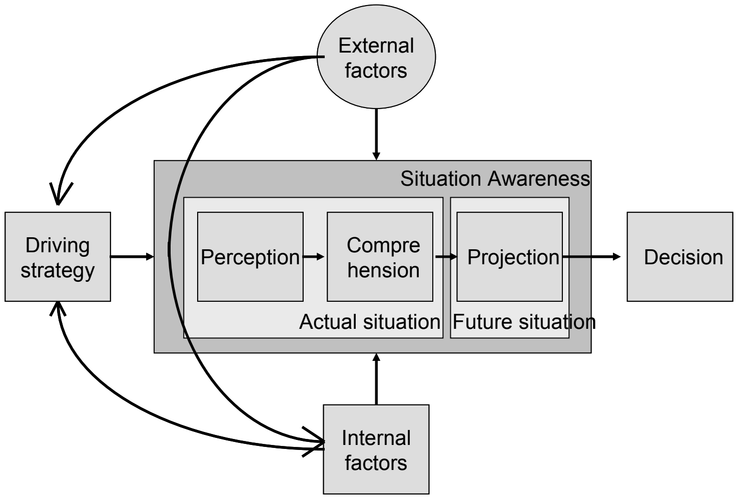
To identify the effect of omitting dynamic information and time pressure in the first experiments, the results of both experiments were 'generally' compared (the performance indicators were measured in different ways). The hypothesis tested was:

Hypothesis 3 *Dynamic information and time pressure do not affect train driving decision-making.*

6.2 Static simulation experiment

This experiment was focused on train driving decision-making, without dynamic information and time pressure. Based on the overall task performance model of train driving, a simplified research model was developed (see *Figure 6-2*). As it was a static simulation experiment, actions (and also feedback) were not available. The external factors dynamic information and time pressure were omitted, and new external information (context information) was added. To emphasise these changes in external factors, it is depicted as a circle.

Figure 6-2 Simplified task performance model of train driving for static experiment



Context information conditions were the input variables, and driving decisions were the output variables. Information processing and empowerment were measured by questionnaires. The main hypothesis for the present research was:

Hypothesis 1

Context information improves train driving decision-making.

The train drivers also had to complete pre- and post-survey questionnaires about the main constructs of the research. The interviews and workshops with train drivers and railway signallers identified the difference between context information about the actual situation, and context information about the future situation. These two situations were taken as conditions of context information. The idea behind this is that information about the future increases projection and thus strategic decision-making.

6.2.1 Case material

Two context information designs were prepared to test the hypothesis. The context information display is called ‘Radar’. Radar I focused on actual context information and Radar II on context information about the future. The two Radar designs were an extension of the information a train driver currently views in the real-life situation.

At the moment, a train driver views on-board information and such outside information as signs and signals (= external information). Radar I and II enhanced this situation by providing context information. Radar I provided information about the context of the situation: surrounding tracks and trains, upcoming signals, speed, delay and number of other trains. Radar II provided the same information as Radar I, complemented with information about the future (= path predictions and orders of trains.).

For practical reasons, the order of information provision in the experiment was fixed for all participants: from no context information via Radar I to Radar II (information is enriched). The practical reasons were that train drivers had to participate during their lunch or coffee break (maximum 30-45 minutes), and that train drivers from various locations participated in the study, so route knowledge could not be available in the scenarios. Therefore, only prototypical scenarios were developed, in which typical route knowledge was not relevant. The train drivers in the workshop decided that – as route knowledge could not be available – all normal and abnormal scenarios boiled down to the same thing: it was not useful to develop several scenarios (Biemans, Swaak, Van der Velde & Huppertz, 2005). The normal scenario was about an international train (high priority train) running a few minutes late, while all other trains were on schedule. The abnormal scenario was about a train in a congestion or traffic jam.

Example of normal scenario

In the normal scenario, train drivers got static simulations of the following case, accompanied with a picture (=actual situation) (*Figure 6-3*):

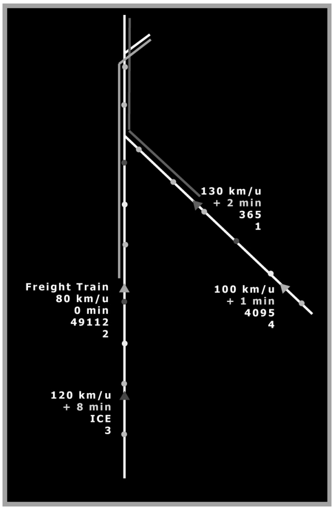
Figure 6-3 Overview of the actual situation of the normal scenario (signal switches between red and yellow)

You are driving a High Speed Train
Maximum track speed 140 km/h
After a number of signals
there will be a junction
Your delay is 8 minutes (+8 min),
your speed is 120 km/h

After having seen the case description and the picture(s), the train drivers were asked what kind of driving decision they would take. Subsequently, this situation was enhanced with context information (respectively Radar I and Radar II) and the same question was asked. *Figure 6-4* provides the example of Radar II. It depicts for all trains the tracks, signals, path predictions, train numbers (or freight and ICE), speed, delay and order of trains.

Figure 6-4 Overview of normal scenario and Radar II

You are driving a High Speed Train
Maximum track speed 140 km/h
After a number of signals
there will be a junction
Your delay is 8 minutes (+8 min),
your speed is 120 km/h



6.2.2 Method

Experimental design

A 3x2 within-subjects design was applied, comprising three information conditions (actual situation, Radar I and Radar II) and two scenarios (normal and abnormal). The main hypothesis for this study was that context information improves railway performances. Railway performances are safety, punctuality, energy costs, job satisfaction and passenger satisfaction. As depicted in the research model (*Figure 6-2*), context information directly affects decision-making through better projection, and indirectly through empowering train drivers. This static simulation was also used to identify the train drivers' perceptions of the use of context information in their everyday practices.

Participants

Fifty-seven train drivers (55 male and 2 female) participated in the study on a voluntary basis during their coffee or lunch breaks. On average, their age was 45.0 years (SD=8.8), and they had 18.3 years (SD=9.9) of train driving experience. This is a representative sample of the Dutch railway population, as on average their age was 46.5 years, and on average their experience as a train driver was 20.6 years.[6]

Tasks

Train drivers viewed two static simulations with the three conditions of context information in fixed order. They had to indicate their driving decision and whether they would inform their passengers. Pre- and post-test questionnaires were conducted to measure the train drivers' perceptions of the effect of context information on the performance indicators. Furthermore, questionnaires were used to gain insights into empowerment, the decision-making process and information preferences.

6.2.3 Results

The results are divided into two sections: those directly related to the driving decisions of the static simulations, and those related to the questionnaires.

Driving decisions

Driving decisions were measured after each scenario, supported by a certain type of information. The possible driving decisions were: maintain speed, brake, accelerate, call railway signaller, 'other'. Many train drivers indicated

[6] Personal communication *Nederlandse Spoorwegen,* 26 July 2005.

that they would 'keep on rolling' in the 'other' category, which means power off. Therefore, in the analysis, this was a separate category. The chosen driving decision was scored 1, and the other 0. Not only the kinds and the numbers of driving decisions were relevant, but also whether an individual train driver changed his decision upon receipt of enriched context information. The methodological set-up (fixed order of conditions) provided opportunities for an analysis of driving-decision transitions on an individual level. The assumption is that qualified experts take appropriate decisions in order to satisfice railway performances (Simon, 1957). *Table 6-2* depicts the mean and SD of driving decisions per information condition for the normal scenario.

Table 6-2 Overview of train driving decisions of the three information conditions in the normal scenario

Train driving decisions in the normal scenario			
	Actual situation Mean (SD)	Radar I Mean (SD)	Radar II Mean (SD)
Constant speed	0.07 (.26)	0.16 (.37)	0.16 (.37)
Brake	-	0.46 (.50)	0.47 (.50)
Accelerate	0.93 (.26)	0.02 (.13)	0.02 (.13)
Call railway signaller	-	0.04 (.19)	0.05 (.23)
Rolling	-	0.32 (.47)	0.28 (.45)
Other	-	0.02 (.13)	0.02 (.13)

The results of the three information conditions were statistically compared with the Wilcoxon signed-ranks test. The actual situation differed significantly from Radar I on braking ($Z=-4.491$, $p=0.00$), accelerating ($Z=-5.980$, $p=0.00$) and rolling ($Z=-3.638$, $p=0.00$). The actual situation also differed from Radar II on these same driving decisions; braking ($Z=-4.600$, $p=0.00$), accelerating ($Z=-5.857$, $p=0.00$) and rolling ($Z=-3.207$, $p=0.01$). Radar I and Radar II did not differ statistically from each other. All other comparisons were not significantly different.

In general, enriched context information resulted in different driving decisions. The difference in context information between Radar I and Radar II did not affect decision-making. Enriched context information resulted in a divergence of driving decisions; from two different decisions in the actual situation to six different decisions in the radar conditions.

Table 6-3 depicts the mean and SD of train driving decisions per information condition for the abnormal scenario.

Table 6-3 Overview of train driving decisions of the three information conditions in the abnormal scenario

Train driving decisions in the abnormal scenario			
	Actual situation Mean (SD)	Radar I Mean (SD)	Radar II Mean (SD)
Constant speed	0.11 (.31)	0.07 (.26)	0.05 (.23)
Brake	0.58 (.50)	0.70 (.46)	0.84 (.37)
Accelerate	0.18 (.38)	0.09 (.29)	-
Call railway signaller	0.14 (.35)	0.14 (.35)	0.11 (.31)
Rolling	-	-	-
Other	-	-	-

The results of the three information conditions were statistically compared with the Wilcoxon signed-ranks test. The actual situation differed significantly from Radar I on acceleration ($Z=-2.000$, $p=0.046$).

The actual situation differed from Radar II on braking ($Z=-3.357$, $p=0.001$) and accelerating ($Z=-3.742$, $p=0.00$). Radar I differed from Radar II on accelerating ($Z=-2.449$, $p=0.014$). All other comparisons were not significantly different.

In general, enriched context information resulted in different driving decisions for the two types of radars. These driving decisions converged from four different decisions in the actual situation to three different decisions in the Radar II condition.

The results support hypothesis 1, which means that context information leads to different driving decisions. Remarkably, in the normal scenario, the variety of decisions was enlarged by the context information, while in the abnormal scenario the variety of decisions was reduced by the context information. This might originate from the different goals: in an abnormal situation, the train driver tries to catch up with his timetable. In a normal situation, the train driver has various goals to aim at, for example energy costs and punctuality.

Questionnaires

Railway performance indicators

In pre-test and post-test questionnaires, the perceptions of the contribution and status of the railway performance indicators (i.e. safety, punctuality, energy costs and passenger satisfaction) were measured. The pre-test questions were about perceptions of the actual situation, as the train driver experiences in everyday practice. The post-test questions were about perceptions as the train drivers would be working with the context information in their everyday practice. The train drivers had to indicate their contribution to each construct on a 5-point Likert scale (1= no contribution, 2= little contribution, 3= average contribution, 4= much contribution, and 5= very personal contribution). For the status of the four constructs, report marks of between 0 and 10 were asked for (0= very bad, 10= excellent). *Table 6-4* depicts the results of the pre- and post-test questionnaires.

Table 6-4 Pre-test and post-test results of the perceived contribution and status of the railway performance indicators

Perceptions of railway performance indicators								
	Safety		Punctuality		Energy costs		Passenger satisfaction	
	Pre-test mean (SD)	Post-test mean (SD)	Pre-test mean (SD)	Post-test mean (SD)	Pre-test mean (SD)	Post-test mean (SD)	Pre-test mean (SD)	Post-test mean (SD)
Contribution (1-5)	4.05 (.51)	3.54 (1.05)	3.77 (.76)	3.95 (.61)	3.56 (.91)	4.05 (.55)	3.86 (.81)	3.86 (.67)
Status (0-10)	7.37 (1.38)	7.88 (1.13)	6.91 (.94)	7.81 (.97)	6.26 (1.35)	7.98 (1.11)	6.58 (1.19)	7.96 (.91)

The results of the pre- and post-test were compared with the Wilcoxon signed-ranks test. For safety, significant differences were found for the train driver's contribution (Z=-3.106, p=0.002) and the status (Z=-3.008, p=0.003). For punctuality, a significant difference was found for status (Z=-4.495, p=0.000). For energy costs, significant differences were found for status (Z=-5.099, p=0.000), and the train driver's contribution to energy costs (Z=-3.508, p=0.000). For passenger satisfaction, a significant difference was found for the status (Z=-5.591, p=0.000).

These results show that train drivers expected that using context information would have positive effects on the status of all railway performance indicators (i.e. safety, punctuality, energy costs and passenger satisfaction). Moreover, they expected their personal contribution to energy costs to increase, and their personal contribution to safety to decrease.

Empowerment

Empowerment is defined as the increased intrinsic task motivation manifested in a set of four cognitions reflecting an individual's orientation towards his or her work role; these cognitions are meaning, competence, self-determination and impact of the working activities (Thomas et al., 1990). The train drivers had to indicate their agreement with four statements on empowerment (Spreitzer, 1995) on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). In the post-test questionnaires, they had to indicate whether the context information would increase these four cognitions. To compare the pre- and post-test results, the post-test results were coded as: -2, -1, 0, 1 or 2. The sum of the pre-test figures and the coded post-test are listed as the post-test results. *Table 6-5* depicts the results of the pre- and post-test on the 4 cognitions of the empowerment construct.

Table 6-5 Pre- and post-test results of empowerment

Pre- and post-test results of empowerment		
	Pre-test mean (SD)	Post-test mean (SD)
Meaning	4.16 (.82)	4.95 (1.25)
Competence	4.72 (.49)	5.67 (.95)
Self-determination	3.30 (1.13)	3.89 (1.56)
Impact	2.84 (1.08)	3.14 (1.84)

To identify the effect of the use of context information on empowerment, the results of the pre- and post-test were compared. The Friedman test allowed the rejection of the null hypothesis that all cognitions were equally supported by the context information (N=57, χ=213.305, df=7, p=0.000).

Detailed analyses with the Wilcoxon signed-ranks test showed significant differences for meaning (Z=-5.646, SD=0.000), competence (Z=-5.776, SD=0.000) and self-determination (Z=-4.605, SD=0.000). The results for impact were not significant (p=0.062).

These results indicate that train drivers expect a positive effect of the context information on the meaning, competence and self-determination of their work, and no effect on the impact of their work.

Task support

The perceived effectiveness of task type, the context information provided was measured with the PETT (Wilson et al., 2001). Post-test, the train drivers had to indicate their agreement with four questions on task support on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). *Table 6-6* depicts the results.

Table 6-6 Post-test results of the Perceived Effectiveness of Task Type of the information support

Perceived Effectiveness of Task Type		
	Mean	SD
Planning	3.89	0.84
Group performance	3.82	1.05
Complex decision-making	3.70	0.93
Negotiation	3.30	1.03

The Friedman test led to the rejection of the null hypothesis that all group tasks were equally supported by the context information. Context information support affected group task support (N=56, χ=19.911, df=3, p=0.000).

Wilcoxon signed-ranks tests between the four types of tasks showed significant differences between negotiation and planning (Z=-3.786, p=0.000), negotiation and group performance (Z=-3.018, p=0.003), and negotiation and complex decision-making (Z=-2,713 p=0.007).

These results indicate that train drivers perceive the context information to be less suitable for negotiation than for the other three tasks (planning, team performance and complex decision-making).

Information elements

After each scenario, the train drivers were asked about the perceived usefulness of the various information elements provided. *Table 6-7* depicts these results for the normal and abnormal scenario.

Table 6-7 Results of the perceptions of usefulness of information elements

Perceptions of usefulness of information elements		
	Normal scenario Mean (SD)	Abnormal scenario Mean (SD)
Surrounding tracks	0.51 (.50)	0.47 (.50)
Surrounding trains	0.72 (.49)	0.70 (.46)
Speed of trains	0.56 (.50)	0.68 (.47)
Path predictions of trains	0.77 (.42)	0.82 (.38)
Order of trains	0.82 (.38)	0.47 (.50)

The results of both scenarios were statistically compared. The Wilcoxon signed-ranks test showed that the perceived usefulness differed significantly for speed of trains (Z=-2.141, p=0.035) and order of trains (Z=-4.095, p=0.000). No other analysis showed significant results.

These results indicate that train drivers perceive differences in the usefulness of the information elements between the normal and abnormal scenario for speed and order of trains.

Observations during static simulations

The simulations ran on two days and at two locations each day. During the simulation sessions, the following observations were made (Van der Velde et al., 2004).

- Discussions between train drivers afterwards revealed that not all train drivers immediately understand what they can do with the context information. '*What was the difference between Radar I and II? When you know path predictions, you can better anticipate. Oh, do we have to put it that way?'*
- Train numbers indicate the kinds of trains (local train, intercity train or freight train). Not all train drivers can translate the numbers into the kinds of trains (with specific relevant characteristics).
- Train drivers apply various driving strategies:
    - From three interviews with freight train drivers it became clear that freight train drivers focus on anticipation, because of their train characteristics. Based on their experiences and route knowledge, they develop various kinds of possible driving scenarios. They continuously match the available context information with their scenarios and drive according to the most likely scenario.
    - In contrast, some train drivers confessed that they only 'drive one signal away'.
- Some train drivers were frightened that providing context information could introduce a 'Big Brother' effect: '*All my colleagues can see what I am doing, and what my performance is, that is, punctuality and ATP interventions.*'

6.2.4 Summary of results

The goal of the static simulation experiment was to test the effect of context information on train driver's decision-making (task performance), without dynamic information and time pressure. In a static simulation, a 3x2 within-subjects design was applied, comprising three information conditions and two route conditions. All train drivers received the information in a fixed order.

In general, enriched context information led to different driving decisions on both normal and abnormal routes. As experts make satisficing decisions (Simon, 1957), changes in driving decisions between information conditions indicate that better decisions are made, which will increase railway performances. Thus, these results support hypothesis 1. On normal routes, driving decisions changed upon context information about the actual situation, while context information about the future situation did not affect decision-making. On abnormal routes, context information about both the actual and the future situation affected decision-making.

Another difference between both routes is that the kinds of decisions diverged in normal situation and converged in abnormal situations. It is likely that in normal situations various decisions can be taken to drive safely and on time (perceived most relevant performance indicators), while under abnormal situations the aim is to be punctual (and of course safe), thus more coherent driving decisions should be taken.

Train drivers expect context information to have a positive effect on all railway performance indicators (i.e. safety, punctuality, energy costs, personnel and passenger satisfaction). Moreover, they perceive that their personal contribution to energy costs increases (better anticipation), while their personal contribution to safety decreases. This decrease is because the context information display in their cabin might distract their attention from the outside information, so they could miss relevant line-side signs, which would decrease safety. The train drivers' personal contribution to railway performances relates to empowerment. For empowerment a positive effect was found. Train drivers expected that context information would increase meaning, competence, self-determination and impact of their work.

Train drivers used various information elements in both scenarios. Speed of trains was perceived as more useful in the abnormal scenario, while the order of trains was perceived as more useful in the normal scenario.

In this experiment all train drivers were given two driving situations, with enriched information in fixed order. Therefore, the results provide only an indication that context information can have an effect on train driving decision-making. What might be of more importance is that train drivers perceive that context information can affect railway performances, and that they feel empowered, and thus intrinsically motivated by using the context information.

6.3 Dynamic simulation experiment

The static simulation tested the general hypothesis about the effect of context information on railway performances, while dynamic aspects and time constraints were excluded. In this experiment, context information support for train drivers was tested in a more naturalistic setting, namely a multi-user simulation with dynamic information and time constraints. In the simulator, both train drivers and a railway signaller took part. Thus, also interaction or communication between train drivers and railway signallers could take place. The following hypotheses were applied to this experiment:

Hypothesis 1

Context information improves train driving decision-making.

Hypothesis 2

Context information improves communication between train drivers and railway signallers.

One of the main findings from the static simulation is that the effect of context information on safety was perceived differently. In general, the drivers found that context information would increase the overall safety positively, while it would decrease their personal contribution to safety, mainly because their attention would be distracted. As safety is a crucial construct for innovation research in railway, this concept needs more study.

While testing the simulator with train drivers, it became evident that some train drivers found the Radar cluttered. The combination of safety and cluttered display provided the context information conditions for the dynamic experiment. Radar I provided the relevant context information elements of their personal train, and Radar II provided these elements of all trains. This made Radar II more cluttered than Radar I. Moreover, the interviews and workshops with train drivers and railway signallers revealed that route knowledge is extremely relevant for railway decision-making. In neither simulation (static and dynamic) could route knowledge be incorporated into the case material. To compensate for the lack of route knowledge, path predictions and the order of trains are always available on Radar.

6.3.1 Case material

A PC-based multi-user train simulator was developed to test the effect of context information railway performances. In several interactive sessions with train drivers, it was decided to keep the user interface of the train drivers' cabin in the simulator very schematic. Train drivers emphasised that otherwise they would keep looking for mismatches with reality. Moreover, the train drivers indicated that the combination of the schematic view and the control handles (2 joysticks) provided enough realism to give them the idea that they were really driving a train. *Figure 6-5* provides a screenshot of train driver's cabin and outside view in the simulator. *Figure 6-6* provides a screenshot of the user interface of the railway signaller in the simulator.

Figure 6-5 Screenshot of train driver's cabin in the simulator

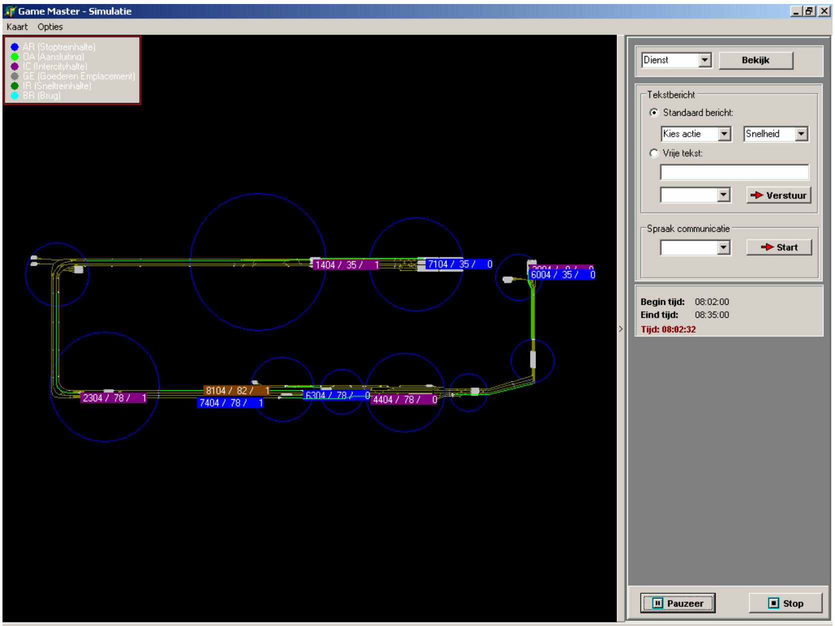


Figure 6-6 Screenshot railway signaller's part of the train simulator

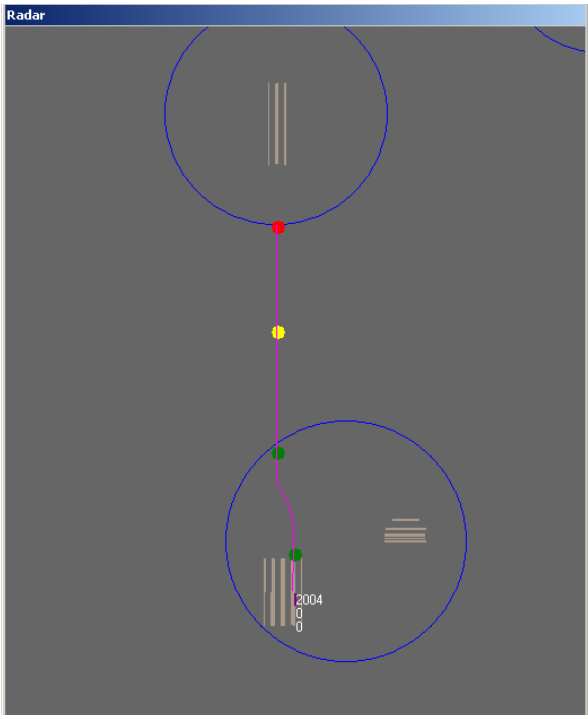
The information elements on both radars are given in *Table 6-8*. As route knowledge was not available in the simulation, path predictions were provided.

Table 6-8 Overview of the information elements in both Radars

Radar I: Information about own train	Radar II: Information about all trains
Upcoming signals	Upcoming signals
Path predictions	Path predictions
Punctuality	Punctuality
Speed	Speed
	Order of trains
	Type of train (train number)

Figure 6-7 depicts a screenshot of Radar I. The blue circles are the areas under the control of railway signallers; the other parts (the 'free route') do not contain switches and are controlled automatically. The green, yellow and red coloured bullets are the upcoming signals. The pink line is the path prediction of the train. 2004 is the train number, and the zeros refer to the speed and punctuality of the train. The 2004 is standing still at a station and is on time.

Figure 6-7 Screenshot of Radar I



The same colours and figures are used in Radar II (see *Figure 6-8*). Radar II also depicts the information about other trains. The path predictions of each train have a different colour, their corresponding signals are presented with a white spot in the middle (to contrast with the signals from their own train). Train drivers have the possibility of zooming in and out on the screen.

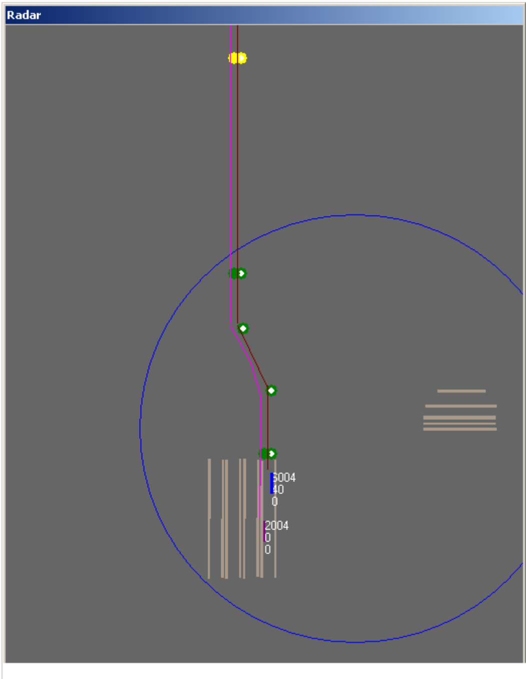


Figure 6-8 Screenshot of Radar II

6.3.2 Method

Experimental design

A 3x2 within-subjects design is applied of three information conditions (actual information, Radar I and Radar II), and two types of scenarios (normal and abnormal). Three instantiations of each type of scenario were available. In each session, four train drivers drove on a simulated track called Texel. Each train driver had to perform six tasks (= rides on Texel), consisting of randomised information condition and scenario (3x2=6). The Texel track was based on the Dutch track between Rotterdam and Dordrecht. All participating train drivers had route knowledge of that particular track. Only one train driver recognised the origin of the Texel track; the others believed that they were driving on a fictive track.

In each task, both human trains and robot trains were available. Robot trains drove according to their timetables. The four human trains in each task were driven by train drivers.

Participants

Forty-eight train drivers (47 male, 1 female) and three railway signallers (2 male, 1 female) participated in the present research. Twelve simulation sessions were conducted. Each session took four hours and involved the participation of four train drivers and one railway signaller. For planning reasons, each railway signaller took part in several sessions. This was an advantage because their part of the simulator was totally different from what they are used to (training time was longer). Train drivers and railway signallers participated on a voluntarily basis. The experiment took place during their normal working hours. As only three railway signallers took part in the experiment, their questionnaire data were not used for the analysis.

On average, the train drivers were 45.6 years old (SD=8.1). Their average working experience was 17.6 years (SD=9.9) (NS average =20.6 years). 31 of the 48 train drivers used to drive freight trains. Their average freight train driving experience was 11.1 years (SD=5.5). Five train drivers had also taken part in the static simulations. 33 (69%) of the train drivers used their computer on a daily basis, 12 (25%) on a weekly basis, 2 (4%) on a monthly basis and 1 (2%) hardly ever uses a computer.

Tasks

At the start of each four-hour session, the train drivers had to complete a pre-test questionnaire with general questions about their age, experience and computer use, and questions about their empowerment and the contribution and the status of the railway performance indicators (safety, punctuality, energy costs and passenger satisfaction).

Furthermore they received 30-minutes training in how to drive in the simulator. They drove on a simple track to get to know how to handle the joysticks and to get used to the user interface of the train. They also drove with both radar displays.

After having completed each ride of the experiment, the train drivers had to indicate their perceived mental workload and the information elements they used.

After having performed all six rides, they had to complete a post-test questionnaire. The questions concerned the quality, the contribution and the status of railway performance indicators, assuming they could use the context information, empowerment, user acceptance, the applicability of the context information in practice and their preferences for the various information elements.

6.3.3 Results

The results are divided in two sections, namely those directly related to the driving performances in the train simulator and those related to the questionnaires. On the first day of the experiment, a lot of computer crashes occurred. Therefore, the log data of that day were omitted from the analysis. The simulator crashes occurred 7 to 12 minutes into each ride. The train drivers got an impression of the context information provided. Therefore, the questionnaire data of those train drivers was incorporated in the questionnaire analysis.

Driving performances

The log data of the trains provided two types of driving performance data: punctuality and energy costs. Punctuality relates to the arrival and departure punctuality of trains, according to their timetables. Departure punctuality depends on the interaction between train driver and train guard and their passengers. As they were not part of the simulation, the departure punctuality was excluded from the analysis. Punctuality was computed as the difference between the measured arrival time and the arrival time stated on the timetable (in seconds).Two types of trains were involved in the experiment: local trains with stops at every station, and intercity trains, with one stop in the simulation. To make the punctuality of local and intercity trains comparable, the punctuality of the local trains was computed as the mean of the number of punctual arrivals per ride. *Table 6-9* provides the results for punctuality in seconds. For consistency reasons the means are described with two decimals. The punctuality results do not differ very much. The largest difference is only one minute.

Table 6-9 Punctuality results (in seconds) for route and information conditions

Punctuality		
	Mean	SE
Normal route	23.84	6.61
Abnormal route	96.62	6.87
Actual Situation	58.24	8.51
Radar I	62.95	8.40
Radar II	59.51	7.84

The means and SEs of the interaction between route and information conditions for punctuality are given in *Table 6-10*.

Table 6-10 Punctuality results of the interaction effects between route and information conditions

Interactions for punctuality			
Route	Information condition	Mean	SE
Normal	Actual	18.30	12.13
	Radar I	27.239	11.40
	Radar II	25.99	10.78
Abnormal	Actual	98.17	11.94
	Radar I	98.66	12.34
	Radar II	93.03	11.40

The statistical results of the main and interaction effects of route, information condition and route and information condition are given in **Error! Reference source not found.**.

Table 6-11 ANOVAs results on main and interaction effects for punctuality

ANOVAs main and interaction effects for punctuality			
	F-value	Df	P
Route	58.271	1	0.00
Information condition	0.084	2	0.92
Route x Information condition	0.156	2	0.86

ANOVAs (univariate analyses of variance) between the route conditions showed a significant difference (F=58.271, df=1, p=0.000). ANOVAs between the context information conditions did not show a significant difference (p=0.919). ANOVAs between the interaction of route and context information did not show a significant difference (p=0.855).

These results indicate that context information does not affect punctuality. Punctuality is affected only by route conditions, which originates from the conditions provided to the train drivers. Thus, in general, providing context information to train drivers does not affect punctuality in railway.

Energy costs relate to the acceleration (m/s^2) of trains: accelerations cost power and thus energy. Energy costs cannot be related to an external standard. It is only relevant to compare the different rides. Therefore, energy costs were analysed by comparing the mean accelerations per route and information condition. To make the various trains (Intercity and local train) and the various instantiations of the normal and abnormal routes comparable, the mean acceleration per instantiated route was measured (= reference route), regardless of the information condition. Each individual route was then compared to this reference route. The results of the differences between each route and its reference route were analysed on route and information condition. *Table 6-12* provides these results.

Table 6-12 Energy cost results for route and information conditions (in deviation per reference route).

Energy costs		
	Mean	SE
Normal route	0.000	0.002
Abnormal route	0.000	0.002
Actual Situation	-0.003	0.003
Radar I	0.004	0.003
Radar II	-0.001	0.003

The means and SEs of the interaction between route and information conditions for energy costs are given in *Table 6-13*.

Table 6-13 Energy cost results of the interaction effects between route and information conditions (in deviation per reference route).

Interaction effects for energy costs			
Route	Information condition	Mean	SE
Normal	Actual	0.001	0.004
	Radar I	0.003	0.004
	Radar II	-0.003	0.004
Abnormal	Actual	-0.006	0.004
	Radar I	0.006	0.004
	Radar II	0.001	0.004

The statistical results of the main and interaction effects of route, information condition and route and information condition are given in *Table 6-14*.

Table 6-14 ANOVAs results on main and interaction effects for energy costs

ANOVAs main and interaction effects for energy costs			
	F-value	Df	P
Route	0.003	1	0.954
Information condition	1.818	2	0.165
Route x Information condition	1.316	2	0.270

ANOVAs between the route conditions did not show a significant difference ($p=0.954$). ANOVAs between the context information conditions did also show no significant difference ($p=0.165$). ANOVAs between the interaction of route and context information did also show no significant difference ($p=0.270$).

These results indicate that energy costs are not affect by context information and by route conditions, or their interactions.

Signal passed at danger (SPAD)

SPADs (signal passed at danger) refers to passing a danger signal (= red signal). SPADs were related to route and information conditions. In total, four SPADs took place: two when no information was provided, one with Radar I and one with Radar II. For route conditions, two SPADs took place in normal routes, and two in abnormal routes. SPADs seem to be equally divided over the conditions. All four SPADs seemed to be caused by pulling the brakes too slowly, which might be caused by a lack of route knowledge.

Automatic Train Protection (ATP) interventions

In total, 32 ATP interventions took place – twice because the response to the dead-man's handle was too late, and 30 times because the speed limit was exceeded. Three train drivers got three ATP interventions, three got two ATP interventions, seventeen got one ATP intervention, and seventeen got no ATP intervention during their six rides. Fifteen of the ATP interventions took place on normal routes and seventeen on abnormal routes. Ten ATP interventions took place without context information, and eleven interventions in each Radar context. ATP interventions seem to be equally divided over the route and information conditions. Thus, likely no effects of route and context information were available.

Dead-man's handle

The dead-man's handle is a safety system. To indicate that the train driver is still able to drive his train, every sixty seconds he has to push a handle with his legs; he is prompted to do so by both a visual and an auditory signal. This was simulated in the simulator by pushing a button on the joy stick every sixty seconds. Every sixty seconds, the dead-man's blue light went on; after two seconds, it was accompanied by an auditory signal. If there was no reaction within five seconds after the auditory signal, an ATP intervention took place. The following illustrates the ecological validity of the simulator. During the experiment, one laptop went black. It became clear that the train driver had pushed the power cable out of the laptop as a reaction to the dead-man's signal.

The reaction time (RT) is considered to be an indication of mental effort: the more the train driver is mentally loaded, the slower his response will be to the dead-man's handle. As stated in the previous section, in total two dead-man's handle interventions took place. *Table 6-15* depicts the results of the reaction measurements (in seconds).

Table 6-15 Reaction time (in seconds) on the dead-man's handle per route and information condition

Reactions to dead-man's handle		
	Mean	SE
Normal route	1.45	0.017
Abnormal route	1.44	0.017
Actual Situation	1.36	0.021
Radar I	1.46	0.021
Radar II	1.52	0.020

The means and SEs of the interactions between route and information conditions for the reaction times are given in *Table 6-16*.

Table 6-16 Reaction time (in seconds) on dead-man's handle for interactions between route and information

Interactions for reactions to the dead-man's handle			
Route	Information condition	Mean	SE
Normal	Actual	1.34	0.031
	Radar I	1.48	0.029
	Radar II	1.52	0.030
Abnormal	Actual	1.38	0.030
	Radar I	1.43	0.029
	Radar II	1.51	0.027

The statistical results of the main and interaction effects of route, information condition and route and information condition are given in *Table 6-17*.

Table 6-17 ANOVAs results on main and interaction effects on reactions on dead-man's handle

ANOVAs main and interaction effects for reactions to dead-man's handle			
	F-value	Df	P
Route	0.026	1	0.873
Information condition	14.846	2	0.000
Route x Information condition	1.244	2	0.288

Reaction times (in seconds) were analysed for route and information conditions, and their interaction. ANOVAs between route conditions did not show significant differences ($p=0.873$). ANOVAs between the context information conditions showed a significant difference ($F=14.846$, $df=2$, $p=0.000$). ANOVAs for the interaction between route x context information also showed no significant difference ($p=0.288$).

These results indicate that providing context information increases the reaction time. Route conditions and the interaction between route condition and context information condition did not affect reaction times.

In general, the more context information is provided, the slower the response to the dead-man's handle. All means of the reaction times are within the limits of ATP interventions.

Mental workload

The train drivers had to indicate their perceived mental effort on the mental effort rating scale (BSMI in Dutch) after each ride. The scale goes from 0 (no effort at all) to 150 (112 is labelled 'very much effort').The mental workload scores are divided into rides 1 to 6, and on route and information conditions. *Table 6-18* provides the means and SDs for the perceived mental workload per ride.

Table 6-18 Results of the perceived mental workload per ride

Mental workload						
	Ride 1 Mean (SD)	Ride 2 Mean (SD)	Ride 3 Mean (SD)	Ride 4 Mean (SD)	Ride 5 Mean (SD)	Ride 6 Mean (SD)
Mental workload	42.66 (24.33)	34.73 (22.03)	34.02 (20.99)	31.61 (21.12)	28.56 (20.88)	27.52 (19.96)

The trend is that the level of perceived mental effort decreased over the rides. A paired samples t-test of the perceived mental effort between the first and the second ride differed significantly ($t=2.943$, $df=39$, $p=0.005$).

Table 6-19 provides the results for the perceived mental workload for the route and context information conditions.

Table 6-19 Results of perceived mental workload per route and information condition

Perceived mental workload		
	Mean	SE
Normal route	33.45	3.51
Abnormal route	33.03	2.00
Actual situation	34.33	2.43
Radar I	31.73	2.51
Radar II	33.65	2.27

The means and SEs of the interactions between route and information conditions for the perceived mental workload are given in *Table 6-20*.

Table 6-20 Perceived mental workload for interactions between route and information

Interaction effects for perceived mental workload			
Route	Information condition	Mean	SE
Normal	Actual	35.23	3.51
	Radar I	31.02	3.39
	Radar II	34.10	3.08
Abnormal	Actual	33.43	3.35
	Radar I	32.44	3.70
	Radar II	33.21	3.35

The statistical results of the main and interaction effects of route, information condition and route and information condition are given in *Table 6-21*.

Table 6-21 ANOVAs results on main and interaction effects for mental workload

ANOVAs main and interaction effects for mental workload			
	F-value	Df	P
Route	0.023	1	0.880
Information condition	0.297	2	0.743
Route x Information condition	0.113	2	0.894

ANOVAs were performed for route conditions and context information conditions. The two analyses did not show significant results, respectively route conditions ($p=0.880$) and context information conditions ($p=0.743$). ANOVAs for the interaction between route and context information did not show significant difference ($p=0.894$).

These results indicate that the train drivers perceived a higher mental workload during their first ride than during their second ride, which can be explained as a training effect or getting acquainted with the simulator and the route. The context information conditions and route conditions did not affect perceived mental workload.

Four safety-related concepts were measured. The measured SPADs and ATP interventions were not related to the route and context information conditions, but possibly to the lack of route knowledge in the simulation. Mental workload measures indicated only a training effect of using the simulator: difference between ride one and two. However, mental workload was also assessed by how long it took to react to the dead-man's handle. These secondary task performance measures indicated that the more context information was provided, the more the train driver was mentally loaded and the slower his response was to the dead-man's handle.

The reaction times indicate that train drivers had to make more effort when supported by context information. This difference in results between the two mental workload indicators might be caused by the already mentioned disadvantage of measuring mental workload, namely that participants are prone to forget certain parts of the task where variations in their mental workload might have occurred (Stanton et al., 2005).

Questionnaires

The completed questionnaires of all 48 train drivers were analysed. The contribution and status of the performance indicators were measured pre- and post-experiment. The train drivers had to indicate their contribution to each construct on a 5-point Likert scale. For the status of the 4 constructs, report marks between 0 and 10 were asked for (0= very bad, 10= excellent). Pre- and post-experiment questionnaires were compared to identify the effect of context information on the contribution and status of the four constructs.

Table 6-22 Results of the perceived contribution and status of the railway performance indicators

Railway performance indicators								
	Safety		Punctuality		Energy costs		Passenger satisfaction	
	Pre-test mean (SD)	Post-test mean (SD)	Pre-test mean (SD)	Post-test mean (SD)	Pre-test mean (SD)	Post-test mean (SD)	Pre-test mean (SD)	Post-test mean (SD)
Contri-bution (1-5)	4.00 (.55)	3.65 (.79)	3.71 (.54)	3.56 (.56)	3.52 (.77)	3.69 (.77)	3.79 (.58)	3.56 (.62)
Status (0-10)	7.29 (1.06)	7.55 (1.37)	6.95 (.68)	7.46 (1.19)	6.00 (.86)	7.67 (1.31)	6.53 (1.10)	7.60 (1.11)

Statistical analyses were performed with the Wilcoxon signed-ranks tests. The tests showed a significant difference for the train driver's contribution to safety (N=47, Z=-2.782, p=0.005). No significant results were found for the perceptions of contributions to the other performance indicators.

The status of the performance indicators was also tested pre- and post-experiment. The Wilcoxon signed-ranks test showed significant differences for the perceived status of punctuality (N=43, Z=-3.059, p=0.002), the perceived status of energy costs (N=44, Z=-4.807, p=0.000), passenger satisfaction in railway (N=47, Z=-4.943, p=0.000). The results for safety were not significant (p=0.341).

These results show that train drivers expected positive effects of using context information on the status of punctuality, energy costs, and passenger satisfaction. They also expected their personal contribution to safety to decrease.

Situation awareness

SA is measured in relation to the three information processing aspects: perception, comprehension and projection. No objective criteria can be defined for SA (Endsley, 1995a), so the results were compared between the context information and route conditions, to gain insights into the decision-making process to explain and understand the results of the task performance measures.

In total, 240 rides took place in which SA was measured. In 144 rides, the train drivers indicated that nothing special had happened. In the other 96 rides, train drivers indicated that more than 30 remarkable situations that took place were directly related to the simulator instead of remarkable situations related to their rides. Asking train drivers why they did not indicate that they had been waiting for several minutes before a red signal, they answered that such situations are not remarkable, because they experience them in practice every day. Moreover, this also relates to the disadvantage of post-task measurement; subjects are prone to forget certain aspects of the tasks where variations in their mental workload might have occurred (see *section on mental workload*).

Thus, in general the SA assessment instrument developed for the present research did not work properly: no valuable data was gathered, so no analyses were conducted.

Empowerment

The four cognitions of empowerment (Spreitzer, 1995) were pre-test measured on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). In the post-test questions, train drivers had to indicate whether context information would have a positive, a negative or no effect on their perception of decision-making. The specific answer possibilities were: much worse, worse, equal, better, much better. The post-test answers were coded as -2, -1, 0, 1 or 2. The sum of the pre-test figures and the coded post-test figures are listed as the post-test results. *Table 6-23* depicts the pre- and post-test results on the four cognitions of empowerment.

Table 6-23 Pre-test and post-test results of empowerment

Empowerment		
	Pre-test Mean (SD)	Post-test Mean (SD)
Meaning	4.38 (.87)	4.64 (1.03)
Competence	4.71 (.77)	5.00 (.95)
Self-determination	3.40 (1.13)	3.96 (1.40)
Impact	3.02 (1.23)	3.50 (1.37)

To identify the effect of the use of context information on empowerment, the results of the pre- and post-test were compared. The Friedman test led to the rejection of the null hypothesis that all cognitions were equally supported by the context information (N=47, χ=132.385, df=7, p=0.000).

Detailed analyses with the Wilcoxon signed-ranks test showed significant differences for all four cognitions: meaning (Z=-3.153, p=0.002), competence (Z=-3.500, p=0.000), self-determination (Z=-4.700, p=0.000) and impact (Z=-4.426, p=0.000).

The results indicate that train drivers perceived being more empowered to perform their job (meaning, competence, self-determination and impact) by the context information.

Communication

In the simulations, train drivers were asked to adhere to procedures as they normally do in practice. During the experiment, it turned out that communication between train driver and railway signaller did not depend on the experimental condition (context information and route condition), but on personal aspects of the train driver. Some train drivers were disciplined and adhered to the rules, that is, they contacted the railway signallers when they stopped for a red signal, while others did not contact the signaller at all, regardless of the experimental condition. Thus, analysing these communications would not have been a meaningful measure. Therefore, cooperation was analysed as the perceived quality of the communication during a ride. After each ride, both railway signallers and train drivers were asked to indicate this with a report figure from 1 to 10 (1= very bad, 10= very good). *Table 6-24* lists the results on the perceived quality of the communication for both train drivers and railway signallers.

Table 6-24 Results perceived quality of communication

	Train drivers Mean (SD)	Railway signallers Mean (SD)
Normal routes	7.85 (1.35)	7.79 (1.05)
Abnormal routes	7.79 (1.70)	6.73 (2.41)
Actual information	7.75 (1.46)	7.20 (2.35)
Radar I	7.88 (1.61)	6.86 (2.04)
Radar II	7.82 (1.56)	7.11 (2.31)

These results indicate that in general, train drivers perceive the quality of the communication as rather good (+/- 7.8 on a 10-point scale), with almost no differences between the conditions. Railway signallers perceived, for all conditions, lower quality of the communication. As only three signallers participated, statistical analyses were not useful. On abnormal routes, heavy communication can be expected. That might be a reason for the railway signallers' perceived difference in quality of communication between normal and abnormal routes.

In the post-experiment questionnaire, both train drivers and railway signallers also had to indicate on a 5-point Likert scale whether the context information would affect the amount of communication (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). A lot of communication distracts railway signallers from their main task: planning routes. Train drivers indicated that the amount of communication would not increase (mean=2.25, SD=0.86). Of the three railway signallers participating, two indicated that the amount of communication would not change, and one indicated that the amount of communication would increase.

During the simulations, many participating train drivers and railway signallers said that taking part in the simulations has improved the understanding of each other and their jobs. They indicated that this would have a positive effect on their future cooperation in practice.

Task support

The perceived effectiveness of task type the context information provided is measured with the PETT (Wilson et al., 2001). Post-experiment, the train drivers had to indicate their agreement with four questions on task support on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). *Table 6-25* depicts the results.

Table 6-25 Perceived Effectiveness of Task Type of the Context information support

Perceived Effectiveness of Task Type		
	Mean	SD
Planning	3.85	0.92
Group performance	3.79	1.09
Complex decision-making	3.40	0.82
Negotiation	3.19	0.89

The Friedman test led to the rejection of the null hypothesis that all group tasks were equally supported by the context information. Context information support affected group task support (N=48, χ=28.573, df=3, p=0.000).

Wilcoxon signed-ranks tests between the four types of tasks showed significant differences between complex decision-making and planning (Z=-3.475, p=0.001), negotiation and planning (Z=-3.423, p=0.001), complex decision-making and group performance (Z=-2.771, p=0.006), and negotiation and group performance (Z=-3.275, p=0.001).

These results indicate that train drivers believe that context information supports planning and group performance more than it supports negotiation and complex decision-making.

User acceptance

User acceptance was measured by an adapted version of the UTAUT questionnaire (Venkatesh et al., 2003). UTAUT measures the acceptance of new technology by identifying the intention to use the technology. The relevant constructs for the intention to use in this experiment were performance expectancy (four questions), effort expectancy (three questions) and social influence (three questions). The UTAUT questionnaire consisted of 10 statements with which the train drivers had to indicate their agreement on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). *Table 6-26* lists the results of the various constructs of UTAUT.

Table 6-26 Post-experiment results of UTAUT

Unified Theory of Acceptance and Use of Technology		
	Mean	SD
Performance expectancy	3.11	0.56
Effort expectancy	3.91	0.86
Social influence	2.61	0.64

The Friedman test led to the rejection of the null hypothesis that all constructs are equally important for the intention to use. Intentions to use differed in their importance (N=47, χ=49.124, df=2, p=0.000).

A Wilcoxon signed-ranks test showed significant differences between all constructs: performance expectancy and effort expectancy (Z=-4.938, p=0.000), and performance expectancy and social influence (Z=-3.680, p=0.000), and effort expectancy and social influence (Z=-5.652, p=0.000).

These results indicate that the behavioural intentions of train drivers to use the context information will be based firstly on the effort expectancies, secondly on the performance expectancies and thirdly on social influence.

Applicability of context information in practice

Two types of context information were tested: Radar I (context information about own train) and Radar II (context information about all surrounding trains). For both types of context information, the usefulness and applicability and the prediction of usage of the specific information items signals, punctuality and path prediction were measured by agreements on a 5-point Likert scale (1= totally disagree, 2= partially disagree, 3= neutral, 4= partially agree, 5= totally agree). *Table 6-27* provides an overview of the means of each information element, per radar condition.

Table 6-27 Perceptions of the usefulness and applicability and predicted use of three information elements of Radar I and Radar II.

Usefulness and applicability			Predicted use		
	Radar I My train Mean (SD)	Radar II All trains Mean (SD)		Radar I My train Mean (SD)	Radar II All trains Mean (SD)
Signals	3.87 (.75)	2.93 (.88)	Signals	4.00 (.96)	2.91 (1.06)
Punctuality	3.74 (.77)	2.83 (1.00)	Punctuality	3.64 (1.10)	2.72 (1.10)
Path-predictions	3.48 (1.10)	2.89 (.95)	Path-predictions	3.62 (1.01)	2.94 (1.05)

A Wilcoxon signed-ranks test was conducted to test the perceived differences between both radars. The usefulness and applicability of all information elements showed significant results between Radar I and Radar

II: signals (Z=4.871 p=0.000), punctuality (Z=4.871 p=0.000), and path predictions (Z=4.871 p=0.000). Also the predicted use all information elements between Radar I and Radar II showed significant results: signals (Z=4.795 p=0.000), punctuality (Z=4.785 p=0.000), and path predictions (Z=3.931 p=0.000).

These results show that train drivers perceived that the usefulness and applicability of signs, punctuality and path predictions were more relevant for their own train than for the surrounding trains. They expect to use the information about their own train more than the information about all trains.

Observations during dynamic experiments

The user interface of the train simulator provides an abstract view of the tracks and signs and signals. Other real-world elements (e.g. houses, farms, cars, bridges, animals and crossings) are not available. Questions can be posed about the ecological validity of the simulator. The main question is whether train drivers really perceived they were performing their job. Some observations during the experiment indicated that the ecological validity was sufficient:

- One laptop broke down during the experiments. The train driver had kicked out the power because he was reacting to the dead-man handle. In practice, the response to the dead-man's handle is to push a button with a foot. The train driver also responded with his feet to the dead-man's sound, and kicked out the power supply.
- At a certain moment, one train driver caused a SPAD: he passed a red signal. The train driver was so upset about this, that he was no longer able to drive properly; he needed a cup of coffee and a cigarette to recover.

6.3.4 Summary of results

The aim of the dynamic simulation experiment was to test whether context information affects train driving decision-making (task performance) and whether it improves communication between train drivers and railway signallers. A third objective was to identify the train drivers' perceptions of using context information in practice. In a multi-user train simulator, a 3x2 within subjects design was applied of three context information conditions and two route conditions.

Context information improved only one performance indicator: job satisfaction. Train drivers indicated that the context information empowered them. Thus, in general, the context information did not affect train driving decision-making. This means that railway performances will not be improved. These results allowed hypothesis 1 to be rejected. Safety related aspects like SPADs, ATP interventions, mental workload and how long it took to react to the dead-man's handle were hardly influenced by the context information; it took longer to react to the dead-man's handle when more context information was provided. Overall, this is a positive finding, as safety is the most important construct for innovation research in railway. Innovations that influence safety in a negative way are unacceptable.

Moreover, train drivers have a positive attitude towards using the context information. Train drivers expected that the context information would positively affect punctuality, energy costs and passenger satisfaction. However, they expected their own contribution to safety to decrease. They might have the idea that the inside context information might distract their attention from the relevant outside information (signs and signals). Although signals are also displayed on the context information display, they are not allowed to trust the on-board information.

The perceptions of the quality of the communication between railway signallers and train drivers did not change on the context information. These results led to the rejection of hypothesis 2. However, railway signallers perceived the quality in general as lower than the train drivers did. Train drivers and railway signallers expected no effect of context information on the amount of communication. But overall, they said that meeting each other and physically working together has improved their understanding, which might affect their future cooperation.

Overall, the train drivers have a positive attitude towards the user acceptance of the context information. Effort expectancies are the highest, training and introduction should focus on performance expectancies and social influence. In general, they found the information about their own train more applicability to practice than information about all trains.

6.4 Summary

The overall research objective of these experiments was to establish whether context information improves railway performances. The railway performances identified are safety, punctuality, energy costs, passenger satisfaction and job satisfaction. In this section, first the results of both experiments were related to these performance indicators, and second, the detailed hypotheses for both experiments were discussed. In the static simulation experiment, safety, punctuality and energy costs were not measured directly. Only the change in driving decisions could indicate that the train drivers intended to change these aspects. Also perceptions of these performance indicators were measured.

Performance indicators

Safety is the most important construct for innovation research in railway. In an objective way, safety was only measured in the dynamic simulation experiment. Safety was operationalised in four different ways (i.e. SPADs, ATP interventions, mental workload and how long it took to react to the dead-man's handle). Most results indicated that the context information did not affect safety. Only the reaction time to the dead-man's handle increased under information conditions, indicating more mental effort for the train driver. In the static experiment, train drivers perceived the overall safety to increase, while in the dynamic simulation they expected no effect on safety. In both experiments, the train drivers perceived that their personal contribution to safety would decrease; train drivers expected to be distracted by the context information in their cabin.

Punctuality was directly measured only in the dynamic simulation experiments. Context information did not affect punctuality. In both experiments, train drivers perceived that punctuality would increase because of the context information, while their personal contribution to punctuality would remain the same.

Energy costs were also measured directly in the dynamic simulations. These results indicated no effect of context information on energy costs. In both experiments, train drivers perceived energy costs to increase when providing context information. Only in the static simulation did train drivers expect their personal contribution to energy costs to increase.

Job satisfaction was operationalised as empowerment, and in both experiments the perceptions were positive: almost all cognitive constructs increased in both experiments. This means that providing context information would increase the intrinsic task motivation, which might serve as a stimulus to use the information in practice.

Passenger satisfaction relates to punctuality and information provision. The objective results of punctuality did not change in the dynamic simulations. The perceptions of passenger satisfaction were positive in both experiments, although in both experiments train drivers also indicated that their personal contribution to passenger satisfaction will not change. This is remarkable, as context information, and especially future context information, provides train drivers with the opportunity to inform passengers about what is taking place on the track and to give an indication of the possible delay.

Hypotheses testing

The overall research question (Does context information increase railway performances?) was refined in three hypotheses:

Hypothesis 1 *Context information improves train driving decision-making.*

Hypothesis 2 *Context information improves communication between train drivers and railway signaller.*

Hypothesis 3 *Dynamic information and time pressure do not affect train driving decision-making.*

These hypotheses were tested in three methodological ways: static simulation experiment, dynamic simulation experiment, and a comparison of the two. Three kinds of data were collected: train driving decisions, log data and questionnaire data.

Comparing the results of the two experiment provided indications for hypothesis 3. The kinds of train driving decisions taken in the experiments were different, namely indications of train driving decisions, and train driving actions taken (with feedback). Therefore, these results are not really comparable. What can be said is that train driving decisions changed upon the context information in the static experiment, while the train driving actions in the dynamic experiment did not show much effect on the performance indicators. Thus hypothesis 3 was rejected: dynamic information and time pressure do affect train driving decision-making.

Hypothesis 1 was supported by the static experiment but not by the dynamic experiment. The differences between the two experiments were not only the dynamic information and time pressure, but also the kind of context information provided, and the methodological set-up: the numbers and kinds of traffic situations provided, and the experimental set-up (fixed order and randomised order).

Hypothesis 2 was tested only in the dynamic simulation. The measured communication results led to the rejection of hypothesis 2: context information did not improve the communication between railway signallers and train drivers.

A general comparison of both experiments shows that the objective data (train driving decision-making and train driving actions) were not comparable. The subjective data, however, showed an overlap. This might indicate that cost-effective methods (e.g. static simulation experiments) are beneficial in this kind and phase of innovation research for identifying the perceptions of potential end-users. The perceptions of potential end-users are relevant: if they do not accept or find it useful, the innovation has the potential to fail, no matter what the performance expectancies are. On the other hand, if a concept is perceived very positively, but the performance expectancies are low, investment should be carefully considered.

Thus, in general the experiments do not fully prove that context information improves railway performances. Train drivers expect that the main railway performances would improve, while safety would not be affected. The main reasons for this might be that it is easy to use, they expect that it will not take much effort to use the context information, and they are intrinsically motivated to use it, which might serve as a stimulus.

Chapter 7

Overview and reflections

This, the final, chapter starts with an overview of the results of the user experiments in home care and railway. The two experiments per case are then compared. It is concluded that low-cost experiments are useful in the first phases of IT innovation research. The effect of extra information on proficient decision-making is considered by focusing on the recognition primed decision model, the decision strategy applied, and knowledge encapsulation. This is followed by a discussion of the various aspects of the applied user-centred design approach and the user experiments. The implications of the results for practice are given. For home care, one should create dedicated wound care teams, continuously supported by a dermatologist, and provide training in how to use the new technologies. For railway, the decision latitude should be increased and training on the feed-forward driving style should be provided.

7.1 Cross case analyses

The overall research objective was to identify the effect of information and communication support on the task performance of distributed professionals. The research was applied to two cases: home care and railway. For each case, a domain analysis was performed and a mobile application was designed. During these user-centred design phases also a task performance model was developed and the main case-specific performance indicators were identified. Two user experiments were performed per case to answer the research questions:

1. What is the added value of information and communication support for distributed professionals on the main task performance indicators?
2. What is the user acceptance and perceived applicability of information and communication support for distributed professionals?

7.1.1 First experiments: survey and static simulation

In home care, technology use is rather rare. Therefore, the concept of information support was first tested in a paper survey, without using technology. In railway, dynamic information and time pressure put heavy cognitive demands on train driving decision-making in real-life settings. Therefore, the concept of information support was first tested in a less cognitive demanding situation, namely a static simulation. The main results of these first, simplified experiments in the two domains are given in *Table 7-1*. The results are compared only on the level of the constructs.

Table 7-1 Overview of the results of the first experiments: survey for home care and static simulations for railway

Construct	Home care survey	Railway static simulation
Decision-making	History information did not affect treatment decision-making.	Context information changed driving decisions.
Perceptions of performance indicators	History information had positive effects on perceptions of the status of quality and efficiency of care and on patient friendliness. History information had positive effects on perceptions of their personal contribution to efficiency of care, and a negative effect on their personal contribution to patient friendliness.	Context information had positive effects on perceptions of safety, punctuality, energy costs and passenger satisfaction. Context information had a positive effect on perceptions of their personal contribution to energy costs, and a negative effect on their personal contribution to safety.
Empowerment	History information had a positive effect on the district nurses' perceptions of the meaning, competence, self-determination and impact of their work.	Context information had a positive effect on the train drivers' perceptions of the meaning, their competences, and their self-determination in their work.
Task support	History information is perceived to support complex decision-making and negotiation.	Context information is perceived to support planning, group performance and complex decision-making.
Information preferences	Perceived preferences for most enriched history information condition: report and pictures.	The perceived usefulness of the context information differed between the normal and abnormal scenarios in speed and order of trains.

The main difference found between the two simplified user experiments relates to decision-making: history information in home care does not affect treatment decision-making, while context information in railway affects train driving decision-making. The professionals' perceptions of the effect of the information support on the performance indicators, on empowerment, on task support and information preferences differ slightly between both experiments.

Relating these results to the research objectives shows that for home care the answer to the first research question is twofold: there is no objectively measured added value of history information on task performances, while the subjectively measured added value is in general positive: district nurses perceive that history information will increase their task performances.

When relating the results to the research objectives for railway, the answer to the first research question is positive: both objectively measured and subjectively measured, the added value of context information on task performances is positive, assuming that changes in driving decisions result in better driving decisions.

The other results of the first two experiments show that the professionals feel empowered by the extra information. In both cases, the extra information supported complex decision-making, while for home care also negotiation was supported, and for railway planning and group performances were supported.

7.1.2 Second experiments: WoundLog and dynamic simulation

For home care, the second experiment was conducted immediately after the first experiment and with the same group of district nurses. In the second experiment, treatment decision-making was performed using technology: wound logbook, mobile phones and WoundLog. For railway, the results of the static simulation provided input for designing the train simulation experiment. In the second experiment, train driving was performed in a PC-based multi-user train simulator supported by two different radar configurations. *Table 7-2* lists the results of the second user experiments. Again, the results between the two cases are only compared on construct level.

Table 7-2 Overview of the results of the second experiments: WoundLog for home care and Train simulator for railway

Construct	Home care WoundLog	Railway Train simulator
Decision-making	History information did not affect treatment decision-making. The combination of history information and communication support did not affect treatment decision-making.	Safety: No direct effects; only increased mental effort for enriched information conditions. Punctuality: Context information did not affect punctuality. Energy costs: Context information did not affect energy costs.
Perceptions of performance indicators	History information and communication support led to positive expectations on the status of the quality and efficiency of care and patient friendliness. History information and communication support led to positive expectations on their personal contribution to quality and efficiency of care.	Context information support led to positive expectations on punctuality, energy costs, and passenger satisfaction. Context information support led to negative expectations on their personal contribution to safety.

Construct	Home care WoundLog	Railway Train simulator
Empowerment	History information had a positive effect on the district nurses' perceptions of the meaning, competence, self-determination and impact of their work.	Context information had a positive effect on the train drivers' perceptions of the meaning, competence, self-determination and impact of their work.
Mental workload	Information and communication support had no effect on mental workload. The initiator of a conference experienced more mental workload than the receiver in conditions with no visual information exchange.	Context information support had no effect on mental workload. Train drivers perceived a higher mental workload during their first ride than during their second ride.
Situation awareness	Assessment method was considered to be inappropriate.	Assessment method was considered to be inappropriate.
Task support	History information is perceived to support complex decision-making and negotiation.	Context information is perceived to support planning and group performance.
Information preferences	Perceived preferences for most enriched history information: report and pictures.	Context information about their own train is perceived to be more relevant than that about all trains.
User acceptance	The behavioural intentions of using the history information and communication support were mostly based on effort and performance expectancy and to a lesser extend on social influence.	The behavioural intentions of using the context information support were mostly based on effort expectations, to a lesser extend on performance expectations and least on social influence.
Applicability	History information and communication support were perceived to positively influence the main performance indicators. Negative aspects are associated with patient acceptance and technical issues (camera handling and taking care of the device).	Context information about their own train is perceived to be more relevant than that about all trains; train drivers also expect to use this information more.

The main results of the second experiments indicate that the information and communication support provided did not affect decision-making. The general trends of the professionals' perceptions of the information and communication support are that in general they have positive expectations for their task performances, it empowers them and it hardly affects mental workload. The perceptions of the kinds of tasks supported by the extra information are completely opposite for both cases in relation to the circumplex model of McGrath (1983); complex decision-making and negotiation for home care and planning and group performance for railway.

Relating these results to the research objectives shows comparable answers for both cases. The answer to the first research question is twofold:

there is no objectively measured added value of information and communication support on task performances, while the subjectively measured added value is in general positive. The answer to the second question is also comparable between the two cases: user acceptance and applicability are in general perceived positively, and the outcomes provided input for case-specific improvement.

In the following sections, the results of the two case-specific experiments are compared. This provides insights into the experimental methods applied.

7.1.3 User experiments in home care

For home care, the reason to perform two experiments was to identify whether technology use was a factor in treatment decision-making, because technology is not commonly used in home care. *Table 7-3* lists the results of the survey and the WoundLog user experiment in home care.

Table 7-3 Comparison of survey and WoundLog for home care

Constructs	Survey	WoundLog
Decision-making	History information did not affect treatment decision-making.	History information did not affect treatment decision-making. The combination of history information and communication support did not affect treatment decision-making.
Perceptions of performance indicators	History information had positive effects on perceptions of the status of quality and efficiency of care and on patient friendliness. History information had positive effects on perceptions of their personal contribution to efficiency of care, and a negative effect on their personal contribution to patient friendliness.	History information and communication support led to positive expectations on the status of the quality and efficiency of care and patient friendliness. History information and communication support led to positive expectations on their personal contribution to quality and efficiency of care.
Empowerment	History information had a positive effect on the district nurses' perceptions of the meaning, competence, self-determination and impact of their work.	History information had a positive effect on the district nurses' perceptions of the meaning, competence, self-determination and impact of their work.
Task support	History information is perceived to support complex decision-making and negotiation.	History information is perceived to support complex decision-making and negotiation.
Information preferences	Perceived preferences for most enriched history information condition: report and pictures.	Perceived preferences for most enriched history information: report and pictures.

The same district nurses participated in both experiments. The results in this section are based on a comparison of the data. However, the district nurses were not explicitly asked about their perceived differences between the survey and the WoundLog experiment.

The results show that, in general, the survey provided the same kinds of results on using the concept of history information as the technology-mediated user experiment. Moreover, the statistical tests showed that the results for history information support for treatment decision-making with and without technology were comparable (see *Section 5.3.2*). Testing the history information concept with and without technology had no effect on the district nurses' perceptions of the main performance indicators. As the same district nurses participated in the survey and the experiment, a Wilcoxon signed-ranks test could be performed between the post-survey and the post-experiment results on these perceptions. The analyses did not show significant differences ($p \geq 0.157$). These results are in line with the rejection of hypothesis 3: technology is not an interfering factor for treatment decision-making by district nurses.

Technology-mediated task performance affected the district nurses' perceptions of the support the technology provides. District nurses expect less support for planning under technology conditions ($Z=-2.50$, $p=0.012$). Their self-determination decreases in technology-mediated situations ($Z=-2.814$, $p=0.005$).

In general, the comparison of the survey and the user experiment showed that performance-related indicators (e.g. treatment decisions, quality and efficiency of care and patient-friendliness) were not affected by technology use. While aspects related to perceptions (e.g. empowerment and task support) were negatively affected by technology use. This might indicate that when using a survey, district nurses perceive only the positive effects of the information concept, while when using the technology, they are also confronted with its disadvantages. An alternative explanation is that three hours of treatment decision-making exhausted the district nurses, which resulted in a less positive attitude.

Thus, a low-cost survey seems to be an appropriate instrument to test the added value of the concept of history information on treatment decision-making in the first phase of IT innovation research. The recommendation is to use this type of method more often in the first phases of IT innovation research.

7.1.4 User experiments in railway

Dynamic information and time pressure put strong cognitive demands on train driving. Therefore, the concept of information support was first tested in a less cognitively demanding situation, namely a static simulation. Second, adapted versions of the context information were tested in the train simulator. *Table 7-4* lists the results of the two simulations.

Table 7-4 Comparison of the results of static and dynamic simulation in railway

Construct	Static simulation	Dynamic simulation
Decision-making	Context information changed driving decisions.	Safety: No direct effects; only increased mental effort for enriched information conditions. Punctuality: Context information did not affect punctuality. Energy costs: Context information did not affect energy costs.
Perceptions of performance indicators	Context information had positive effects on perceptions of safety, punctuality, energy costs and passenger satisfaction. Context information had a positive effect on perceptions of their personal contribution to energy costs, and a negative effect on their personal contribution to safety.	Context information support led to positive expectations on punctuality, energy costs, and passenger satisfaction. Context information support led to negative expectations on their personal contribution to safety.
Empowerment	Context information had a positive effect on the train drivers' perceptions of the meaning, their competences, and their self-determination in their work.	Context information had a positive effect on the train drivers' perceptions of the meaning, competence, self-determination and impact of their work.
Task support	Context information is perceived to support planning, group performance and complex decision-making.	Context information is perceived to support planning and group performance.
Information preferences	The perceived usefulness of the context information differed between the normal and abnormal scenarios in speed and order of trains.	Context information about their own train is perceived to be more relevant than that about all trains.

The two experiments were conducted with different train drivers at different moments. Moreover, the results of the static simulation were input for the dynamic simulation. The main difference between the experiments was that the dynamic information and time pressure were absent from the static simulation. Roughly comparing the results should give an indication of the use of dynamic information and time pressure on train driving decision-making.

The main difference found between the two experiments relates to decision-making: in the static simulation experiment, train drivers changed their driving decisions under the enriched information conditions, while in the dynamic simulation experiment train drivers' driving actions did not affect railway performance indicators. This result may be attributed to the way decision-making was measured, in addition to the influence of the absence of dynamic information and time pressure.

In the static simulation experiment, train drivers said what kinds of decisions they would take in such situations, while in the dynamic simulation experiment, the results of the actions taken were measured. The railway task performance model describes the relationship between the decisions and actions. Decisions are the result of the perception, comprehension and projection phase, and actions are the results of the decisions. Thus, the difference in the decision-making results between the two experiments might partly be caused by the different measurements: decisions versus actions.

Dynamic information and time pressure are train driving aspects that are typical of expert task performance in naturalistic settings (Zsambok et al., 1997). Omitting those kinds of aspects might result in a different task, or at least different decision-making strategies can be applied. Paquette and Kida state (1988) that with low task complexity (static simulation), full processing decision-making strategies can be used, while with increased task complexity (dynamic simulation) reduced processing strategies are used, to reduce cognitive effort. As the results differ between the static and the dynamic simulation, dynamic information and time pressure might have a great effect on railway decision-making. The results of the static simulation suggest that context information affects train driving decision-making. Combining that with the results of the dynamic simulation, it might be an indication that the typical aspects of real-life settings (dynamic information and time pressure) obstruct a smooth and immediate adaptation of the context information in their driving routines. Possible reasons for this are described in the following section.

When comparing the perception-based results of the two experiments, the same trend as in home care can be identified. In a small, low-cost experiment, in which the information concept is only visualised, perceptions are more positive than after a few hours of using the information concept in a more realistic task environment.

Explanations for the overall results for both cases are provided in the following sections on theoretical, practical and methodological considerations and implications, resulting in a revisited task performance model.

7.2 Considerations on theory

The decision-making results of three of the four experiments show that task performances are not improved by information and communication support. Only the decision-making results of the static simulation experiment in railway show that information support changes decision-making. Presumably, using information and communication support in working routines requires more lengthy training. Two anecdotes from the user experiments are described to illustrate this. The first anecdote took place during a mobile phone conference in the home care experiment. This anecdote emphasises that not all relevant information is used properly during the experiments.

Anecdote from home care experiment **Setting: mobile phone conference**
Two district nurses are discussing a patient. The receiver of the conference call provides firm advice to the initiator of the conference. The receiving district nurse answers all questions promptly and states explicitly her opinion on what to do with this patient (=treatment decision). After hanging up, the district nurse looks at the leader of the experiment and asks (herself) 'Was the patient a diabetic?'

The second anecdote took place in the dynamic train simulation experiment. A train driver (who participated in some workshops on the design of the radar) was waiting for a red signal. He started to comment to the experiment leader on what was going on. This anecdote emphasises the difference between stating what to do (= decision-making) and applying this to practice (= acting). It shows that although the train driver knew what was going on in front of him, he just applied his usual driver routine when faced by a green signal.

Anecdote from railway experiment **Setting: at the start of a radar supported ride**
'The radar looks nice. Look, now I can see what is happening. I am waiting for a red signal, because a freight train is in front of me. If my signal improves, and I wait for a green one, all I have to do is to adjust my speed to the train in front of me'. After a short while, his signal improves, and even turns green. The train driver applies full power and has to brake strongly for the next signal.

When relating the results of the experiments to the conceptual framework of task performance, three interrelated reasons can be found for the lack of effect of the provided information and communication support. These three reasons concern the recognition primed decision-making (RPD) model, the applied decision-making strategy and knowledge encapsulation. Each of these is described below.

7.2.1 Recognition primed decision-making

Naturalistic decision-making (NDM) often applies the recognition primed decision-making (RPD) model (Klein, 1999). Experts rely on prior experiences and apply the principle of first recognising a situation, and then selecting the appropriate course of action they applied before. The first feature of the RPD model is recognising cases as typical (Klein, 1989). Assuming that proficient decision makers work according to the RPD model, the extra information and communication support create new situations, which cannot be directly recognised as typical. Such situations cannot be recognised as they have never had this kind of information before. If professionals do not assess such situations as new, but recognise them as typical or experienced before, an action is selected that matches an old situation. In this way, the extra information and communication support is not used. Likely, new information and communication support requires new decision-making, which needs more time and effort to develop than available in the applied user experiments.

7.2.2 Decision-making strategy

A decision-making strategy provides an understanding of the different ways activities can be accomplished (Vicente, 1999). The applied decision-making strategy determines the flow of information (Vicente, 1999). Moreover, the task performance model describes that the way the information is perceived, comprehended and projected depends on the strategy applied. It is possible that the decision-making strategies applied in home care and railway do not take information into account that relates to the kinds of extra information provided.

For railway, two driving strategies were identified, namely the feedback and the feed-forward driving style. When a feedback driving style is applied, the train driver responds to ATP signals and does not focus on the comprehension and integration of the various information sources. Thus, in a feedback driving style, context information will not be part of the information flow and thus will not be used, and therefore will show no effect on train driving performances. The question is whether train drivers actually apply the feedback style, as they were trained to use the feed-forward style. In practice, their decision latitude has been changed over the years. The density of the railway system resulted in a convergence of maximum track speed and timetable speed. This means that their decision latitude has decreased; there is hardly any possibility to make up a delay. Using their skills, train drivers looked for other opportunities. By not anticipating signals but just reacting to ATP indications, they gain some extra time. In this way, they act as responsible drivers who want to drive as punctually as possible; they have adopted a feedback driving style.

For home care, two typical treatment decision-making strategies can be applied (analytical and intuitive), as can some mixed forms. In analytical decision-making, decision-making proceeds in accordance with a certain systematic process, and the decision can be reached by analysis of the situation (Lauri et al., 2002). While intuitive decision-making relates to opposite characteristics (e.g. rapid information processing, simultaneous cue use, pattern recognition), intuitive decision-making relates to human problem-solving that relies on the prior knowledge an individual has gained about the issues and areas concerned (Lauri et al., 2002). This way of describing intuitive decision-making resembles the description of recognition primed decision-making (Klein, 1999). The kinds of tasks at hand determine the kinds of processes used in decision-making (Hammond, 1996). It is likely that the district nurses in the experiment used the intuitive strategy, which relates to using prior knowledge and means that new information is not taken into account (cf. RPD).

7.2.3 Knowledge encapsulation

The considerations on RPD and decision-making strategies showed that new information and communication concepts cannot be directly incorporated into existing work routines. The question is posed what kinds of cognitive mechanisms underlie this incorporation. Schmidt and Boshuizen (1993) state that in expertise development in medicine, knowledge is transferred through experiences with real cases into narrative structures. The cognitive mechanisms responsible for this transition are encapsulation and tuning. Knowledge encapsulation refers to the encapsulation of elaborated knowledge into high-level but simplified causal models or even diagnostic categories and tuning through the inclusion of contextual information (Schmidt & Boshuizen, 1993).

For the present research, the concept of knowledge encapsulation was interpreted as district nurses who have to learn to encapsulate the history information in their internal model of wound care knowledge. Their wound care knowledge comprises intuitive, technical and evidence-based knowledge (see *Table 3-2*). Training in how to use the history information properly should result in new causal models.

In relation to the results of the experiments, history information does not affect treatment decision-making. District nurses make decision in the way they are used to making them, based on their existing internal factors of wound care.

It is likely that a similar process takes place in the railway sector. Expertise development in railway is rather unknown. In train driving education, driving and train handling procedures are taught theoretically, and a lot of emphasis is placed on on-the-job training, namely acquiring experience-based knowledge. Experienced-based training is probably also useful for learning how to use the context information and to incorporate that information into their driving routines (knowledge encapsulation).

Thus, knowledge encapsulation refers to encapsulating new information in existing knowledge structures, which results in enhanced causal models. Knowledge encapsulation takes place in the internal factors of the task performance model.

7.3 Implications for practice

In this section, the results in home care and railway are translated into practical advice per domain.

7.3.1 Home care

The district nurses who participated in the experiments all worked for the same home care organisation, and delivered wound care on an individual basis. Compared to other home care organisations, a very large number (200) of district nurses were qualified to deliver wound care. As a result, they did not deliver wound care very often. Other organisations have teams of district nurses dedicated to delivering wound care. These teams are continuously educated by a dermatologist. Moreover, introducing WoundLog in practice cannot be accomplished by equipping 200 district nurses with laptops and cameras.

Before testing WoundLog in a test pilot, emphasis should be placed on technical aspects and patient acceptance. The camera should be replaced by a wireless, easy-to-use camera. The laptop should be light-weight and easy to carry. The patients should be well informed about the impact of WoundLog: fewer exhausting visits to the hospital and a chronological overview of their personal wound healing process for motivation.

The recommendation based on the present research is to continuously train district nurses to deliver wound care by creating dedicated teams supervised by dermatologists. These district nurses should be trained in using an improved version of WoundLog for delivering wound care. Before introducing it into practice, it should be tested in a pilot experiment.

7.3.2 Railway

Over the years, the density of the Dutch railway system has resulted in a convergence of maximum track speed and timetable speed. Nowadays, timetables are developed in such a way that hardly any decision latitude is left for train drivers: timetable speed is often equal to the maximum track speed. This means that their decision latitude has decreased; there is hardly any possibility to make up for a delay or to focus on energy costs. Increasing the decision latitude (by making the timetable speed lower than the maximum track speed) provides opportunities for train drivers to employ their skills and to increase railway performances. Moreover, in relation to the demand-control model of Karasek (1979), the combination of low decision latitude and high job demands is associated with mental strain and job dissatisfaction.

Context information seems beneficial for the feed-forward driving style. As their decision latitude has decreased over the years, train drivers sought other opportunities to employ their skills. When they did not anticipate signals but just reacted to ATP indications, they gained some extra time. In this way, to act as responsible drivers who want to drive as punctually as possible, they adopted a feedback driving style. To use the context information properly, the feed-forward driving style (as train drivers are used to apply) needs to be readopted. Train drivers feel empowered by the context information, which might serve as a stimulus to use it in practice.

In the 1990s, European legislation required the Nederlandse Spoorwegen (NS) to split up into two separate companies, one responsible for the rail infrastructure (ProRail) and one responsible for the transportation (NS). People who used to work for NS on the infrastructure became employees of ProRail. As a result, employees of both companies complain that communication has become complicated: too many people need to communicate to get something done (see also *Section 2.6.1*). During the experiments, train drivers and railway signallers indicated that physically working together improved their understanding of each other and of each others work, and that this would affect positively their cooperation in practice.

The recommendation resulting from the present research is to increase the decision latitude of train drivers, as this can improve their personal contribution to railway performances and their job satisfaction. This decision latitude should be supported by context information, and train drivers should be taught to adopt the feed-forward driving style and to use the context information properly. Before introducing this in practice, tests should be performed in which route knowledge is incorporated. Besides, train drivers and railway signallers should be trained together (e.g. by using the train simulator) to increase their understanding of each other and of each other's job.

7.4 Considerations on methodology

The present research consisted of two phases: user-centred design and user experiments. Each of these phases was applied to the two domains; home care and railway.

7.4.1 User-centred design

The goals of the user-centred design phase were to identify (1) the specific performance indicators, (2) the design of the dedicated mobile applications, and (3) the cognitive constructs for task performance (see *Section 2.1*).

The types of performance indicators that were identified in home care and railway differed a lot. For railway, five distinct performance indicators were identified with a small overlap (punctuality is also part of passenger satisfaction). While for home care all four identified indicators were related to treatment decision-making. This observation is in line with the design of the information support: train drivers found various types of information useful, while district nurses clearly identified the types of useful information.

The professionals had positive perceptions of the user acceptance and applicability of the two mobile applications. This means that involving potential end-users in all phases of the design of the information and communication support resulted in applications that were well accepted. However, the results revealed that decision-making was not improved by the information and communication support. This might stem from problems with recognition primed decision-making, decision-making strategy and knowledge encapsulation. What is the relation between these theoretical considerations and the applied user-centred design?

In the user-centred design phases, interviews and workshops were conducted with potential end-users. In such calm and quiet settings, the professionals had enough time to reflect on their activities. This resulted in design solutions that were appropriate to situations in which decisions were made in time-consuming ways, in which all information capacity could be used for decision-making. Moreover, even the observations at the workplace took place in quiet circumstances. For railway, the interviewers got the safety instructions that they were not allowed to observe train drivers and railway signallers or to ask them questions in busy or stressful situations. For home care, the frequency of semi-urgent and urgent wound care patients was low, so only normal visits were observed. It is likely that in such situations, full cognitive effort could be allocated to information processing and decision-making. This means that the design of the information and communication support matches 'ideal' situations in which there is enough time to use all cognitive effort for decision-making.

There are two ways in which mobile applications can be improved to match cognitively demanding real-life situations. The first is described in the implications for practice section. For railway, develop new timetables that increase the decision latitude of train drivers and train them to apply the context information in their feed-forward driving style. For home care, create dedicated wound care teams who are continuously educated by a dermatologist and train them to use the WoundLog application in their daily practice.

The second way is to redesign the applications by first identifying real-life proficient decision-making. Besides the question whether this identification is possible (prescriptive models of naturalistic decision-making are lacking), the difficulty is that the types of design become very tailor-made (home care: decision-making strategy is task dependent) and that even undesirable behaviour (railway: feedback driving style) is supported this way.

7.4.2 User experiments

Performing user experiments that focus on practical situations put constraints on the power of the experiment. Power is the probability of correctly rejecting a false null hypothesis. If the power of an experiment is low, then there is a good chance that the experiment will be inconclusive. Power depends on the number of factors in an experimental design and the sample size. As the user experiments of the present research involved practical situations in which professionals normally work, a large number of factors were involved (cf. ecological validity). At the same time, only a small number of professionals took part in the experiments. Therefore, the power of the experiments was low.

In all user experiments, real-life tasks of district nurses and train drivers were simulated. Of course, the experimental tasks did not exactly match real-life tasks. Mismatches might influence experimental task performance.

For home care, authentic patient material was collected in picture format. Authentic patient material is useful for user experiments, as it is the only way to keep cases identical (cf. (Boshuizen, Schmidt, Custers, & Van de Wiel, 1995). However, authentic patient material does not imply that authentic wound care tasks are performed. District nurses had to make treatment decisions based on these 'paper' patients. First, district nurses are not used to making treatment decisions from paper: a two-dimensional representation of a person's wound provides other information than a real person. Moreover, colour is essential in wound care. Although the best attempts were made, the colours on the prints of the patients' wounds might differ from the colours in practice. Also in relation to real-life situations, sensory perceptions (smell and feel), which help to assess a patient's condition, were missing.

In railway, the main difference from real-life settings in the simulator was probably the absence of route knowledge. Such knowledge is relevant to train drivers in determining and using their decision latitude. Despite its lack of route knowledge, the train simulator can be seen as an appropriate method to test innovations in railway in a cost-effect and safe way.

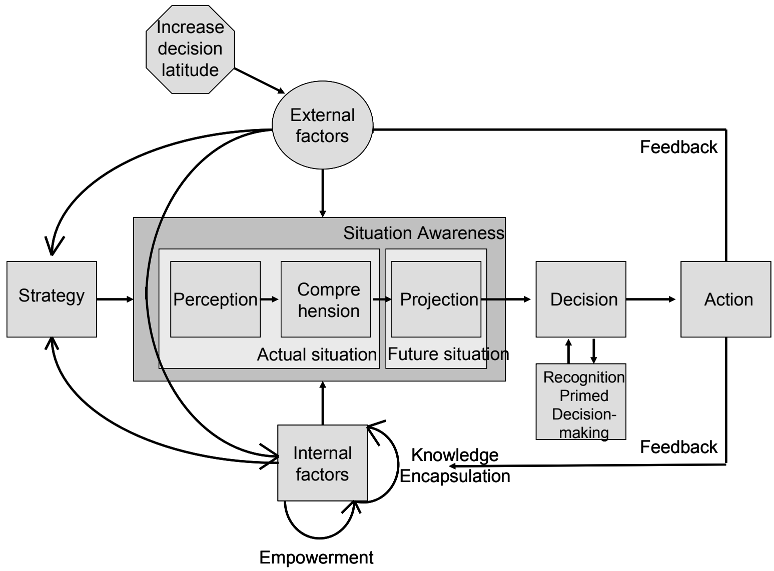
7.4.3 Research in practice

The present research focused on professionals. Studying proficient decision-making requires the participation of professionals in the experiments. It is not only difficult to get professionals to participate, but they tend to have only a limited amount of time available (and thus are expensive). Thus, the number of tasks performed is rather low in comparison with for example user experiments with student subjects (cf. external validity). Moreover, the overall research objective of this study also puts authenticity demands on the types of tasks that need to be performed and the environments in which they need to be performed (cf. ecological validity). Complex real-life situations need to be simulated. Besides, understanding user behaviour in complex real-life situations means that many, often interrelated constructs are involved that need to be identified, measured and interpreted (cf. construct validity). This requires creative and innovative data collection that goes beyond the one-to-one relation between construct and assessment. So in general, performing user experiments related to real-life situations is complex and time-consuming, but provides insights into real-life behaviour.

7.5 Towards a revisited model of task performance

Three theoretical explanations are presented to explain the results of the experiments. All theoretical explanations refer to training in how to use the new information, especially in relation to existing cognitive patterns or structures. Recognition primed decision-making and knowledge encapsulation are relevant theoretical constructs for both cases. The applied driving strategy was especially important for railway. Moreover, specific to railway is the proposal to increase the drivers' decision latitude, in order to make the feed-forward driving style useful. As a result of the theoretical considerations and the implications for practice, the general task performance model is revisited and the three theoretical aspects are incorporated into it.

Figure 7-9 Revisited task performance model



The increased decision latitude changes the external factors, which in turn affect the internal factors and the strategy applied. Recognition primed decision-making (RPD) is not part of decisions, but interacts with decisions. RPD relates to the toolbox of experiences the professional can use for proficient decision-making. RPD is not part of the decision, as situations of low cognitive effort can take place in which time and effort-consuming information processing strategies can be applied in which new situation can be trained or experienced, which fill the toolbox of RPD. Knowledge encapsulation is applied to provide training in how to use new information and, more precisely, it changes the existing knowledge structure. Knowledge encapsulation refers to a more dynamic process than the way the internal factors are presented at this moment. Moreover, as knowledge encapsulation is a process that is fed by training or experiences, the feedback of the actions provides input into knowledge encapsulation.

7.6 In conclusion

The present research studied the effect of information and communication support on the task performance of distributed professionals. Often, technological developments alone are enough reason to introduce new applications in practice. However, to make applications useful, profitable and cost-effective, and to use them to their full potential, the design should be based on the cognitive and affective processes of the end-users, especially if the end-users are professionals. Professionals exploit different ways of working (or strategies) which require different support. This study also shows the relevance of cost-effective methods in IT innovation research: early identification of performance expectancies and suggestions to improve the design for applicability in practice. User acceptance and perceived applicability are valuable indications of the success of a new application, as they can work as a stimulus to use it in practice. However, effects on task performances are also valuable indicators for the match with cognitive structures and productivity improvements. The home care sector is recommended to create dedicated teams for wound care that are continuously educated by dermatologists. The rail sector is recommended to increase the decision latitude of train drivers, and to provide training and support for the feed-forward driving style.

Appendix A

Assessment instruments

Home care: Casus beschrijving

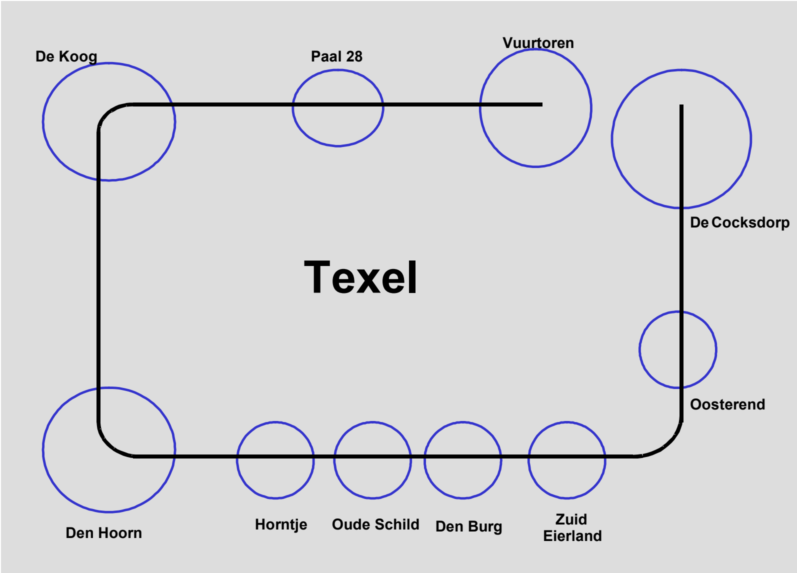
Beschrijf de belangrijkste karakteristieken van de wond:

Welke informatie uit de casus is het meest belangrijk voor u:

Hoe denkt u dat het ziektebeeld van deze cliënt zich zal ontwikkelen?

Railway: Situatieschets

Geef aan waar de meest opmerkelijke situatie van de afgelopen rit plaatsvond:



Beschrijf of teken -in het kort- die situatie, en geef aan welke treinen erbij betrokken waren:

Beschrijf of teken hoe deze situatie zich oploste:

Was er een betere oplossing mogelijk?

Railway: Technologie gebruik

Geef op een schaal van 1 tot 5 aan in hoeverre u het eens bent met de volgende stellingen (1=helemaal mee oneens, 2=gedeeltelijk mee oneens, 3=neutraal, 4=gedeeltelijk mee eens, 5=helemaal mee eens).

		1	2	3	4	5
1	Ik kan mijn werk sneller uitvoeren met de radar					
2	Ik kan mijn productiviteit verbeteren met de radar					
3	Ik vind de radar nuttig voor mijn werk					
4	Het gebruik van de radar vergroot mijn kansen op een salarisverhoging					
5	Mijn interactie met de radar zal duidelijk en begrijpelijk zijn					
6	Het zal eenvoudig voor mij zijn om bekwaam met de radar om te gaan					
7	Ik vind de radar eenvoudig te gebruiken					
8	Ik vind het leren omgaan met de radar eenvoudig					
9	Ik vind het een goed idee de radar te gebruiken in mijn werk					
10	Mensen die mijn gedrag beïnvloeden vinden dat ik de radar moet gebruiken					
11	Mensen die belangrijk zijn voor mij vinden dat ik de radar moet gebruiken					
12	Mijn management zal behulpzaam zijn bij het gebruik van de radar					

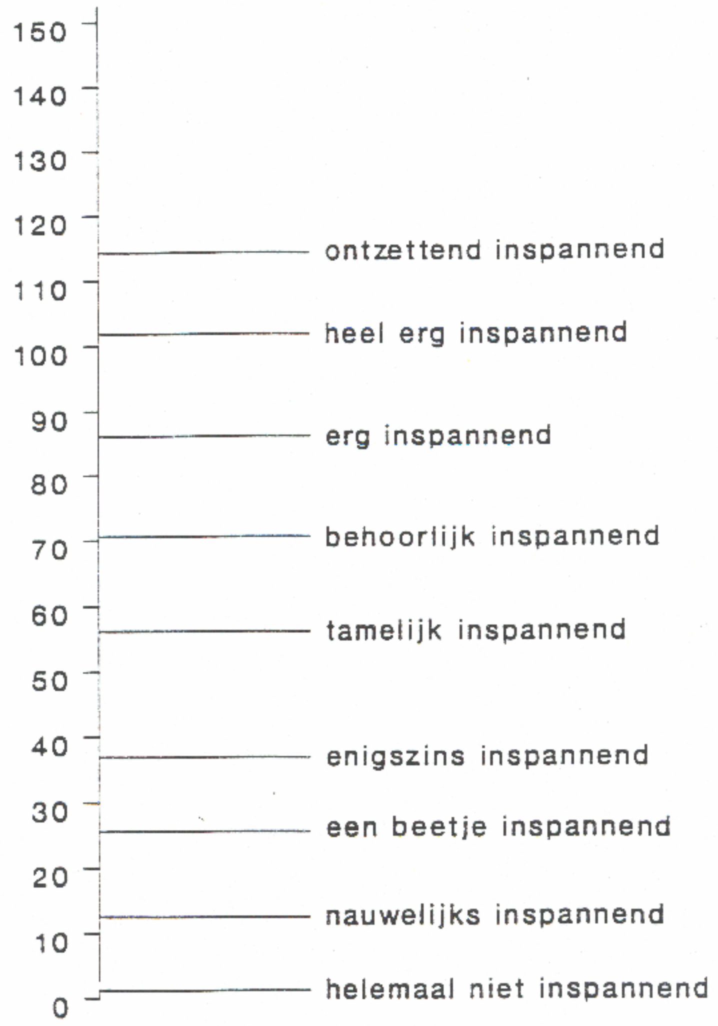
Home care: Technologie gebruik

Geef op een schaal van 1 tot 5 aan in hoeverre u het eens bent met de volgende stellingen (1=helemaal mee oneens, 2=gedeeltelijk mee oneens, 3=neutraal, 4=gedeeltelijk mee eens, 5=helemaal mee eens).

		1	2	3	4	5
1	Ik kan mijn werk sneller uitvoeren met deze technologie					
2	Ik kan mijn productiviteit verbeteren met deze technologie					
3	Ik vind deze technologie nuttig voor mijn werk					
4	Het gebruik van deze technologie vergroot mijn kansen op een salarisverhoging					
5	Mijn interactie met deze technologie zal duidelijk en begrijpelijk zijn					
6	Het zal eenvoudig voor mij zijn om bekwaam met deze technologie om te gaan					
7	Ik vind deze technologie eenvoudig te gebruiken					
8	Ik vind het leren omgaan met deze technologie eenvoudig					
9	Ik vind het een goed idee deze technologie te gebruiken in mijn werk					
10	Mensen die mijn gedrag beïnvloeden vinden dat ik deze technologie moet gebruiken					
11	Mensen die belangrijk zijn voor mij vinden dat ik deze technologie moet gebruiken					
12	Mijn management zal behulpzaam zijn bij het gebruik van deze technologie					

Mentale inspanning

Wilt u door middel van het zetten van een streepje op de verticale lijn aangeven hoeveel inspanning het u heeft gekost deze taak te verrichten?



Samenvatting

Technologische en maatschappelijke ontwikkelingen zorgen ervoor dat het samenwerken in professionele omgevingen verandert. Mobiele diensten worden ontwikkeld zodat mensen op verschillende locaties en op verschillende momenten met elkaar kunnen communiceren, samenwerken en relevante informatie kunnen raadplegen. Daarnaast bestaat de trend dat men het vakmanschap - de kennis en capaciteiten - van de professional beter wil benutten. Dit wordt ook wel kwaliteit van de arbeid of 'empowerment' genoemd.

Dit onderzoek bekeek in hoeverre informatie- en communicatie-ondersteuning het beslisgedrag van vakmensen beïnvloedt en wat vakmensen vinden van het gebruik en de toepasbaarheid ervan in hun werkpraktijk. Het onderzoek spitste zich toe op twee verschillende casussen: de thuiszorg en de spoorwegen. Beide casussen zijn voorbeelden van sociaal technische systemen, waarin mensen (sociaal) en systemen (technologie) samenwerken. In de thuiszorg ligt de nadruk op het samenwerken tussen mensen ondersteund door technologie. Bij de spoorwegen ligt de nadruk op de combinatie van mens en systeem, die gezamenlijk een groot technisch systeem besturen.

De onderzoeksvragen waren:

1. Wat is de toegevoegde waarde van informatie- en communicatieondersteuning op de taakprestaties van gedistribueerd werkende vakmensen?
2. Wat is de gebruikersacceptatie en de perceptie van de toepasbaarheid van de informatie- en communicatieondersteuning voor vakmensen?

Thuiszorg en spoorwegen

De thuiszorg voert bij patiënten thuis medische (bijvoorbeeld bloedsuikermetingen) en niet-medische (bijvoorbeeld wassen en schoonmaken) handelingen uit. In dit onderzoek lag de focus op de verzorging van chronische open wonden. Momenteel beschikken wijkverpleegkundigen over een mobiele telefoon om vanuit het huis van een patiënt contact op te nemen met de behandelend arts (dermatoloog of huisarts) of een collega voor advies. Op basis van interviews met 7 wijkverpleegkundigen en 2 dermatologen is een mobiele dienst ontwikkeld,

WondLog genaamd. Hiermee kunnen wijkverpleegkundigen bij een patiënt thuis beschikken over een drietal functionaliteiten: (1) gestructureerde wondrapportages en wondfoto's bijhouden (wondlogboek), (2) multimedia berichten met gerichte vragen naar dermatologen of collegae sturen en antwoorden ontvangen, en (3) videoconferenties uitvoeren met dermatologen of collegae. Op basis van deze informatie- en communicatiemogelijkheden zou een wijkverpleegkundige in staat moeten zijn om in het huis van de patiënt een betere behandelingsbeslissing te nemen.

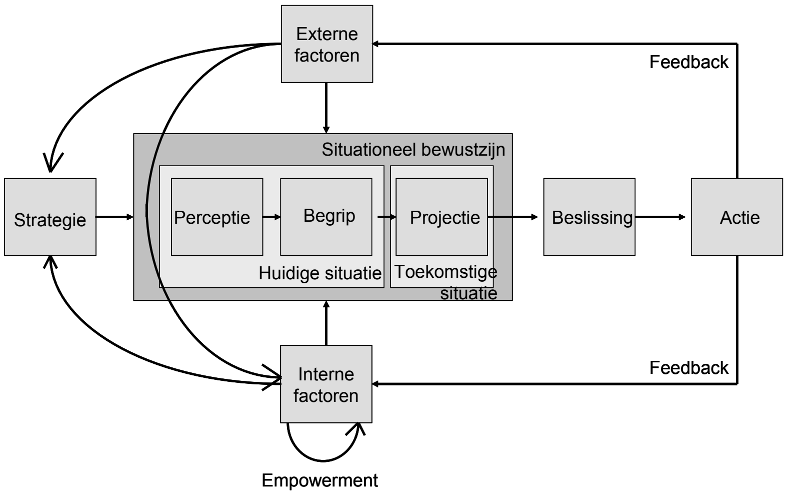
Op het spoor bedient de treindienstleider de seinen en wissels en bepaalt aan de hand van planregels welke trein waar mag rijden en wanneer. Een machinist volgt seinbeelden op en zorgt voor het veilige en efficiënte vervoer van passagiers. Bij de spoorwegen lag de nadruk op de informatieomgeving van een machinist. Relevante informatie voor een machinist op de bok is zijn buitenbeeld bestaande uit seinen en borden, zijn cabine informatie, zoals snelheidsmeter, Automatische Trein Beïnvloeding (ATB) en wegbekendheid. Wegbekendheid is door ervaring opgebouwde kennis over het baanvak: waar moet ik gaan remmen, wanneer kan ik een laag sein verwachten etc. Op basis van interviews met 6 machinisten en 2 treindienstleiders is een informatiescherm ontwikkeld, radar genaamd. Hierop kunnen machinisten zien wat hun toekomstige seinbeeld en rijweg zal zijn, welke treinen er in hun directe omgeving zijn en wat daarvan het toekomstige seinbeeld en rijweg zal zijn, en wat hun huidige snelheid en punctualiteit is. Hiermee kunnen machinisten beslissingen nemen over hun rijhandelingen (remmen, optrekken en uit laten rollen), zij kunnen hun reizigers informatie geven, en gaan communiceren met de treindienstleider.

Het idee achter beide mobiele applicaties is dat zij meer informatie verschaffen over de toekomstige situatie, en zodoende het beslisgedrag van de vakmensen positief beïnvloeden. Wijkverpleegkundigen kunnen aan het wondlogboek de recente ontwikkeling van de wond zien. Machinisten kunnen op de radar zien hoe hun toekomstige rijweg eruit zal zien, en of zij rekening kunnen houden met verkeer achter zich.

Theoretisch achtergrond

Twee uiteenlopende voorbeelden van sociaal technische systemen zijn geselecteerd om te identificeren wat het effect is van informatie- en communicatieondersteuning op de taakprestatie van gedistribueerd werkende vakmensen. Het cognitieve proces dat ten grondslag ligt aan de taakuitvoering op basis van extra informatie- en communicatieondersteuning is beschreven in het taakprestatiemodel. Het taakprestatiemodel is gebaseerd op het informatieverwerkingsmodel van Wickens (1992). Het generieke principe is dat mensen informatie waarnemen, een beslissing nemen en een actie uitvoeren. Door hun ervaringen hanteren vakmensen vaak een bepaalde strategie op basis waarvan ze hun taak uitvoeren. Figuur 1 geeft een overzicht van het taakprestatiemodel dat toegepast is in dit onderzoek.

Figuur 1 Taakprestatie model



Het belangrijkste cognitieve construct in dit onderzoek is het beslisgedrag. Beslisgedrag van vakmensen wordt het best omschreven door de Natuurlijk Beslisgedrag theorie (*Naturalistic Decision Making*) (Zsambok en Klein, 1997). Deze theorie bestudeert het beslisgedrag van vakmensen, die individueel of in groepen werken, in omgevingen met veel (soms tegenstrijdige) informatie. Deze vakmensen staan vaak ook onder tijdsdruk om een beslissing te nemen. Hoe vakmensen beslissingen nemen beschrijft het Herkenning Gebaseerd Beslismodel (*Recognition Primed Decision model*) (Klein, 1989). Dit model stelt dat vakmensen –in het algemeen- een situatie vaststellen en beoordelen of zij dit eerder hebben meegemaakt. Vervolgens maken zij zich een voorstelling om relevante acties te identificeren, gebaseerd op een 'goed genoeg' principe.

Taakprestaties

Om te kunnen identificeren wat de toegevoegde waarde van de informatie- en communicatieondersteuning is op de taakprestaties dienen de taakprestaties eerst gedefinieerd te worden. Vicente (1999) beschrijft dat er 3 generieke prestatiematen zijn voor sociaal technische systemen: veiligheid, productiviteit en veiligheid. Middels de interviews met de wijkverpleegkundigen en machinisten zijn deze prestatiematen doorvertaald naar casus specifieke prestatiematen. De prestatiematen voor wijkverpleegkundigen zijn: kwaliteit en efficiëntie van de zorg, patiënt gerichtheid en kwaliteit van de arbeid voor de wijkverpleegkundige. De prestatiematen voor de machinisten zijn: veiligheid, punctualiteit, energiezuinigheid, klantgerichtheid en kwaliteit van de arbeid voor de machinisten.

Om de toegevoegde waarde van de informatie- en communicatieondersteuning op de taakprestaties te kunnen bepalen hebben wijkverpleegkundigen en machinisten gebruikersexperimenten uitgevoerd. In een gecontroleerde omgeving moesten zij authentieke taken uitvoeren, waarbij zij ondersteund werden door de informatie- en communicatieondersteuning. Taakprestaties met en zonder informatie- en communicatieondersteuning zijn met elkaar vergeleken. Om een adequaat beeld te krijgen van de taakprestaties en de percepties van de vakmensen betreffende de toepasbaarheid en bruikbaarheid van de informatie- en communicatieondersteuning is een combinatie van – zowel kwantitatieve als kwalitatieve en zowel objectieve als subjectieve – methoden en instrumenten toegepast.

Gebruikersexperimenten

In beide casussen zijn twee gebruikersexperimenten uitgevoerd. In iedere casus was het eerste experiment een eenvoudig en kortdurend experiment. Het tweede experiment in beide casussen kwam meer overeen met de werkzaamheden in de praktijk.

Het algemene doel van de gebruikersexperimenten met de wijkverpleegkundigen was te identificeren of de historische wondinformatie en de communicatie met een collega de behandelingsbeslissing positief zou beïnvloeden. Behandelingsbeslissingen die in de praktijk voorkomen zijn: normaal, semi-urgent en urgent. Normaal verwijst naar patiënten die normaal behandeld kunnen worden, semi-urgent verwijst naar patiënten waarvoor later die dag een advies van een collega of arts noodzakelijk is, en urgent verwijst naar patiënten die onmiddellijk gezien moeten worden door de behandelende arts. Het casus materiaal van 18 patiënten met open wonden was verzameld. Om deze casussen te classificeren hebben wijkverpleegkundigen dit casusmateriaal onafhankelijk van elkaar beoordeeld: een vakgroeplid en een duo van ervaren wijkverpleegkundigen. De overeenkomst tussen deze beoordelingen was erg laag. Op basis hiervan is besloten 2 referentiekaders op te stellen om de behandelingsbeslissingen van de wijkverpleegkundigen te beoordelen: normaal, semi-urgent en urgent en normaal en (semi)urgent. Beide referentiekaders zijn gedefinieerd op basis van de scores van het vakgroeplid.

Wijkverpleegkundigen maken in hun huidige werkzaamheden weinig gebruik van technologische hulpmiddelen (alleen een mobiele telefoon). De ontwikkelde dienst WondLog draait op een laptop met camera. De onervarenheid in het omgaan met de laptop zou wel eens een negatief effect kunnen hebben op de behandelingsbeslissing (hoe krijgt men de juiste informatie in beeld). Daarom kregen 20 wijkverpleegkundigen eerst een papieren vragenlijst voorgelegd, waarop zij patiënten casussen met en zonder historische wondinformatie moesten beoordelen. Daarnaast dienden zij vragen te beantwoorden over de toepasbaarheid van deze informatie in de praktijk. Het belangrijkste resultaat is dat historische wondinformatie geen effect heeft op behandelingsbeslissingen, getoetst met beide referentiekaders. De percepties van de wijkverpleegkundigen waren echter positief: zij verwachten positieve effecten van de historische wondinformatie op de kwaliteit en efficiëntie van zorg en op de klantgerichtheid. Tevens verwachtten zij dat de kwaliteit van de arbeid toe zal nemen.

In het tweede experiment kregen 20 wijkverpleegkundigen wederom patiënten casussen voorgelegd. Zij kregen op 4 verschillende manieren informatie- en communicatieondersteuning. De informatie werd conditioneel aangeboden zonder technologie (vragenlijst van het eerste experiment) en op laptop via WondLog. De communicatie werd conditioneel aangeboden via de mobiele telefoon (zonder wondinformatie), en op de laptop via WondLog (informatie en communicatie). Wederom dienden zij vragen te beantwoorden over het gebruik en de toepasbaarheid van WondLog in de praktijk. Het belangrijkste resultaat is dat de informatie- en communicatieondersteuning geen effect had op de behandelingsbeslissing. Wederom waren de percepties van de wijkverpleegkundigen positief. Zij verwachtten dat Wondlog een positief effect heeft op de kwaliteit en efficiëntie van zorg en de klantgerichtheid. Tevens verwachtten zij dat de kwaliteit van de arbeid toe zal nemen.

Het algemene doel van de gebruikersexperimenten met de machinisten was te identificeren of de informatie- en communicatieondersteuning het rijgedrag van machinisten positief zou beïnvloeden. In de praktijk heeft een machinist weinig vrijheidsgraden in zijn rijgedrag: optrekken, afremmen, uit laten rollen en communiceren met de treindienstleider. Machinisten krijgen tijdens het rijden verschillende soorten dynamische informatie aangeboden (cabine informatie, seinen, borden etc.) en werken onder tijdsdruk. De dynamiek van de informatie en de tijdsdruk zijn van grote invloed op hun beslisgedrag. Aangezien de ontwikkelde radar informatie nieuw is voor machinisten, kan dit gaan interfereren met de dynamische informatie en tijdsdruk. Daarom bestond het eerste experiment uit een statische simulatie. Machinisten (58 in totaal) kregen foto's van authentieke situaties op het spoor aangeboden verrijkt met radar informatie. De radar bood onder andere informatie over de positie en snelheid van andere treinen en het toekomstige seinbeeld van hun eigen trein. De volgorde van de aangeboden informatie werd hierbij steeds rijker. Naast rijbeslissingen dienden de machinisten ook vragen te beantwoorden over het gebruik en de toepasbaarheid van de radars. De belangrijkste resultaten zijn dat de machinisten andere rijbeslissingen nemen op basis van de radarinformatie. Tevens verwachtten de machinisten dat de contextinformatie een positief effect heeft op alle prestatiematen: veiligheid, punctualiteit, energiezuinigheid en klantgerichtheid. Voor veiligheid verwachtten zij echter tegelijkertijd dat de radar hun aandacht zal afleiden van het belangrijke seinbeeld buiten op de rails. Tevens verwachtten zij dat de kwaliteit van de arbeid toe zal nemen.

Dynamische informatie en tijdsdruk waren wel aanwezig in het tweede experiment In een treinsimulator - waar vier machinisten tegelijk rijden onder leiding van één treindienstleider - zijn treinritten uitgevoerd, conditioneel ondersteund door radarinformatie. In totaal deden 48 machinisten mee aan het experiment, waarbij iedere machinist zes ritten reed. Het belangrijkste resultaat is dat radarinformatie geen effect heeft op de prestatiematen. Wederom hebben de machinisten wel positieve verwachtingen van de contextinformatie in hun werkpraktijk. Zij verwachten dat de punctualiteit, energiezuinigheid en klantgerichtheid erdoor zal toenemen. Maar zij verwachtten ook afgeleid te worden door de radar van het belangrijke seinbeeld buiten. Tevens verwachtten zij dat de kwaliteit van de arbeid toe zal nemen.

In het algemeen laten de gebruikersexperimenten geen effecten zien van de informatie- en communicatieondersteuning op de belangrijkste taakprestaties. De verwachtingen van de vakmensen zijn echter positief. Verklaringen voor deze algemene resultaten op theoretisch, praktisch en methodologisch gebied staan in de volgende paragraaf beschreven. Tevens is de relevantie van deze studie voor ICT innovatie onderzoek beschreven.

Theoretische inzichten

Het Herkenning Gebaseerd Beslismodel (Recognition Primed Decision model) stelt dat vakmensen beslissingen nemen op basis van eerdere ervaringen. Zij bepalen eerst of ze een situatie herkennen en dan selecteren ze de acties die ze destijds hebben genomen. Op basis van de aangeboden informatie- en communicatieondersteuning ontstaan nieuwe situaties, die (nog) niet herkend kunnen worden door de vakmensen. Als de vakmensen deze situaties niet als nieuw detecteren, maar als situaties die ze eerder beleefd hebben, wordt dus een actie geselecteerd die gericht is op de oude situatie. Op deze manier wordt de informatie- en communicatieondersteuning dus niet meegenomen in het beslisproces. Waarschijnlijk vereist nieuwe informatie- en communicatieondersteuning een nieuw beslisproces, waar men door de geringe tijdsduur van het experiment niet aan toegekomen is.

De toegepaste taakprestatiestrategie bepaalt de manier waarop informatie wordt waargenomen, begrepen en geprojecteerd (zie Figuur 1). Bekende strategieën van wijkverpleegkundigen zijn analytisch en intuïtieve strategieën om tot beslishandelingen te komen. Bij een analytische strategie vindt een zorgvuldige afweging plaats van alle variabelen, terwijl bij een intuïtieve strategie een ervaringsgebaseerde afweging plaats vindt. Beide strategieën zijn niet tegenstrijdig, maar komen in de praktijk vaak gecombineerd voor. Machinisten hanteren grofweg twee verschillende strategieën of rijstijlen: anticiperen op seinen en situaties die komen gaan, of reageren op ATB signalen. Het is mogelijk dat door de toegepaste strategieën van de vakmensen de nieuwe informatie- en communicatieondersteuning niet meegenomen wordt in hun waarneming- en beslisproces. Voor wijkverpleegkundigen zal de intuïtieve strategie de nieuwe informatie niet meenemen. Machinisten met een rijstijl van reageren op ATB signalen zullen de nieuwe informatie ook niet meenemen.

Bovenstaande verklaringen geven aan dat nieuwe informatie- en communicatieondersteuning niet onmiddellijk wordt opgenomen in werkroutines, althans niet tijdens de duur van het experiment. Het cognitieve mechanisme dat hieraan ten grondslag kan liggen is kennisinkapseling (knowledge encapsulation). Kennisinkapseling verwijst naar het inkapselen van gedetailleerde kennis in causale modellen onder invloed van context informatie. Op basis van leerervaringen wordt dus nieuwe kennis omgezet in causale modellen. Voor wijkverpleegkundigen betekent dit dat zij problemen hadden de historische wondinformatie in hun interne wondkennismodel in te kapselen. De leerervaring in dit experiment was waarschijnlijk niet lang genoeg. Hoewel het begrip kennisinkapseling voortkomt uit de medische wetenschap, kan het in de toegepaste betekenis doorvertaald worden naar het gedrag van machinisten in dit onderzoek. Machinisten ondervonden problemen met het inkapselen van radarinformatie in hun interne kennismodel. Ook hier was de leerervaring in dit experiment waarschijnlijk niet lang genoeg.

Aandachtspunten bij de gebruikte methodologie

Twee verschillende soorten methoden zijn toegepast in dit onderzoek: ontwikkeling van informatiedisplays met eindgebruikers en gebruikersexperimenten. Beide methoden worden besproken en kanttekeningen worden geplaatst.

Beide informatiedisplays werden ontwikkeld door en voor vakmensen. De displays kwamen tot stand in verschillende interviewsessies en workshops. De gebruikersacceptatie en de percepties over de toepasbaarheid van de informatiedisplays waren positief. Echter, de taakprestaties werden niet beïnvloed in de experimenten. Een mogelijke verklaring kan zijn dat de informatiedisplays gebaseerd waren op taakuitvoeringstrategieën waarin de vakmensen over voldoende tijd en mogelijkheden beschikten om tot beslissingen te komen (interviews en workshops). In authentieke taken - zoals uitgevoerd in de tweede reeks experimenten - staan de vakmensen onder tijdsdruk en krijgen veel informatie aangeboden. Waarschijnlijk hanteren zij in die experimenten een andere strategie om tot een taakuitvoering te komen, waardoor niet alle displayinformatie benut wordt.

Ten tweede vonden gebruikersexperimenten plaats met de informatiedisplays waarin getracht werd de praktijkomgeving zo goed mogelijk na te bootsen. Echter, dat kan bij beide casussen nog verbeterd worden. Om het experimentele materiaal zo constant mogelijk te houden kregen wijkverpleegkundigen papieren patiëntencasussen ter beoordeling voorgelegd. De foto's op papier kunnen qua kleur enigszins afwijken van echte patiënten. Daarnaast zijn wijkverpleegkundigen niet gewend patiënten vanaf papier te beoordelen. Wegbekendheid is een cruciaal construct voor machinisten om hun vakkennis in te zetten en te gebruiken. Helaas was het niet mogelijk in de huidige simulator wegbekendheid op een adequate manier weer te geven. Alle genoemde factoren kunnen van invloed zijn geweest op de resultaten van de experimenten.

De inzet van vakmensen doet concessies aan het aantal personen dat deelneemt aan de experimenten en de tijdsduur van de experimenten. Het relatief kleine aantal personen dat deelneemt, heeft een negatief effect op de validiteit van het onderzoek. De beperking in de tijdsduur van het experiment kan een negatief effect hebben op de training in het omgaan met de nieuwe applicatie en informatie.

ICT innovatie onderzoek

Dit onderzoek bestudeerde het effect van informatie- en communicatieondersteuning op de taakprestaties van gedistribueerd werkende vakmensen. Vaak is het feit dat technologische applicaties ontwikkeld zijn al voldoende reden om ze in de praktijk te introduceren. Echter, om informatie- en communicatie ondersteuning bruikbaar, nuttig en kosteneffectief in te zetten moeten zij afgestemd zijn op de cognitieve en affectieve aspecten van de toekomstige gebruikers. Vooral als deze gebruikers vakmensen zijn. Vakmensen passen namelijk verschillende strategieën toe om hun taken uit te voeren, die verschillende ondersteuning behoeven.

Dit onderzoek toonde de relevantie aan van kosteneffectieve methoden voor ICT-innovatie-onderzoek. Namelijk de vroegtijdige onderkenning van verwachte taakprestaties en suggesties om de ontwerpen aan te passen op wensen uit de praktijk. Gebruikersacceptatie en de perceptie van de toepasbaarheid zijn waardevolle indicaties van het succes van een nieuwe applicatie in de praktijk. Positieve percepties van eindgebruikers kunnen namelijk als een stimulans werken om het in de praktijk toe te passen.

Aanbevelingen voor de praktijk
In vergelijking tot andere thuiszorgorganisaties mogen veel wijkverpleegkundigen van de thuiszorgorganisatie die participeerde in dit project wondzorgtaken verrichten. Op deze manier ziet iedere wijkverpleegkundige weinig open wonden tijdens haar werkzaamheden. Andere organisaties hebben speciale teams opgericht om wondzorgtaken te verrichten. Deze teams worden begeleid en getraind door dermatologen. Zo ontstaan teams die gespecialiseerd zijn in het verrichten van wondzorg. Dus wordt geadviseerd om de wondzorg kwaliteit te verhogen door specialistische wondzorg teams samen te stellen. Voordat WondLog in de praktijk geïntroduceerd of getest wordt dienen enige technische aanpassingen verricht te worden (o.a. het gebruikersgemak van de camera), tevens dient men aandacht te besteden aan de acceptatie door patiënten.

Door toename in de bezettingsgraad op het spoor is de regelruimte van machinisten in de afgelopen jaren afgenomen. De baanvaksnelheid en de dienstregelingsnelheid zijn (bijna) aan elkaar gelijk, waardoor machinisten bijna geen mogelijkheden hebben om vertragingen in te lopen of aan energiezuinig rijden te denken. Handelingsruimte voor machinisten kan gecreëerd worden door de baanvaksnelheid en dienstregelingsnelheid van elkaar te laten verschillen.

Machinisten hanteren grofweg twee verschillende rijstijlen: anticiperen op seinen en situaties die komen gaan, of reageren op ATB signalen. De rijstijl waarmee men reageert op ATB signalen geeft machinisten als het ware iets meer tijd om een achterstand in te lopen. Op deze manier creëert hij voor zichzelf iets meer handelingsruimte om een vertraging in te lopen. De radar die ontwikkeld is in dit onderzoek ondersteund de anticiperende rijstijl. Dus, voordat radarinformatie in de praktijk getoetst kan worden wordt geadviseerd de beschikbare handelingsruimte van machinisten te vergroten door de baanvaksnelheid en dienstregelingsnelheid van elkaar te laten verschillen. Machinisten dienen dan de anticiperende rijstijl aan te houden om de radarinformatie effectief in te kunnen zetten.

References

Aris, P., & Van den Dikkenberg, R. (2003). *Kick-off workshop Rij-Optimalisatie* [Kick-off workshop Driving Optimalisation]. Internal report ICT Breedspoor.

Baker, C. P. (2001). The WOC nurse in home care. *Journal of Wound, Ostomy and Continence Nursing*, 28, 270-273.

Beach, L. R. & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review*, 3, 439-449.

Beaudouin-Lafon, M., & MacKay, W. (2003). Prototyping Tools and Techniques. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* (pp. 1006-1031). Mahwah, NJ: Lawrence Erlbaum Associates.

Biemans, M., Swaak, J., Hettinga, M., & Schuurman, J. G. (2005). Involvement matters: On the proper involvement of users and behavioural theories in the design of a medical teleconferencing application. In K. Schmidt, M. Pendergast, M. Ackerman, & G. Mark (Eds.), *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*, pp. 304-312

Biemans, M., Swaak, J., Van der Velde, S., & Huppertz, M. (2005). Empowering train drivers: designing and testing context information. In C.Sandom & I MacLeod (Eds.), *Proceedings of the IEE Human Factors for Engineers Professional Network*,. London, pp. 17-24

Blandford, A. & Wong, B. L. W. (2004). Situation Awareness in emergency medical dispatch. International *Journal of Human-Computer Studies,* 61, 421-452.

Boshuizen, H. P. A., Schmidt, H. G., Custers, E. J. F. M., & Van de Wiel, M. W. (1995). Knowledge development and restructuring in the domain of medicine: The role of theory and practice. *Learning and Instruction*, 5, 269-289.

Bricon-Souf, B., Anceaux, F., Bennani, N., Dufresne, E., & Watbled, L. (2005). A distributed coordination platform for home care: analysis, framework and prototype. International *Journal of Medical Informatics*, 74, 809-825.

Bryans, A. & McIntosh, J. (1996). Decision making in community nursing: an analyis of the stages of decision making as they relate to community nursing assessment practice. *Journal of Advanced Nursing*, 24, 24-30.

Buss, I. C., Halfsen, R. J. G., Huyer Abu-Saad, H., & Kok, G. (2004). Pressure ulcer prevention in nursing homes: views and beliefs of enrolled nurses and other health care workers. *Journal of Clinical Nursing*, 13, 668-676.

Bye, R., Farrington-Darby, T., Cox, G., Hockey, G. R. J., Wilson, J. R., & Clarke, T. (2005). Grounding research: From work analysis to a macrocognitive representation of integrated rail operations. In *Proceedings of the Second Rail Human Factors Conference*, London.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of Management Review*, 13, 40-52.

Chu, P. C. & Spires, E. E. (2003). Perceptions of accuracy and effort of decision strategies. *Organizational Behavior and Human Decision Processes*, 91, 203-214.

Currey, J. & Botti, M. (20063). Naturalistic decision making: A model to overcome methodological challenges in the study of critical care nurses' decision making about patients' hemodynamic status. *American Journal of Critical Care,* 12, 206-211.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, September, 319-340.

De Bakker S.C. (2006). *Organizational members in the information society: coping with information overload*. Ph.D. Thesis, Universiteit van Amsterdam.

De Groot, A. D. (1961). *Methodologie* [Methodology]. Den Haag: Mouton.

De, P. & Scarpello, J. H. B. (1999). What is the evidence for effective treatment of diabetic foot ulceration? *Practical diabetes international: the journal for diabetes care teams worldwide,* 16, 179-183.

Dowding, D. & Thompson, C. (2003). Measuring the quality of judgement and decision-making in nursing. *Journal of Advanced Nursing*, 44, 49-57.

Endsley, M. R. (1995a). Measurement of Situation Awareness in dynamic systems. *Human Factors*, 37, 65-84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37, 32-64.

Eneroth, M., Larsson, J., Apelqvist, J., Reike, H., Salomon, M., Gough, A., Foster, A., & Edmonds, M. E. (2004). The challenge of multicenter studies in diabetic patients with foot infections. *The Foot*, 14, 198-203.

ERTMS (2005). *European Rail Traffic Management System.* http://www.ertms.com [On-line].

Friedberg, E. H., Harrison, M. B., & Graham, I. D. (2002). Current home care expenditures for persons with leg ulcers. *Journal of Wound, Ostomy and Continence Nursing*, 29, 186-192.

Gaba, D. M. & Howard, S. K. (1995). Situation Awareness in anaesthesiology. *Human Factors*, 37, 20-31.

Gatley, E. P. (1992). From novice to expert: the use of intuitive knowledge as a basis for district nurse education. *Nurse Education Today*, 12, 81-87.

Hamilton, W. I. & Clarke, T. (2005). Driver performance modelling and its practical application to railway safety. *Applied Ergonomics,* 36, 661-670.

Hammond, K. R. (1996). *Human judgement and social policy. Irreducible uncertainty, inevitable error, unavoidable injustice.* New York: Oxford University Press.

Haram, R., Ribu, E., & Rustøen, T. (2003). The views of district nurses on their level of knowledge about the treatment of leg and foot ulcers. *Journal of Wound, Ostomy and Continence Nursing*, 30, 25-32.

Hettinga, M., Biemans, M., Peddemors, A., Salden, A., Van der Spek, J., & Benz, H. (2003). *Pilot Draadloze Zorg: Gebruikersstudie* [Pilot Wireless Care: User studie]. Telematica Instituut, Enschede, The Netherlands. Retrievable from: https://doc.telin.nl/dscgi/ds.py/ViewProps/File-35348.

Hollnagel, E., & Woods, D. D. (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*. Boca Raton: CRC Press, Taylor & Francis Group.

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: implications for job redesign. *Administrative Science Quarterly*, 24, 285-308.

Kecklund, L., Ingre, M., Kecklund, G., Söderström, M., Åkerstedt, T., Lindberg, E., Jansson, A., Olsson, E., Sandblad, B., & Almqist, P. (2001). The Train project: Railway safety and the train driver information environment and work situation. In *Proceedings of Second Signalling Safety*, London, 26-27 February.

Klein, G. A. (1989). Recognition-primed decisions. *Advances in Man-Machine Systems Research*, 5, 47-92.

Klein, G. A. (1999). *Sources of power. How people make decisions*. London, England: The MIT press.

Klein, G. A. & Calderwood, R. (1991). Decision models: Some lessons from the field. *IEEE Transactions on Systems, Man, and Cybernetics*, 21, 1018-1026.

Lamminen, H., Voipio, V., & Ruohonen, K. (2001). Trends in clinical practice: Telemedicine framework and applications in dermatology and ophthalmology. *Annals of Medicine*, 33, 222-228.

Landauer, T. K. (1995). *The trouble with computers*. MIT Press, London, England.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.

Lauri, S. & Salanterä, S. (2002). Developing an instrument to measure and describe clinical decision-making in different nursing fields. *Journal of Professional Nursing*, 18, 93-100.

Lipshitz, R., Klein, G., Orasuma, J., & Salas, E. (2001). Focus article: Taking stock of naturalistic decision-making. *Journal of Behavioral Decision Making*, 14, 331-352.

Livingstone, H., Gipson, T., & Luther, R. (2005). Development of a risk-based approach to training and assessment in safety critical roles. In *Proceedings of Training, Education & Simulation International (TESI 2005)*, Maastricht, The Netherlands.

Lucas, E. (2005). *Jaarbericht. ProRail aan het werk in 2004* Utrecht: ProRail. Retrieved from www.prorail.nl

Malone, T. W. (2004). *The future of work*. Harvard Business School Press.

May, C., Finch, T., Mair, F., & Mort, M. (2005). Towards a wireless patient: Chronic illness, scarce care and technological innovation in the United Kingdom. *Social Science & Medicine*, 61, 1485-1494.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Prentice-Hall, New Jersey.

McIntosh, J. (1996). The question of knowledge in district nursing. *International Journal of Nursing Studies*, 33, 316-324.

McLeod, R. W., Walker, G. H., & Moray, N. (2005). Analysing and modelling train driver performance. *Applied Ergonomics*, 36, 671-680.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement.* (3rd ed., pp. 13-104). New York: American Council on Education.

Michel-Verkerk, M. B., Schuring, R. W., & Spil, T. A. M. (2003). USE IT or Leave IT: A model to reveal user satisfaction in healthcare process ex ante and ex post. In *Proceedings of the Third International Conference of Healthcare and Medical Technology,* London, UK.

Ministerie van VWS (2005). *Thuiszorg* [Home care]. Ministerie van Volksgezondheid, Welzijn en Sport [On-line]. Retrieved from http://www.minvws.nl/dossiers/thuiszorg/default.asp

Nielsen, J. (1993). *Usability engineering.* San Fransisco: Morgan Kaufmann Publishers.

NS (2005). *NS Dienstregelingen en informatie* [Dutch Railway Timetables and Information]. http://www.ns.nl/servlet/Satellite?cid=1074533960277&pagename=www.ns.nl%2FPage%2FArtikelPage_www.ns.nl&lang=nl&c=Page [On-line].

O'Donnel, C. R., & Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Volume II: Cognitive processes and performance ed., pp. 42.1-42.29). New York: Wiley.

Oakley, A., & Wootton, R. (2002). Introduction. In R. Wootton & A. Oakley (Eds.), *Teledermatology* (pp. 3-9). London: The Royal Society of medicine Press limited.

Oprins, E., & Schuver, M. (2003). Competentie gericht opleiden en beoordelen bij LVLN (in Dutch). *Human Factors Advisory Group* (HUFAG) Nieuwsbrief, Winter 2003, 2-4.

Paquette, L. & Kida, T. (1988). The effect of decision strategy and task complexity on decision performance. *Organizational Behavior and Human Decision Processes*, 41, 128-142.

Patistea, E. & Siamanta, H. (1999). A literature review of patients' compared with nurses' perceptions of caring: Implications for practice and research. *Journal of Professional Nursing*, 15, 302-312.

Perednia, D. (2002). Foreword. In R. Wootton & A. Oakley (Eds.), *Teledermatology* (pp. xiii-xv). London: The Royal Society of Medical Press limited.

Pieper, B., Templin, T., Dobal, M., & Jacox, A. (1999). Wound prevalance, types, and treatments in home care. *Advances in Wound Care,* 12, 117-126.

Pinelle, D., & Gutwin, C. (2003). Designing for loose coupling in mobile groups. In *Proceedings of the ACM. Group '03*, November 9-12, (pp. 75-84) Sanibel Island, Florida.

Pirolli, P. & Card, S. (1999). Information Foraging. *Psychological Review*, 106, 643-675.

ProRail (2005). ProRail: *Feiten en cijfers.* http://www.prorail.nl/ProRail/Over+ProRail/Feiten+en+cijfers.htm [On-line].

Raad voor de Transport Veiligheid (2001). *Botsing tussen twee reizigerstreinen in Dordrecht 28 november 1999* [Collision between two passenger trains in Dordrecht, November 28th, 1999] Den Haag. Retrieved from www.rtvt.nl

Ramsay, M., & Nielsen, J. (2000). *WAP Usability: Déjà Vu: 1994 all over again*. Nielsen Norman Group, Fremont, USA.

Randel, J. M. & Pugh, H. L. (1996). Differences in expert and novice situation awereness in naturalistic decision making. *International Journal of Human-Computer Studies*, 45, 579-597.

Rasmussen, J. (1976). Outlines of a hybrid model of the process plant operator. In T. B. Sheridan & G. Johannsen (Eds.), *Monitoring behavior and supervisory control*. (pp. 371-383). New York: Plenum.

Rasmussen, J., Pejtersen, A. M., & Schmidt, K. (1990). *Taxonomy for cognitive work analysis.* (Rep. No. Risø-M-2871). Risø National Laboratory.

Ribu, E., Haram, R., & Rustøen, T. (2003). Observations of nurse's treatment of leg and foot ulcers in community healthcare. *Journal of Wound, Ostomy and Continence Nursing*, 30, 342-350.

Robbins, T. L., Crino, M. D., & Fredenhall, L. D. (2002). An integrative model of the empowerment process. *Human Resource Management Review*, 12, 419-443.

Roe, B. H., Luker, K. A., Cullum, N. A., Griffiths, J. M., & Kenrick, M. (1993). Assessment, prevention and monitoring of chronic leg ulcers in the community: report of a survey. *Journal of Clinical Nursing*, 2, 299-306.

Salvucci, D. D. & Macuga, K. L. (2002). Predicting the effects of cellular-phone dialling on driver performance. *Cognitive Systems Research*, 3, 95-102.

Schaffers, H., Brodt, T., Pallot, M., & Prinz, W. (2006). *The future workspace. Perspectives on mobile and collaborative working.* MOSAIC consortium, Telematica Instituut, The Netherlands.

Schmidt, H. G. & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, 5, 205-221.

Schotanus, B., & Zigterman, L. (2004). De orde van het seinhuis en kan de mens het spoor veiliger en efficiënter maken. *Colloquium Vervoersplanologisch Speurwerk 2004*.

Schön, D. A. (1991). *The reflective turn: Case studies in and on educational practice*. Teachers College, Columbia University, New York.

Simon, H. A. (1957). *Models of man: social and rational: mathematical essays on rational human behavior in a social setting*. New York: Wiley and Son.

Smits, R. (2002). Innovation studies in the 21st century: Questions from a user's perspective. *Technological Forecasting & Social Change*, 69, 861-883.

Speer, P. W. & Peterson, N. A. (2006). Psychometric properties of an empowerment scale: Testing cognitive, emotional and behavioral domains. *Social Work Research*, 24, 109-118.

Spreitzer, G. M. (1995). Psychological empowerment in the workplace: dimensions, measurement and validation. *Academy of Management Journal*, 38, 1442-1465.

Stanton, N., Salmon, P., Walker, G., Baber, C., & Jenkins, D. (2005). *Human Factors Methods: A practical guide for engineering and design*. Ashgate, England.

Thomas, K. & Velthouse, B. (1990). Cognitive elements of empowerment: An "interpretive" model of intrinsic task motivation. *Academy of Management Review*, 15, 666-681.

UMTS forum (2000). *The UMTS third generation market - Structuring the service revenue opportunities*. (Rep. No. 9).

Van der Velde, S., Huppertz, M., & Biemans, M. (2004). *Verkenning de nieuwe werkpraktijk* [Exploration of new ways of working in practice] ICT Breedspoor: Het Spoor Meester.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27, 425-478.

Verwey, W. B. & Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology*, 2, 270-285.

Vicente, K. J. (1999). *Cognitive Work Analysis. Toward safe, productive and healthy computer-based work*. Lawrence Erlbaum Associates, Publishers London.

Visco, D., Shalley, T., Wren, S. J., Flynn, J. P., Brem, H., Kerstein, M. D., & Fitzpatrick, J. J. (2001). Use of telehealth for chronic wound care: A case study. *Journal of Wound, Ostomy and Continence Nursing*, 28, 89-95.

Wickens, C. D. (1992). *Engineering psychology and human performance.* (2nd edition ed.) HarperCollins Publisher.

Wilson, E. V., & Connolly, J. R. (2001). Effects of group task pressure on e-mail and face-to-face communication effectiveness. In Proceedings of the *ACM Group 2001*, Boulder Colorado, USA (pp. 270-278).

Wilson, E. V., & Morrison, J. P. (2000). A measure of task-technology fit for computer mediated communication. In S. Clarke & B. Lehaney (Eds.), *Human Centered Methods in Information Systems: Current Research and Practice.* Hershey, PA.

Wilson, J. R. & Norris, B. J. (2005a). Rail human factors: Past, present and future. *Applied Ergonomics,* 36, 649-660.

Wilson, J. R. & Norris, B. J. (2005b). Special issue on rail human factors. *Applied Ergonomics*, 36, 647-648.

Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*, 37, 60-82.

Woods, D. D. (1988). Coping with complexity: The psychology of human behaviour in complex systems. In L. P. Goodstein, H. B. Andersen, & S. E. Olsen (Eds.), *Tasks, errors, and mental models. A festschrift to celebrate the 60th birthday of Professor Jens Rasmussen* (pp. 128-148). London: Taylor & Francis.

Wootton, R. (1999). *Introduction to Telemedicine*. London: The Royal Society of Medicine Press limited.

Yin, R. K. (1994). *Case study research: Design and methods.* (2nd ed.) London: Sage Publications.

Zijlstra, F. R. H., & van Doorn, L. (1995). *The construction of a subjective effort scale* (Internal Report). Delft University of Technology: Department of Social Sciences and Philosophy.

Zsambok, C. E., & Klein, G. (1997). *Naturalistic decision-making*. New Jersey: Lawrence Erlbaum Associates, Publishers.

Telematica Instituut Fundamental Research Series

001 G. Henri ter Hofte, *Working apart together: Foundations for component groupware*

002 Peter J.H. Hinssen, *What difference does it make? The use of groupware in small groups*

003 Daan D. Velthausz, *Cost-effective network-based multimedia information retrieval*

004 Lidwien A.M.L. van de Wijngaert, *Matching media: Information need and new media choice*

005 Roger H.J. Demkes, COMET: *A comprehensive methodology for supporting telematics investment decisions*

006 Olaf Tettero, *Intrinsic information security: Embedding security issues in the design process of telematics systems*

007 Marike Hettinga, *Understanding evolutionary use of groupware*

008 Aart T. van Halteren, *Towards an adaptable QoS aware middleware for distributed objects*

009 Maarten Wegdam, *Dynamic reconfiguration and load distribution in component middleware*

010 Ingrid J. Mulder, *Understanding designers, designing for understanding*

011 Robert J.J. Slagter, *Dynamic groupware services: Modular design of tailorable groupware*

012 Nikolay K. Diakov, *Monitoring distributed object and component communication*

013 Cheun N. Chong, *Experiments in rights control expression and enforcement*

014 Cristian Hesselman, *Distribution of multimedia streams to mobile Internet users*

015 Giancarlo Guizzardi, *Ontological foundations for structural conceptual models*

016 Mark van Setten, *Supporting people in finding information: Hybrid recommender systems and goal-based structuring*

017 Remco Dijkman, *Consistency in Multi-viewpoint Architectural Design*

018 João P.A. Almeida, *Model-driven design of distributed applications*

019 Margit C.M. Biemans, *Cognition in Context: The effect of information and communication support on task performance of distributed professionals*

See also: http://www.telin.nl/publicaties/frs.htm