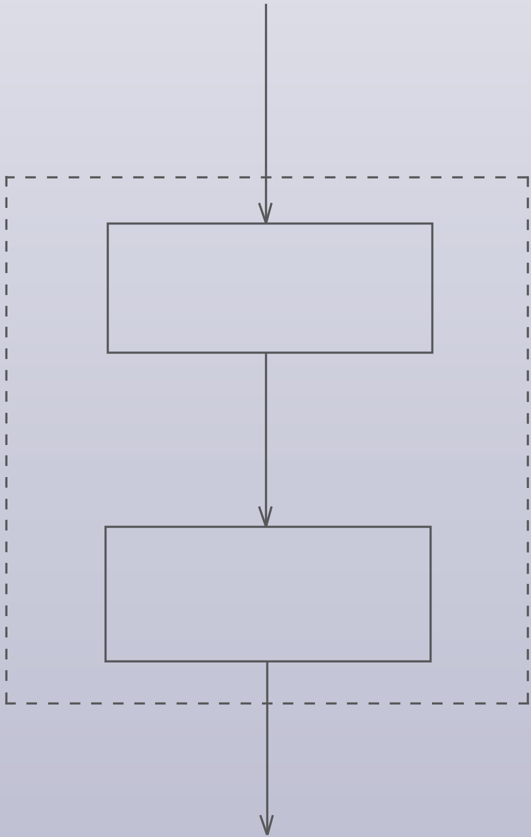
Smart Tailoring of Real-Time Physical Activity Coaching Systems



Harm op den Akker

SMART TAILORING OF REAL-TIME PHYSICAL ACTIVITY COACHING SYSTEMS

Harm op den Akker

Address of correspondence
Harm op den Akker
Roessingh Research and Development
PO Box 310
7500 AH Enschede
The Netherlands
h.opdenakker@rrd.nl

The publication of this thesis was financially supported by:

Roessingh Research and Development

Cover illustration: Rieks op den Akker & Bob Dickhoff (www.whataboutbob.nl)

Printed by: Gildeprint, Enschede

ISBN: 978-90-365-3762-9
DOI: 10.3990/1.9789036537629
ISSN: 1381-3617 (CTIT Ph.D. thesis Series No. 14-324)

CTIT

CTIT Ph.D. Thesis Series No. 14-324
Centre for Telematics and Information Technology
P.O. Box 217, 7500 AE
Enschede, the Netherlands

Thesis Number: / 200

SMART TAILORING OF REAL-TIME PHYSICAL ACTIVITY COACHING SYSTEMS

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op 17 oktober 2014 om 12:45 uur

door

Harm op den Akker
geboren op 20 februari 1985
te Enschede

Dit proefschrift is goedgekeurd door:

Prof. dr. ir. H.J. Hermens (promotor)
Dr. Valerie M. Jones (dagelijks begeleider)

Samenstelling promotiecommissie

Voorzitter/secretaris

Prof. Dr. P.M.G. Apers, Universiteit Twente

Promotor

Prof. Dr. H.J. Hermens, Universiteit Twente

Dagelijks begeleider

Dr. Valerie M. Jones, Universiteit Twente

Leden

Prof. Dr. D.K.J. Heylen, Universiteit Twente

Prof. Dr. M.M.R. Vollenbroek-Hutten, Universiteit Twente

Prof. Dr. P.J.M. Havinga, Universiteit Twente

Prof. Dr. B.J.A. Kröse, Universiteit van Amsterdam

Dr. Katarzyna Wac, Université de Genève

Contents

List of Figures

List of Tables

Chapter 1

General introduction

Chapter 1

Healthcare for the aging population

The burden on healthcare services in the Western world and developing countries is increasing due to an aging population and the associated increase in prevalence of chronic diseases. The current demographic changes that contribute to this problem have three major causes. The first major cause is that more people are surviving into old age, and the average age of death has increased. Life expectancies have increased yearly with a rate of 0.24 years over the past century in the United States, Europe and Japan (Christensen et al. 2009). The second cause is that there are relatively fewer and fewer newborns in the Western world. Simple and effective systems for birth control have been developed and gained widespread cultural acceptance, leading to a decrease in birth rates (Bös & Von Weizsäcker 1989). These two factors cause a change in the age distribution of the population, with a relatively larger proportion of the population over retirement age and needing care and economic support from the relatively dwindling population of working age. But there is a third factor currently at play; the baby boom generation is now reaching the age of retirement and an age at which many chronic conditions are becoming more prevalent. As a result, our society is dealing with a large population in need of chronic care, and a smaller workforce carrying the burden of maintaining this care. The growing demand on the healthcare system thus calls for treatments that are economically viable and reduce the demand on trained healthcare professionals.

Western society is turning towards technology to provide the solution. Information and Communication Technologies (ICT) to enable self-care for patients (eHealth) and technology that allows professional care to be delivered to patients in their home environment (telemedicine) form a fast growing field of research (Ekeland et al. 2010), and are widely regarded as a promising paradigm in reducing the cost of healthcare. A major focus in the fields of eHealth and telemedicine is the maintenance of a healthy lifestyle in terms of everyday physical activity. Physical *in*activity has an effect on a number of major non-communicable diseases and is estimated to cause 9% of premature mortality worldwide (Lee et al. 2012). If the physical inactivity factor could be eliminated completely, the world's life expectancy is estimated to rise by another 0.68 years. Fortunately, those extra life years are also estimated to be spent in better health, as an elimination of the physical inactivity factor could also lead to a 6% reduction of the burden of disease of coronary heart disease, 7% for type 2 diabetes and 10% for breast and colon cancers (Lee et al. 2012).

Physical activity plays an important role in the prevention of obesity, and chronic disease such as coronary heart disease, type 2 diabetes and breast and colon cancers (Lee et al. 2012). The American College of Sports Medicine recommends that the majority of adults perform moderate-intensity cardio respiratory exercise training for at least thirty minutes each day (Garber et al. 2011). Maintaining a healthy level of physical activity does not only play an important role in prevention. Physical activity often forms a major component of treatment programs for chronic conditions, motivating research into monitoring and promotion of physical activity in patients suffering from Chronic Low Back Pain (CLBP) (van Weering 2011), Chronic Fatigue Syndrome (CFS) (Evering 2013), and Chronic Obstructive Pulmonary Disease (COPD) (Tabak 2014). This Thesis is about physical activity.

Physical activity monitoring

The field of physical activity monitoring in humans is not new. An article by Kohl et al. (2000) reviews various techniques for assessing physical activity, citing articles that date back to 1971. Pioneering work in the field reported studies conducted with simple pedometers, self-reporting or direct observation. But the real rise of the field came with the introduction of accelerometry. A number of much cited papers by Bouten et al. describe the development of a tri-axial accelerometer based sensor for measuring daily physical activity (Bouten et al. 1997) and the validation of this technique by comparing the output of the sensor with energy expenditure measured with doubly labeled water (Bouten et al. 1996). In a 7-day free-living experiment with 30 subjects, body movement was registered with a tri-axial accelerometer and correlated to energy expenditure estimates such as doubly labeled water (Speakman 1998) — considered to be the gold standard in energy expenditure estimation. A 0.73 correlation factor between overall physical activity level and sensor outputs was found. Besides the societal benefits of research into physical activity, the popularity of eHealth and telemedicine applications that target daily physical activity behavior can also, at least partly, be attributed to such recent developments in monitoring technology.

Since then, 3D accelerometers have become smaller, cheaper, and ubiquitously available. Research in the field has also seen a surge in recent years with numerous reviews being published on sensor development, e.g. (Intille et al. 2012), assessment methodologies, e.g. (Kohl et al. 2000, Plasqui et al. 2013, Jones et al. 2013) as well as classification of specific human activities from sensor data, e.g. (Altun et al. 2010). The field has truly gained maturity with many commercial activity trackers that use accelerometers available on the market today. Even modern smartphones contain accelerometers and research shows their applicability in human movement recognition (Kwapisz et al. 2010). For the consumer market, the accelerometers were originally used to automatically rotate the screen when switching the device to portrait- or landscape view — but nowadays they are also used for tracking the movement of the owner and provide input to activity tracking applications such as *FitBit*[1] or *Samsung S Health*[2]. The field of activity sensing has shifted from using large, uncomfortable, technological devices to using small, unobtrusive and sometimes even fashionable lifestyle items, or a readily available consumer electronics such as mobile phones.

Current interventions and coaching

Miniaturized, wireless sensors have enabled monitoring of daily physical activity on a large scale. But although monitoring is an important first step in gaining insights into the daily levels of physical activity of various populations, it is not a goal in itself. The overall objective — and in a broad sense the objective of this research — is to stimulate people in achieving and maintaining a physically active lifestyle. Physical activity coaching applications and interventions aim to achieve this, in general by measuring the user's daily physical activity and providing feedback, coaching, health related information, and incentives to increase, balance, or maintain a certain *healthy* level of activity. Before the emergence of eHealth and telemedicine, health behavior change interventions were delivered through means such as printed material by sending informational leaflets through the mail (Dutton et al. 2008, Short et al. 2011, Noar et al. 2007) or telephone as a form of public service telemarketing (van Keulen et al. 2011, Marcus et al. 2007). Although

[1]fitbit.com
[2]developer.samsung.com/s-health-sdk

valuable lessons can be learned from these early, low-tech interventions, this Thesis focuses on more individualized approaches.

In this work, the focus is specifically on *real-time* physical activity coaching — applications that are able to provide direct, timely coaching to the individual, based on their *current* physical activity behavior. Smartphones are among the most appropriate devices for providing real-time coaching (as opposed to web portals, or PC applications). At Roessingh Research and Development, a large amount of effort has been put into developing real-time physical activity coaching interventions for various patient populations. The system used in these interventions consists of a 3D-accelerometer based activity sensor that connects wirelessly to a smartphone. The smartphone runs the coaching application that can show the user his current level of activity, and provide regular coaching in the form of short motivational messages. These efforts are documented in a number of publications and are summarized in three PhD dissertations on Chronic Low Back Pain (CLBP) (van Weering 2011), Chronic Fatigue Syndrome (CFS) (Evering 2013), and Chronic Obstructive Pulmonary Disease (COPD) (Tabak 2014). The work described in this Thesis can be seen as an effort to tackle some of the issues identified in these dissertations, particularly regarding long-term adherence and short-term compliance to the interventions.

Tailoring

Evaluation of the activity coach described above has shown its effectiveness in stimulating physical activity (Tabak et al. 2014, Dekker-van Weering et al. 2012). There are however two major issues that need to be solved. First, the increase in physical activity tends to diminish after several weeks, and second, the response to the motivational cues has a large variance across subjects and studies. The differing responses to the feedback mechanisms employed can be attributed to the high degree of heterogeneity within the target populations. It is widely believed that tailoring — or *personalization* — can help in increasing the effectiveness of technology that aims to achieve behavior change. In 2008, Hawkins et al. defined tailoring as "*any of a number of methods for creating communications individualized for their receivers, with the expectation that this individualization will lead to larger intended effects of these communications*" (Hawkins et al. 2008).

Since each individual is different, an intervention should therefore preferably use an approach that is tailored to this individual.

This Thesis is about tailoring real-time physical activity coaching systems. Modern sensors allow ubiquitous monitoring of the user's levels of physical activity. Smartphones can receive the data wirelessly and provide coaching to the user at any time and at any place. Smartphones are also becoming smarter, allowing the acquisition of rich contextual data through built in sensors and a constant internet connection. They also contain enough processing power to perform complex data analysis tasks. All of these technological advancements make it possible to do real-time analysis of the user, his context, and his activity behavior, enabling *smart tailoring of real-time physical activity coaching systems*.

Aim and outline of the Thesis

The four main sections of the Thesis each aim to answer a specific research question. These research questions and an outline of the Thesis is presented below.

> **RQ1:** What different aspects of tailoring can be identified that can be used to tailor real-time physical activity coaching systems? How can we model these aspects in order to facilitate design and assessment of tailored activity coaching systems?

This Thesis explores tailoring methods that can be used to increase the effectiveness of physical activity coaching applications. This exploration begins with an analysis of the literature aimed to answer RQ1. In **Chapter 2** an extensive survey of the literature on tailored real-time activity coaching systems is presented. Twelve different applications that each approach tailoring in various different ways are described. A total of seven different tailoring concepts have been identified and a model is developed that shows how each of the concepts relate to each other and how they can affect various properties of communication to the user. By making a distinction between a communication's timing, intention, content and representation, the model can be used to identify which tailoring techniques can be applied to which part of the communication.

> **RQ2:** Can compliance to motivational cues be increased through matching the timing of communication to the user and his context?

In **Chapter 3** the focus is on tailoring the timing aspect of communication. The Kairos system that is described is a self-learning, context aware system, that tailors the timing of motivational messages to individual users. The system constantly gathers information on the context of the user and uses machine learning techniques to predict the optimal timing of presenting a motivational cue. The theoretical foundations of the method, the design and implementation of the system, and the results of a longitudinal trial with 10 COPD patients are described.

> **RQ3:** How can a motivational message be defined, and how can individual message components be tailored?

Given the right timing for presenting the coaching to the user, the next step is to determine *what* to say. In **Chapter 4** a model of motivational messages is constructed, further developing the concepts of message timing, intention, content and representation, and showing practical examples of how tailoring techniques can be used to tailor each individual aspect. A practical framework is developed that can be used to generate motivational messages in a structured, sequential manner. Example implementations of tailoring rules are given, focusing on the message *intention* and *content* aspects.

> **RQ4:** What is the perceived effect of advanced representations of coaching mechanisms for office workers?

Chapter 5 deals with the last remaining aspect of communication: *representation*. In this chapter the focus lies on new ways of representing the coaching to the users. The chapter explores the concept of representation by describing the development and evaluation of two prototype applications. The first prototype demonstrates coaching in a multi-device setting, where the coach migrates with the user across different devices. The second prototype focuses on the use of an embodied agent using speech and facial animations to represent the *virtual coach* concept.

Finally, **Chapter 6** presents a summary of the contributions of the Thesis, including a discussion regarding the implications for future work, as well as a discussion on outstanding issues.

Bibliography

Altun, K., Barshan, B. & Tunçel, O. (2010), 'Comparative study on classifying human activities with miniature inertial and magnetic sensors', *Pattern Recognition* **43**(10), 3605–3620.

Bös, D. & Von Weizsäcker, R. K. (1989), 'Economic consequences of an aging population', *European Economic Review* **33**(2-3), 345–354.

Bouten, C. V. C., Koekkoek, K. T. M., Verduin, M., Kodde, R. & Janssen, J. D. (1997), 'A Triaxial Accelerometer and Portable Data Processing Unit for the Assessment of Daily Physical Activity', *IEEE Transactions on Biomedical Engineering* **44**(3), 136–147.

Bouten, C. V. C., Verboeket-Van De Venne, W. P. H. G., Westerterp, K. R., Verduin, M. & Janssen, J. D. (1996), 'Daily physical activity assessment: comparison between movement registration and doubly labeled water', *Journal of Applied Physiology* **81**(2), 1019–1026.

Christensen, K., Doblhammer, G., Rau, R. & Vaupel, J. W. (2009), 'Ageing populations: the challenges ahead', *Lancet* **374**(9696), 1196–1208.

Dekker-van Weering, M. G. H., Vollenbroek-Hutten, M. M. R. & Hermens, H. J. (2012), 'Do Personalized Feedback Messages about Activity Patterns Stimulate Patients with Chronic Low Back Pain to Change their Activity Behavior on a Short Term Notice?', *Applied Psychophysiology and Biofeedback* **37**(2), 81–9.

Dutton, G. R., Provost, B. C., Tan, F. & Smith, D. (2008), 'A tailored print-based physical activity intervention for patients with type 2 diabetes', *Preventive Medicine* **47**(4), 409–11.

Ekeland, A. G., Bowes, A. & Flottorp, S. (2010), 'Effectiveness of telemedicine: A systematic review of reviews', *International Journal of Medical Informatics* **79**(11), 736–771.

Evering, R. M. H. (2013), Ambulatory Feedback at Daily Physical Activity Patterns: A Treatment for the Chronic Fatigue Syndrome in the Home Environment?, PhD thesis, University of Twente.

Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I.-M., Nieman, D. C. & Swain, D. P. (2011), 'American College of Sports Medicine position stand. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise', *Medicine & Science in Sports & Exercise* **43**(7), 1334–1359.

Hawkins, R. P., Kreuter, M. W., Resnicow, K., Fishbein, M. & Dijkstra, A. (2008), 'Understanding tailoring in communicating about health', *Health Education Research* **23**(3), 454–466.

Intille, S. S., Lester, J., Sallis, J. F. & Duncan, G. (2012), 'New Horizons in Sensor Development', *Medicine & Science in Sports & Exercise* **44**(1 Suppl 1), S24–S31.

Jones, R. A., Hinkley, T., Okely, A. D. & Salmon, J. (2013), 'Tracking Physical Activity and Sedentary Behavior in Childhood: A Systematic Review', *American Journal of Preventive Medicine* **44**(6), 651–658.

Kohl, H. W., Fulton, J. E. & Caspersen, C. J. (2000), 'Assessment of Physical Activity among Children and Adolescents: A Review and Synthesis', *Preventive Medicine* **31**(2), 54–76.

Kwapisz, J. R., Weiss, G. M. & Moore, S. A. (2010), Activity Recognition using Cell Phone Accelerometers, *in* 'Proceedings of the 4th International Workshop on Knowledge Discovery from Sensor Data (SensorKDD-2010)', Washington, DC, USA, pp. 74–82.

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N. & Katzmarzyk, P. T. (2012), 'Impact of Physical Inactivity on the World's Major Non-Communicable Diseases', *Lancet* **380**(9838), 219–229.

Marcus, B. H., Napolitano, M. A., King, A. C., Lewis, B. A., Whiteley, J. A., Albrecht, A. E., Parisi, A. F., Bock, B. C., Pinto, B. M., Sciamanna, C. A., Jakicic, J. M. & Papandonatos, G. D. (2007), 'Examination of print and telephone channels for physical activity promotion: Rationale, design, and baseline data from Project STRIDE', *Contemporary Clinical Trials* **28**(1), 90–104.

Noar, S. M., Benac, C. N. & Harris, M. S. (2007), 'Does Tailoring Matter? Meta-Analytic Review of Tailored Print Health Behavior Change Interventions', *Psychological Bulletin* **133**(4), 673–693.

Plasqui, G., Bonomi, A. G. & Westerterp, K. R. (2013), 'Daily physical activity assessment with accelerometers: new insights and validation studies', *Obesity Reviews* **14**(6), 451–462.

Short, C. E., James, E. L., Plotnikoff, R. C. & Girgis, A. (2011), 'Efficacy of tailored-print interventions to promote physical activity: a systematic review of randomised trials', *The International Journal of Behavioral Nutrition and Physical Activity* **8**(1), 113.

Speakman, J. R. (1998), 'The history and theory of the doubly labeled water technique', *The American Journal of Clinical Nutrition* **68**(4), 932S–938S.

Tabak, M. (2014), Telemedicine for Patients with COPD - New Treatment Approaches to Improve Daily Activity Behaviour, PhD thesis, University of Twente.

Tabak, M., op den Akker, H. & Hermens, H. J. (2014), 'Motivational Cues as Real-Time Feedback for Changing Daily Activity Behavior of Patients with COPD', *Patient Education and Counseling* **94**(3), 372–378.

van Keulen, H. M., Mesters, I., Ausems, M., van Breukelen, G., Campbell, M. K., Resnicow, K., Brug, J. & de Vries, H. (2011), 'Tailored Print Communication and Telephone Motivational Interviewing Are Equally Successful in Improving Multiple Lifestyle Behaviors in a Randomized Controlled Trial', *Annals of Behavioral Medicine* **41**(1), 104–118.

van Weering, M. G. (2011), Towards a New Treatment for Chronic Low Back Pain Patients - Using Activity Monitoring and Personalized Feedback, PhD thesis, University of Twente.

Chapter 2

Tailoring Real-Time Physical Activity Coaching Systems: A Literature Survey and Model

Harm op den Akker, Valerie M. Jones and Hermie J. Hermens, in *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change, 2014*. DOI:10.1007/s11257-014-9146-y

Abstract

Technology mediated healthcare services designed to stimulate patients' self-efficacy are widely regarded as a promising paradigm to reduce the burden on the healthcare system. The promotion of healthy, active living is a topic of growing interest in research and business. Recent advances in wireless sensor technology and the widespread availability of smartphones have made it possible to monitor and coach users continuously during daily life activities. Physical activity monitoring systems are frequently designed for use over long periods of time placing usability, acceptance and effectiveness in terms of compliance high on the list of design priorities to achieve sustainable behavioral change. Tailoring, or the process of adjusting the system's behavior to individuals in a specific context, is an emerging topic of interest within the field. In this article we report a survey of tailoring techniques currently employed in state of the art real time physical activity coaching systems. We present a survey of state of the art activity coaching systems as well as a conceptual framework which identifies seven important tailoring concepts that are currently in use and how they relate to each other. A detailed analysis of current use of tailoring techniques in real time physical activity coaching applications is presented. According to the literature, tailoring is currently used only sparsely in this field. We underline the need to increase adoption of tailoring methods that are based on available theories, and call for innovative evaluation methods to demonstrate the effectiveness of individual tailoring approaches.

2.1 Introduction

The prevalence of chronic diseases is increasing world wide, largely due to demographic changes. The growing demand on healthcare services calls for cost-effective treatments that reduce the demands on healthcare professionals. Provision of eHealth and telemedicine services, in particular technology mediated services which stimulate and support patient's self-efficacy, is a fast growing field of research (Ekeland et al. 2010) and is widely regarded as a promising paradigm to reduce the burden on the healthcare system. An important factor in prevention and treatment of chronic disease and supporting healthy ageing is maintaining a healthy lifestyle in terms of regular physical activity. The American College of Sports Medicine recommends that the majority of adults perform moderate-intensity cardio respiratory exercise training for at least thirty minutes each day (Garber et al. 2011). Monitoring of physical activity and development of eHealth and telemedicine systems to motivate individuals to reach personal activity targets is a large and growing field of research and development in its own right.

2.1.1 Physical activity monitoring

Research into accurate assessment of physical activity levels has been conducted for many decades. Kohl et al. (2000) reviews various assessment techniques, citing articles dating back to 1971, and classifies techniques into six categories, including *self-report* (Sallis & Saelens 2000), *direct observation* (McKenzie 2002), *indirect-, and direct calorimetry* (Bailey et al. 1995), *doubly labeled water* (Speakman 1998) and *electronic or mechanical monitoring* using e.g. pedometers (Saris & Binkhorst 1977, Lutes & Steinbaugh 2010) and accelerometry (Bouten et al. 1997, Plasqui et al. 2013). Activity monitoring tools in this last category have seen a surge in recent years due to their low cost and unobtrusive applicability.

The popularity of low-cost, accelerometer based activity monitoring tools becomes apparent when looking at the wide range of commercially available systems and services (Table 2.1). These commercial fitness applications, designed for use throughout the day, measure physical activity and provide feedback on performance and/or progress to the user in various ways. The products vary mainly on two points: the location where the sensor is worn, and how measured activity data is fed back to the user. This feedback is either visualized on the sensor,

#	**Name**	**Worn**	**SE**	**SP**	**WP**	**PC**	**Website**
1	Polar Active	W	+		+		polarusa.com
2	ActivPal	H			+	+	paltech.plus.com
3	GENEActive	W	+			+	geneactiv.co.uk
4	MoveMonitor	H				+	mcroberts.nl
5	APDM	H A W C				+	apdm.com
6	ActiGraph	H A W C				+	theactigraph.com
7	Philips DirectLife	H C P	+		+		directlife.philips.com
8	FitBit	H P C W	+	+	+		fitbit.com
9	Acti Smile	H C	+				actismile.ch
10	BodyBugg	U	+	+	+		bodybugg.com
11	BodyMedia FIT	W		+		+	bodymedia.com
12	Jawbone UP	W		+			jawbone.com/up/
13	Nike+ Fuelband	W	+	+	+		nikeplus.nike.com

Table 2.1: Representative sample of commercially available accelerometry based activity monitoring systems. The 'Worn' column indicates where the sensor can be worn: on the Ankle (**A**), Chest (**C**), Hip (**H**), Upper arm (**U**), Wrist (**W**) or in the Pocket (**P**). The remaining columns indicate whether the product gives feedback on the sensor (**SE**), smartphone (**SP**), web portal (**WP**) or through a PC program (**PC**).

smartphone, web portal, or PC application. A number of commercially available accelerometry-based activity monitoring and feedback applications are shown in Table 2.1[1].

The feedback modalities used differ in terms of timeliness and richness as visualized in Figure 2.1, with feedback on the sensor being the fastest and least rich, and PC applications capable of the richest, but slowest, feedback. Sensors capable of displaying feedback can do so without any significant delay, however the feedback is simplistic due to the limitations of a small display and low processing power. Smartphones offer more possibilities in terms of screen size and processing, and if a sensor is connected (wirelessly) to a phone (e.g. *Nike+ Fuelband*) feedback via the phone can be presented in real-time. Web portals introduce a delay in feedback since they operate on synchronized data over the internet, however they

[1] This selection (gathered in June, 2012) does not represent an exhaustive list but serves as a representative sample to illustrate the range of available commercial products.

offer a potentially richer experience due to the ability to provide full-screen data visualization. Furthermore they can be accessed from any location. Currently available PC applications usually require the sensor to be physically near a specific computer running specific software. This approach offers the least timely modality for providing feedback. However, the available processing power and screen size make it potentially the richest way of doing so, enabling for example the use of realistically rendered virtual human avatars.

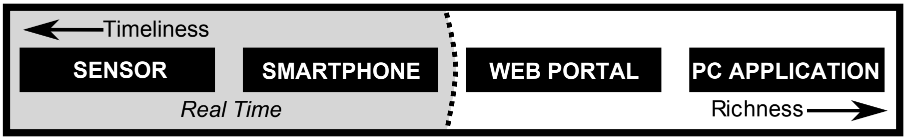


Figure 2.1: Timeliness versus Richness of feedback modalities. Modalities to the left are increasingly readily available, while modalities to the right are increasingly rich in their capabilities of providing feedback. Sensor and smartphone are real time feedback modalities.

2.1.2 Background

A vast range of systems for daily activity monitoring are available; a complete review of these is not the purpose of this article. Rather we focus on the techniques employed in these systems which aim to motivate the user to reach personal activity related goals. We have already shown the diversity in the way different systems provide feedback to the user, ranging from real-time approaches through a display on the sensor or a direct connection to a smartphone, to *offline* methods using e.g. web portals or PC applications. However feedback is only one of the possible ways of stimulating the user to change his activity behavior through generating awareness of current behavior. There are other motivational strategies and different ways of ensuring that the system's goal of changing the user's behavior is understood and followed by the user. Since each individual is different, an intervention should preferably use an approach which matches the individual user. This is better known as **tailoring**, and it is the topic of this survey.

The field of tailoring — also referred to as personalization, or individualization — is broad, even when limited to behavior change interventions. Long before the emergence of eHealth and telemedicine, tailoring techniques were applied to print-based health behavior change interventions. These 'tailored' interventions could range from a simple leaflet that is addressed to an individual by name — i.e. personalized generic communication (Kreuter et al. 1999), a message developed specifically for a certain target group to messages targeted at the individual user. The effects of tailoring in such print-based interventions have been well documented, and the overall merit of this approach is clear (see e.g. (Noar et al. 2007) for a meta-review on tailored print).

The use of printed material however, severely limits the ability to effectively target individual users, a shortcoming that is alleviated by the application of web-based interventions. In a study to promote a healthy lifestyle, targeted at families, Colineau & Paris (2011) used an online webportal to provide family-based goals and tailored feedback to encourage families to submit ideas for improving the family's lifestyle. The feedback messages provided by this online platform are tailored specifically to the situation of the family, or of individual family members through an automatic message composition system. The use of such computerized interventions allows more advanced tailoring, providing more personal and varied motivational messages to the user.

Morandi & Serafin (2007) use various parameters such as stage of change, day of the week, and specific user preferences to tailor motivational messages to the user. Others use more advanced models based on e.g. ontologies (Erriquez & Grasso 2008) possibly in combination with user clustering (Cortellese et al. 2009). A method for clustering users of a health-related persuasive gaming application is discussed in (Orji et al. 2014) in this issue, showing which design strategies fit best to which gamer type. To facilitate personal communication with the user, the field of tailored behavior change interventions is highly related to and benefits greatly from the work done on user modeling, a field that addresses the issue of communicating the *right* thing at the *right* time in the *right* way for each individual user of a system (Fischer 2001). In particular, smartphone based user modeling frameworks such as described in Gerber et al. (2010) offer promising benefits for real time tailored coaching applications, as smartphones offer the best potential for providing rich, real-time coaching (see Figure 2.1).

2.1.3 Scope and Goals

The body of literature on tailoring, as well as on activity monitoring and coaching is extensive, and as stated earlier beyond the scope of this survey. We focus here on the following two key aspects:

1. **Physical activity coaching:** we include only applications which aim to motivate users to change their activity behavior by means of a coaching element.

2. **Real time tailoring:** we include only those systems that offer real-time coaching, and use some form of tailoring to adjust the application's communication to individual users.

The goals of this article are twofold: (**I**) to provide a comprehensive survey of the literature that describes tools and techniques currently used for tailoring in real-time coaching applications for physical activity; (**II**) to define a conceptual framework of these techniques, extending current models of tailoring, that can help guide the design of new tools and applications in this field.

This article is not a systematic review in the sense that we do not aim to derive statistical evidence or conclusions from existing literature. Due to the scoping of the article it is also clear that we neither capture the full body of work on tailoring, nor on physical activity promoting applications. Instead we aim to provide an exploratory and more detailed analysis of the more narrow cross-section between these two fields. Emerging ubiquitous technology brings the promise of continuous coaching that goes beyond e.g. daily or weekly summaries of performance. Smartphones and other intelligent devices have the ability to reach their users at opportune moments throughout the day, and as such are able to provide intensive coaching. This ability has great merit for behavior change systems and as such we add a specific focus on such emerging technologies that offer *real-time* tailored activity coaching.

Various reviews, complementary to the current work, have been published in recent years and will be discussed in Section 2.2. The literature search process is described in Section 2.3, after which an overview of included papers is given in Section 2.4. We then present our model of tailoring (Section 2.5) before providing an analysis of the tailoring concepts (Section 2.6), finishing with a discussion (Section 2.7) and conclusions (Section 2.8).

2.2 Related work

In recent years, a number of scientific review papers were published on various topics related to the theme of tailored real-time coaching for physical activity. In 2008, Lustria et al. (2008) reviewed 30 computer-tailored health interventions delivered over the web, seven of which in the domain of physical activity. Although the modalities (or delivery methods) used for the interventions were not real time, there were some interesting findings in the various tailoring mechanisms adopted. The authors distinguish between three types of tailoring: personalization, feedback, and adaptation (or 'content matching'). These types originate from a 2008 paper by Hawkins et al. (2008) who systematically defined the constructs that encompass tailoring. A detailed description of tailoring concepts, including the ones defined by Hawkins is given in Section 2.5.2. All of the included physical activity interventions in Lustria's review adopted a combination of two or all of these tailoring mechanisms. Overall, the authors conclude that computer-tailored online interventions vary greatly in their strategies to provide users with tailored messages, but also in their delivery mode and -timing, and overall levels of sophistication.

A systematic review published in 2009 by Fry & Neff (2009) focuses specifically on periodic prompts that encourage healthy behavior (including physical activity) and shows the effectiveness of daily, weekly, or monthly messages, reminders or brief feedback in limited contact interventions. The review includes 19 articles, published between 1988 and 2008, 13 of which used email as the medium for sending prompts, and 14 of which employed some form of message tailoring. One of the aspects evaluated was the timing of messages. Most of the included studies used weekly prompts to encourage healthy behavior change, and in one of the studies it was shown that weekly (telephone) prompts performed significantly better in encouraging physical activity than prompts sent every three weeks (Lombard et al. 1995). The authors state that the question remains how prompts issued every day (a situation that would better approach the real-time feedback that we are interested in) would affect behavior change, because such a frequency was not found in any of the included studies.

A review by Enwald & Huotari (2010) on second generation tailored health communication for the prevention of obesity conclude on this regard that "*mobile devices can help to achieve 'kairos', that is, the opportune moment to persuade...*". Real-time feedback through the use of mobile phones thus seems to be a promising, but underexplored field of research. In general, Enwald & Huotari (2010) found in their review of 23 studies, from which seven targeted physical activity, negative or mixed results regarding the effectiveness of tailoring, which according to the authors, is in line with previous studies. A particular focus of interest in their review is the use of theories and models of health behavior change to guide intervention design. Out of the included 23 studies, 14 used Prochaska's Transtheoretical-, or Stages of Change (TTM/SoC) model, including five of the seven physical activity related studies. The SoC model claims that people attempting to change can be categorized as being in one of five stages of the change process: precontemplation, contemplation, preparation, action, and maintenance (Prochaska & Velicer 1997). The theory is often used in physical activity interventions as one of the tailoring criteria, e.g. people in different stages need different motivation or prompts, but the effectiveness of using this model in physical activity interventions is contested (Adams & White 2005). The 'controversy' of using the TTM in the domain of physical activity promotion is again underpinned by a review on tailored print communication by Short et al. (2011) who state that their findings support earlier claims that studies using Social Cognitive Theory (Bandura 1986) or The Theory of Planned Behavior (Ajzen 1991) demonstrate more positive effects in increasing physical activity levels than those applying the theories of the TTM.

The reviews mentioned earlier are relevant to the current work insofar that they deal with tailoring for healthy behavior, including physical activity, but none of them have a focus on (or even mention) real-time technologies. In contrast, a review by Kennedy et al. (2012) deals specifically with what the authors call "active assistance technology", defined as "*any technology involving automated processing* *of health or behavior change information that is ongoing as the user interacts with the technology*". One of the four technology roles selected for inclusion of articles is defined as "*dynamic adaptive tailoring of messages depending on context*", which has potential overlap with the topic of real-time, tailored feedback. The authors found widespread use of dialog systems as active technology (19 out of 41 included studies), out of which eight employed embodied conversational agents as interface to the user. Overall they found that dynamic tailoring was not a major topic in most included studies and concluded that the potential of active technologies for dynamic information processing is currently not fully exploited. The authors stress the need for interdisciplinary collaboration between behavior change researchers and researchers from computer science and cognitive science, but do not provide clear recommendations on the use of dynamic tailoring.

2.3 Search strategy, inclusion-, and exclusion criteria

The systematic literature search was carried out in two phases. An initial systematic search was done in July 2012 (phase one). Due to the limited amount of included articles, a second phase search was performed in August 2013 (phase two).

To capture the literature relevant for the scope and goals (Section 2.1.3) of this survey we formulated our search query to find articles related to "[*personalized*] [*activity*] [*coaching*]". The search terms used can be found in Table 2.2, where all **<u>bold-underlined</u>** terms were added only for the second phase. In order to cover both the health- and technology domains, we performed the initial search on PubMed (www.pubmed.com, 582 results) as well as the ACM Digital Library (dl.acm.org, 623 results). Additionally we included a total of 116 results from a manual search through Google Scholar (scholar.google.com) and our personal libraries, obtaining a total initial set of 1.321 papers. We performed an initial filtering of results by removing duplicates and by looking at titles and abstracts, eliminating all papers that were not primarily about everyday physical activity, novel coaching tools or applications, or methods regarding motivational coaching. This first filtering resulted in a set of 320 papers for which full text articles were

Topic	Terms
personalized	personalized OR personalised OR personalization OR personalisation OR individualized OR individualised OR individualization OR individualisation OR tailored OR tailoring
activity	physical activity OR daily activity OR **walking** OR **exercise** OR **exercising** OR **activities of daily living**
coaching	coach OR coaching OR feedback OR motivate OR motivation OR stimulate OR stimulation OR promote OR promotion
application	**application** OR **system** OR **device**

Table 2.2: Search terms used in the systematic literature search phase. Target literature has to contain at least one of the listed terms from each of the three topics. The **bold-underlined** terms were added in the second literature search phase.

retrieved. When analyzing the full-text articles, we excluded any work that did not describe an application containing real-time communication with its user. Also, papers that targeted very specific physical activities — (e.g. gaming, exercises) instead of regular daily activities — were excluded. The specific inclusion-, and exclusion criteria are listed in Table 2.3. We found a total of **13 papers**, describing **11 applications** to be included in the survey. We also found a total of **40 "background papers"** describing relevant (real-time) tailoring techniques that did not describe its use in a specific application. These papers are used throughout the survey to serve as examples or background literature where relevant.

For the second phase literature search we modified our search query in two ways. First we extended our definition of [*activity*], and second — based on the experience from the first phase of literature search — we limited our search to [*application*] oriented papers. The extended search terms are presented in **bold-underlined** in Table 2.2. The second phase search was carried out in August 2013. The systematic search was carried out on the PubMed and the ACM digital library archives. From these two sources we found a combined total of 986 results, out of which 361 were new compared to the phase one search results. By examining the titles and abstracts we filtered out any articles that were either not written in English, review articles, and articles not targeted at daily physical activity — resulting in a set of 85 potentially relevant new articles. We performed the same filtering on the archives of the User Modeling and User-Adapted Interaction Journal as well as the proceedings of the past conferences on User Modeling, Adaptation and Personalization (UMAP 2009-2012), resulting in an additional 42 articles. For the 127 potentially relevant titles, full-text articles were retrieved and the inclusion of papers was again based on the inclusion-, and exclusion criteria as described in Table 2.3. From the second phase search, an additional **two papers**, describing **two applications** were included. In addition, we added **10 "background papers"** to the body of literature to be used as reference material.

In total, we included **12 applications**, described by 15 different papers found in the literature search. In order to adequately analyze and describe these applications, we performed an additional search targeting the selected applications by searching for the application *name* as well as for additional publications by the authors. This search yielded an additional 36 papers that are used to summarize the works in Section 2.4 below.

Topic	In./Ex.	Criteria
Activity	In.	Describes an application targeting promotion of everyday physical activities including e.g. walking or jogging preferably to be used throughout the day.
	Ex.	Targeted at specific (rehabilitation) exercises, exergames, or specific training of functions/skills.
Real-time	In.	Applications that are able to communicate constantly with the user and can give immediate feedback on measured performance.
	Ex.	Web-based applications (where users can only be targeted after login) or applications that do not have a direct connection between sensor and feedback device.
Tailoring	In.	Applications that use some form of tailoring/personalization.
	Ex.	Applications that do not change their behavior or usage for different individual- or groups of users.

Table 2.3: Listing of the inclusion- and exclusion criteria as used for the selection of final papers in phase one and two of the systematic search process.

Sec.	Application	Reference(s)
2.4.1	The Mobile Personal Trainer	(Buttussi et al. 2006)
		(Buttussi & Chittaro 2008)
2.4.2	MPTrain and TripleBeat	(Oliver & Flores-Mangas 2006)
		(de Oliveira & Oliver 2008)
2.4.3	UbiFit Garden	(Consolvo et al. 2008)
2.4.4	NEAT-o-Games	(Fujiki et al. 2008)
2.4.5	The Mobile Fitness Companion	(Ståhl et al. 2008)
2.4.6	Handheld Exercise Agent	(Bickmore et al. 2009)
2.4.7	Haptic Personal Trainer	(Qian et al. 2010)
		(Qian et al. 2011)
2.4.8	Everywhere Run	(Mulas et al. 2011)
2.4.9	Move2Play	(Bielik et al. 2012)
2.4.10	ActivMON	(Burns et al. 2012)
2.4.11	BeWell+	(Lin et al. 2012)
2.4.12	Analytic, Social, Affect	(King et al. 2013)

Table 2.4: Listing of the 12 included applications and their corresponding papers that were found during the literature search. The order is in ascending year of publication and author name.

2.4 Descriptions of included applications

In this section we will present an overview and summaries of the 12 included applications (see Table 2.4). The summaries focus on (1) functionalities, (2) theoretical foundations, (3) employed tailoring concepts, and (4) algorithmic approaches. For each of the applications we address these four topics where sufficient details are provided from the literature. The analysis of these papers and the creation of the tailoring model described later was done in an iterative way. On the one hand, the concepts and model have been derived from the analysis of the literature, while on the other hand the same literature is later described using the model as framework for structured analysis. After the overview of papers presented here, we will first describe the key concepts and our model of tailoring in Section 2.5. Then, in Section 2.6, we give a detailed analysis of the various tailoring concepts by looking at how they are employed in included applications.

2.4.1 The Mobile Personal Trainer

The Mobile Personal Trainer (MOPET) described in (Buttussi et al. 2006) and (Buttussi & Chittaro 2008) is an embodied virtual trainer that can guide users through outdoor fitness trails (see Figure 2.2.a at the end of this section). The embodied coach, Evita, runs on a smartphone and provides audio navigation, audio and graphical feedback about performance and animated 3D demonstrations of exercises along the trail. The system uses GPS to track the user's position and encourages users to keep up with certain speeds at regular checkpoints using speech synthesis. The 2008 update of the MOPET system describes additional tailoring approaches like context-aware and user-targeting features. The system integrates a user model containing gender, age, weight, height, physiological parameters derived from a guided auto-test, as well as historical information regarding previously completed trails and exercises. The sensed context (location, speed) combined with information from the user model is used to recommend exercises and provide alerts regarding e.g. speed of jogging.

The MOPET user model is initialized at the application's first launch by requiring the user to manually input gender, age, weight and height. Subsequently the user is asked to perform the guided auto-test which consists of stepping on and off a step. The autotest implements an algorithm that estimates the user's maximum volume of oxygen uptake per minute (VO_2Max) based on the user's *Power*, heart rate and some gender/age specific constants. The user's Power calculation starts when the user's heart rate is within a predefined (age-specific) range and is based on weight, the height of the step (as indicated by the user) and the time taken per step. Additional user model information is obtained by storing the number of times a user has completed a certain exercise within-, above- or below the user's heart rate thresholds as well as the number of times the exercise was abandoned prematurely.

The Power and VO_2Max values can be updated by performing subsequent autotests, but are currently only used as an indication of the user's physical condition, and not for further tailoring of the system. Instead, the user's calculated heart rate thresholds and recorded previous experiences are used by the exercise recommender module. When recommending a strengthening exercise, the module keeps track of the current *level* of performance, starting at the beginner's level (fewer repetitions, slower pace). After each completion of an exercise the user model is updated with the performance, and subsequent recommendations of the same exercise can take this history into account by recommending increased repetitions or pace.

As future tailoring capabilities of the system, the authors mention the option of setting specific goals related to e.g. weight loss, cardiovascular training or muscle strength to further guide the recommendations of exercises. In a separate paper, the authors describe the automatic creation of a user generated fitness trail database that can be used in the MOPET system (Buttussi et al. 2009). By storing user-preferences in the trail database, collaborative filtering based algorithms can be used to recommend fitness trails to users of the MOPET system based on personal preferences as well as physical ability. The literature regarding the MOPET system does not give any details on whether any specific theories of behavior change are used as a basis of the tailoring features.

2.4.2 MPTrain and TripleBeat

The MPTrain system described in (Oliver & Flores-Mangas 2006) is a mobile phone based system that uses automatic music selection to encourage the user to reach his/her exercise goals. The system consists of a set of physiological sensors (accelerometer, ECG) connected wirelessly to a mobile phone. The system implements a learning algorithm that automatically determines a mapping between musical features (volume, beat and energy), the user's current exercise level and the user's heart-rate response. While jogging, before the end of the current song, the algorithm determines whether the user needs to speed up, slow down or keep his pace, based on the user's current heartbeat compared to a predefined goal. MPTrain will then select the next song whose tempo is similar, faster or slower than the current song according to the difference between the actual and desired heart rates. The authors have chosen music as a feedback modality based on the theory that music improves gait regularity due to the beat helping individuals to anticipate the rate of movement, and as such causing the body and the music to get synchronized.

An update to the MPTrain system is presented in (de Oliveira & Oliver 2008), dubbed TripleBeat. The TripleBeat system adds two important tailoring approaches: a virtual competition with other runners and a glanceable real-time user interface. The authors aim to increase the motivational effect of the system based on the behaviour change theories described in Fogg (2003) as well as on empirical demonstrations from related research. The authors specifically mention four persuasive strategies, taken from Fogg (2003), as motivation for their work: (1) providing *personal awareness* through feedback on current physiological and activity data — for which there are many examples; (2) leveraging *social factors* through providing real-time information about the performance of other users based on the work of (Maitland et al. 2006, O'Brien & Mueller 2007, Sohn & Lee 2007); (3) providing *enjoyable interaction* by e.g. the use of an appealing 3D virtual trainer as in the MOPET system (see Section 2.4.1) or through the use of virtual game environments (Mokka et al. 2003); and (4) *unobtrusive/intuitive notification*, stressing the need to provide relevant information without interrupting or disturbing the user.

In TripleBeat these persuasive strategies are implemented as follows. The *personal awareness* is created by allowing users to monitor heart rate and pace in real-time, as well as by providing real-time feedback on how to achieve specific workout goals. A more interesting and defining feature is the TripleBeat's implementation of *social factors*. Users can hold a virtual race with virtual runners — other runners who have previously completed a run or the user himself. In order to promote fair competition, the system can automatically match the user with a competitor based on similarity in how well they achieve their goals. This is done using a variation of the k-nearest neighbour algorithm on the *score-vectors* of the user and his potential opponents. In order to provide a challenge, there will always be at least one opponent whose score is higher than that of the user. The *enjoyable interaction* of Triplebeat is 'inherently' delivered through its musical feedback; and its *unobtrusive notifications* are provided through a glanceable user interface that provides simple and clear feedback on current performance. Additional information regarding the basis of the MPTrain/TripleBeat system — the automatic generation of music playlists — can be found in (Oliver & Kreger-stickles 2006*b*), and details regarding the evaluations of the system can be found in (Oliver & Kreger-Stickles 2006*a*).

2.4.3 UbiFit Garden

Consolvo et al. (2008) describe the UbiFit Garden, a mobile phone application that uses a glanceable display of a flowering garden to create awareness of and stimulate regular physical activity (see Figure 2.2.b). The system includes a separate sensor for measuring physical activity. The 'Mobile Sensing Platform' activity sensor takes data from a 3D accelerometer and a barometer and uses boosted decision stump classifiers to distinguish between various types of activities (e.g. walking, running, cycling) four times per second — described in more detail in (Choudhury et al. 2008). This data is sent to the mobile phone over bluetooth where a smoothing algorithm is used to identify longer bouts of activities. In the case that performed activities are not recognized correctly, the user has the option of adding, editing or removing activities in a manual journal feature.

The defining feature of the UbiFit system is its glanceable mobile phone display. The display is implemented as a background image of the phone and represents the user's activity as a flowering garden. Each of the various types of detected activities are represented as a different type/color of flower. These different types of flowers represent different types of activities as recommended by the ACSM for a balanced physically active lifestyle: cardio, resistance training, flexibility, and walking. A specific tailoring approach that is implemented is a feature that allows users to set their own weekly goals in terms of activity types. Upon reaching their weekly goals, a large yellow butterfly appears on the glanceable display. Initial evaluations focused mainly on the activity detection component of the system, but a follow up three-month experiment was conducted later (Klasnja et al. 2009).

In Consolvo et al. (2008) the authors mention that the UbiFit garden application is targeted specifically at users in the contemplation, preparation and action stages of change of the Transtheoretical Model (Prochaska & DiClemente 1986), although it is unclear which specific design decisions were made based on this consideration. The application does not seem to use the user's current stage of change to tailor any specific form of motivational support. In (Consolvo, McDonald & Landay 2009), the authors give background on some of the behaviour change theories behind the UbiFit garden system. For example, the ability for users to set their own goal is based on the Goal-Setting Theory by Locke & Latham (2002), with the specific reason for self-setting of goals being that *...the* *individual needs to have decided that the goal is important to her... rather than being assigned to her with no rationale*. More details regarding the goal setting design decisions can be found in (Consolvo, Klasnja, McDonald & Landay 2009). Furthermore, the authors draw especially from two major psychological works: *Presentation of Self in Everyday Life* (Goffmann 1959) and *Cognitive Dissonance Theory* (Festinger 1957). A detailed description of how these theories can be applied in the design of systems supporting behavioral change can be found in (Consolvo, McDonald & Landay 2009).

2.4.4 NEAT-o-Games

Similar to the TripleBeat's virtual competition system described above, the NEAT-o-Games system by (Fujiki et al. 2008) implements a virtual race as motivator for physical activity. The system's main purpose is to stimulate NEAT — *non-exercise activity thermogenesis* — or daily physical activity, by turning daily life into a game. An activity sensor measures daily physical activity and sends data to a smartphone over Bluetooth. An algorithm on the phone derives "activity points" from the measured movements, which propels the user forward in a virtual race with a networked buddy list, where a winner is declared every day.

The work is motivated by the idea that *strong motivation* and *ubiquity* are the two key drivers for everyday activity behavior change. The motivational aspect is tackled by opting for a gaming approach, while the ubiquity issue is dealt with by not requiring the full attention of the 'player' all the time. In principle, the user is playing throughout the day, by having all of his movements captured and transformed into "activity points". Primarily, these activity points are used to propel players forward in a virtual race with other players, while points can also be spent on hints in a cognitive game (Sudoku). The authors mention a set of four design principles used in the development of the system: *simple*, *informative*, *discreet*, and *motivating*. However, other than background on serious gaming for promoting physical activity, the authors do not seem to draw on any scientific theories of behavior change or tailoring approaches. The distinctive feature of the NEAT-o-Games system is the virtual race to provide motivation through competition. In an earlier pilot study it was shown that the addition of a computerized avatar increased mean activity of the user, and the further addition of a real human opponent increased activity even further (Fujiki et al. 2007). The system does not allow for automatic selection of opponents and requires a manual partnering with a buddy.

The system was evaluated in a short pilot study with eight participants, as well as a four-week field trial with 10 included participants. The authors state that the gaming paradigm employed appeared to be effective as the participants who were classified as ‘consistent users’ reported higher activity levels than other participants.

2.4.5 The Mobile Fitness Companion

The Mobile Fitness Companion described in (Ståhl et al. 2008) consist of a conversational agent running on a smartphone and a stationary system in the user’s home in the form of a Nabaztag rabbit. The mobile interface shows an image of the same rabbit to create a feeling of persistence. The focus of this work is on the creation of a ‘companion’ as well as a natural language interface using automatic speech recognition (ASR) and text-to-speech (TTS).

The context of the work is the execution of fitness tasks. The system uses GPS to track the user throughout the day, deriving distance, pace, duration and calories burned during physical activities. The companion can keep track of a personal user plan and can suggest tasks for the user to perform based on the time of day and the user’s current location (determined by GPS). The user can accept the suggestion or initiate a dialogue with the system to suggest a different exercise (e.g. walking). The technical details of the Mobile Fitness Companion mostly concern the implementation of the ASR and TTS algorithms and functionalities through a client (mobile) and server architecture. As such, the authors do not provide any background regarding theories of behavior change or tailoring aspects (goal setting and context awareness) employed in the system. Details on the evaluation strategies can be found in (Benyon et al. 2008) and further details including evaluation results are presented in (Turunen et al. 2011).

2.4.6 Handheld Exercise Agent

In order to promote physical activity (Bickmore et al. 2009) describes the Handheld Exercise Agent. The authors developed a “general purpose health counseling agent interface" for use on smartphones. The interface consists of an animated agent that can display nonverbal conversational behaviors and produce text balloon output that is synchronized with lip movements. This Embodied Conversational Agent (ECA) does not produce auditory speech due to privacy concerns. The article focuses on a pilot study to assess social bonding between user and agent in the application domain of exercise promotion. The study consists of two conditions: the AWARE condition, in which the agent automatically recognizes walking activities and the NON-AWARE condition in which the user has to explicitly tell the system when walking activities were performed. The agent would provide positive reinforcement after walking bouts of 10 minutes or longer, or provide neutral comments after shorter activity bouts. Also, in the AWARE condition, the user could request the total number of steps walked since midnight as additional feedback. The user interactions in the system are limited to multiple-choice options, and possible interactions are scripted in an XML-based hierarchical state-transition network. The scripts consist of agent utterances in plain text as well as specifications for transitions to different states based on user selection or sensed information (physical activity). Scripts are pre-processed using a text-to-embodied-speech engine as well as a viseme (visual phoneme) generator before being installed on the mobile device.

As a motivation for the work on embodied agents, the authors highlight the need for effective health behavior change interventions to deliver tailored motivational and informational messages based on the context of the user and his user characteristics such as motivational readiness (i.e. stage of change), past behavior, ethnicity and age. The complexity and nuance required to deliver this type of communication is perhaps most effectively delivered through technologies that come closest to the "gold standard" of one-on-one face-to-face counseling. The authors also highlight, from the literature, the importance of health provider empathy and the quality of the provider-patient working relationship in improving patient satisfaction, adherence and health outcomes. Context awareness is another important factor, as it provides the agent with the ability to proactively intervene in certain circumstances, increasing also perceptions of familiarity, common ground, solidarity and intimacy. However, the only "context aware" information currently employed in the system is the detection of bouts of walking. Overall it is hard to discover how certain design decisions are specifically motivated by (behavior change) theories. The system also does not use any specific tailoring approaches, other than that the use of an avatar should give a *personal* experience.

The author's focus is on the (embodied) conversational agents and provide much more background regarding this topic in e.g. (Bickmore & Picard 2005, Bickmore et al. 2005). More recently, the author's further explored the use of conversational agents to promote — amongst others — physical activity in (Bickmore et al. 2013).

2.4.7 Haptic Personal Trainer

The work described in (Qian et al. 2010) and (Qian et al. 2011) takes a low-level approach to feedback on walking behavior for older adults. The solution is a smartphone based application that measures steps taken using the built in accelerometer and provides haptic feedback (vibrations) to the user to stimulate walking faster or slower. The author's argue that haptic feedback is ideal because it removes the need to consult the visual interface of the phone while in motion. Besides this relatively simple assumption, the work does not seem to be based on any specific theories of behavior change or persuasive technologies. The system also doesn't implement any specific tailoring approaches other than providing feedback to the user.

The work is implemented as a phone application (Nokia N95) including a step-counting algorithm that uses the phone's internal accelerometers. The haptic feedback consist of structured vibration pulses with varying durations that are composed into rhythmic units which are detectable by the user. A large focus of the work is on the development of these perceivable tactile icons (so-called 'tactons') that can aid in non-visual interaction between the system and the user. To find an optimal way of using the tactile channel for feedback, several experiments were conducted that are explained in more detail in Qian, Kuber & Sears (2011).

In a second version of the system, the haptic feedback was augmented with auditory feedback and additional work was done in amplifying the vibration signals of the phone. Sixteen older adults participated in an experiment to evaluate the effectiveness of the feedback modalities. For each participant a baseline pace (steps/minute) was recorded, and based on this lower- and upper limits of their ideal desired pace were calculated. During the experiment, participants were asked to walk for 90 seconds in four different feedback conditions (no feedback, audio-, tactile-, and audio + tactile feedback). Results showed that best performance — in terms of additional steps walked per minute — was achieved in a multimodal feedback scenario using both haptic and audio cues.

2.4.8 Everywhere Run

The Everywhere Run smartphone application described by (Mulas et al. 2011) is designed to motivate and support users during running activities. The main goal of the application is to foster social interaction between runners and real personal trainers so that runners can receive personal training plans. The main motivation behind the work is that one of the biggest barriers for beginning runners is the issue of creating a proper workout schedule; and social interaction can help motivate users to exercise. Other than this general assumption there is no reference to theoretical background on e.g. goal setting theory. In the Everywhere Run application, workout plans can be created on the Android based smartphone application, but more importantly they can be sent to the application via email. This way of interaction allows professional trainers to design a personalized, detailed workout plan for the users and send it to the user. When the user starts a session, the application functions as a virtual trainer by making sure the user adheres to his training plan. The application's screen serves as a glanceable user interface, while audio cues guide the user through the workout session. The application promotes social interaction by enabling users to share their workout schedules with others. With the focus on social interaction and setting of exercise goals, the system does not include any other tailoring functionalities.

A 2012 "updated version" of the system, dubbed *Everywhere Race!* is presented in (Mulas et al. 2012). This version focuses on the issue of finding opponents for a virtual race by implementing an integration with the popular social network *Facebook*. The application allows the sharing of created virtual races through Facebook, and allows users to search for and join existing races. The Everywhere Race system does not specify the implementation of algorithms for automatic matching of competitors. The system was evaluated with 35 users over a period of 30 days and results in terms of motivation and physical activity seem to be positive.

#	**Design principle**
1	Give the user proper credit for activity.
2	Provide personal awareness of activity level.
3	Ensure fair play.
4	Provide a variety of motivational tools.
5	Provide feedback on activities done.
6	Consider the practical constraints of users' lifestyles.
7	Provide both short-term and long-term motivation.
8	Support social influence.
9	Provide possibility of integration with existing solutions.
10	Protect users' privacy.

Table 2.5: The 10 design principles based on literature review from social features and game design principles in fitness applications from Bielik et al. (2012).

2.4.9 Move2Play

The Move2Play system described in (Bielik et al. 2012) is a conceptual design of an innovative platform to stimulate healthy living and improve quality of life. A key aspect of the system's design is the integration of different motivational facilities (intrinsic and extrinsic) delivered in a non-obtrusive manner, taking into account the current context of the user. To guide the system's design, the authors first describe a set of 10 design requirements for successful implementation of physical activity encouragement tools, listed in Table 2.5. These design principles are based on literature review regarding social features for wellness applications (Ahtinen et al. 2009) as well as game design principles for fitness and exercise applications (Campbell et al. 2008, Yim & Graham 2007). In their motivation and background the authors refer mostly to similar existing solutions and do not touch upon any behavior change theories.

The authors describe a design that encompasses the full scope of their own design requirements, resulting in a conceptual system that includes many different tailoring approaches and technological innovations related to activity monitoring and coaching. The authors propose to increase the accuracy of activity assessment by combining accelerometers with location information provided by wireless networks. Activity data is fed back to the user via a number of different visualization options, and incorporates goal setting and group feedback. Furthermore the system contains social encouragement through social network integration, gamification techniques such as achievements, badges and unlock-able content, as well as an animated motivational agent to foster natural communication with the user. Although the envisioned design is innovative, for the most part the system appears to be conceptual. Besides a very high-level modular architecture, no technical implementation or algorithmic details are given. From the described modules some specifics are given regarding the features of a user- and domain model that are used for inferring optimal activity recommendations. However, from the description of the evaluation it appears that only minor parts of the described functionality are implemented.

2.4.10 ActivMON

The work in (Burns et al. 2012) focuses on a low-complexity Ambient Light display as feedback device for representing the user's level of physical activity. The paper describes a wrist-worn device that notifies the user of his progress towards his daily activity goal. The device contains a 3D accelerometer and a multicolor LED. The ambient light lights up red at the start of the day, and progressively turns to yellow and green as the user reaches a predefined activity goal — a more recent paper describes in detail the use of a color gradient as feedback modality (Burns et al. 2013). Activity is measured using a 3d accelerometer, where thresholding of the magnitude of acceleration is used to increment an activity counter. Furthermore, the device can connect to a smartphone to enable data synchronization with a server. This connects the user to a social group of users which allows the device to show in near real-time when group members are physically active. The system was evaluated with a group of five colleagues. After a baseline week of measuring, each group member received a target activity goal set to 105% of their baseline activity. Although the device had some usability problems, four of the five users averaged higher activity levels than their personal goals at the end of the second week.

The system targets users with attitudes and behaviors described as "less motivated", and argue that these users are less willing to commit time and effort to monitor detailed information regarding their physical activity performance. This is the motivation for using a less complex interface that is simpler to engage with, such as the ambient display. The decision to include a group-component is based on the recommendations from Consolvo et al. (2006) of supporting social influence. Additional details regarding the ActivMON application are presented in (Burns et al. 2011). Most notably this work describes an interesting tailoring approach in the form of an *Adaptive Goal Setting* algorithm. The device will automatically calculate an average activity level of the first week of use and set a personal goal to 105% of this value. In subsequent weeks, if the user reaches his goal, it is automatically raised by another 5%; if the user fails to reach the goal it is left unchanged. Through the use of a web interface, users can influence the goal-setting behavior using simple "decrease/increase" buttons to alter the given goal. Although the authors do not mention it as motivation, this implementation is a good example of using the principles of the Goal Setting Theory (Locke & Latham 2002) by providing challenging, achievable goals as well as providing a mechanism for users to commit to those goals.

2.4.11 BeWell+

The BeWell+ application described in (Lin et al. 2012) combines physical activity with sleep- and social interaction monitoring and provides feedback along those three dimensions using an ambient display on a smartphone's wallpaper (see Figure 2.2.c). The article focuses on two specific tailoring approaches that aim to improve a previous version of the system, both of these are aimed at tailoring the experience better to individual users. The theoretical background of the work mostly consists of a comparison to earlier work as well as similar existing applications, and includes no mention of specific behavior change theories that are consulted or used to drive any specific design decisions.

The first improvement in this new version of the system is the use of *community adaptive well being feedback*. The idea is that the performance of individual users is compared to other users of the system in terms of their well being scores. Through the use of an algorithm that automatically groups similar users, a more realistic assessment of an individual user's performance is possible. This is achieved by calculating how well he did compared to his peers, instead of comparing with an ideal situation (which may be unachievable). The system effectively matches the user with *positive role models*, users who are similar in behavior, but perform slightly better. The similarity matching is done using a *behavioral similarity network*, a weighted graph structure in which the nodes represent users, and the edge-weight is defined by the similarity between users. This similarity is based on mobility- (measured through GPS tracking), temporal-, and activity patterns (measured by activity inference). Details regarding the lifestyle similarity calculations can be found in (Lane et al. 2011). By repeating the matching progress as new data is available the system constantly adapts these groupings.

The second aspect of the paper focuses on *well being adaptive energy allocation*. As a way of saving battery on the smartphone, the authors developed a system that prioritizes resource allocation to those sensors and modules of the system that are most relevant. For example, if the user has a normal sleep pattern but low physical activity, the system would shift its sensing priorities from sleep to physical activity. Specifically, the system can adapt the frequency of sensor sampling, feature extraction, and activity inference; as well as the frequency of communicating with the back-end cloud infrastructure to send data or collect revised well being scores from the adaptive well being feedback component.

2.4.12 "Analytic, Social, Affect"

Three variations of a daily physical activity coaching tool are described in (King et al. 2013): an *analytically* framed version, a *socially* framed version, and an *affectively* framed version. All three applications work on a smartphone that measures daily physical activity patterns and provides a glanceable display for providing feedback of the current level of activity. The work forgoes technical implementations or descriptions of algorithms for discussion on theoretical background and evaluation.

The analytic application distinguishes itself by adding user-specific goal setting, set by the user himself every week, which was also added to the feedback on the smartphone's display. Users are provided with goal options of increasing difficulty, with the idea that graded goals increases self-efficacy while "nudging" individuals towards their goals (Thaler & Sunstein 2008). When weekly goals were not met, the system provides a "trouble shooting" mechanism that helps users in setting more attainable goals, and also provides additional informational tips on reaching the weekly goal. This version of the application is heavily based on behavior change theories related to self-efficacy and goal setting. The social application focuses on group-feedback displayed by a series of avatars on the smartphone's wallpaper, representing the user and other group members. Each avatar's posture changes based on the activity level of its corresponding user to indicate coarse performance about individuals in the group. In addition, physical activity feedback was presented in relation to the group performance, as well as the performance of a second "competing" group. In order to provide a 'positive role model', each group contained a virtual participant that inhibits healthy activity behavior. The affectively framed version of the application focuses on the use of a virtual avatar, a bird, that changes posture, position, and movements based on the level of physical activity performed by the user (see Figure 2.2.d). Similar to the UbiFit garden system (Consolvo et al. 2008), this visually appealing representation of physical activity becomes more attractive if more physical activity is performed.

2.5 Definition of concepts

The overview of related work given in Section 2.2 already touches upon the issue of definitions in the field of tailoring. The process of our literature search made it even more clear that for example terms like 'personalization', 'tailoring' and 'individualization' are used rather interchangeably throughout the literature. The article by Hawkins et al. (2008) provides some clear definitions, by coining 'tailoring' as an overall umbrella term for various sub concepts defined as "feedback", "personalization" (we use *user targeting*), and "content-matching" (we use *adaptation*[2]). Unfortunately, Hawkins' use of the term 'personalization' to define a very specific tailoring strategy is rather confusing. Therefore we opted to use the term "**user targeting**" — a term taken from the field of advertising (see e.g. (Wang et al. 2011)).

In this section we expand upon Hawkins' concepts by defining an extended model of tailoring. In Section 2.6 the model is used as a framework for the analysis of

[2]The term "adaptation" is taken from Lustria et al. (2008) as an alternative to "content-matching".

Figure 2.2: Example screenshots of four different real time physical activity coaching applications: **a)** the MOPET system by (Buttussi et al. 2006), **b)** the UbiFit Garden by (Consolvo et al. 2008), **c)** the BeWell+ application by (Lin et al. 2012) and **d)** the socially framed application as described in (King et al. 2013).

the papers included in this survey, while at the same time validating the definitions given below.

One of the reasons why Hawkins et al.'s (2008) definitions of tailoring fall short for this survey, is the introduction of recent advanced tailoring techniques such as **context awareness** and **self learning** that are not adequately covered. Subsequently we feel that the concept of **goal setting** does not fit the model of Hawkins and should be treated separately. Thirdly, where Hawkins et al. (2008) shortly treat interaction between users in their description of comparative feedback (comparing results with that of a peer or group of peers), we will elaborate on the idea of **inter-human interaction** as a separate concept. Finally, it is worth noting that the presented model can be seen as a detailed elaboration on part of the persuasive system design methodologies as defined by Oinas-Kukkonen & Harjumaa (2009), who in turn attempt to develop solid design methods for persuasive technology based on the fundamentals that have been laid out by Fogg (2003).

2.5.1 Tailoring & communication

Before defining the various aforementioned tailoring concepts we will provide a working definition of *tailoring* itself. When talking about *tailoring*, we use the word as a transitive verb, meaning we are tailoring *something* to *someone*. In our context of physical activity coaching systems, the *someone* is the user of the system who wants to (or has to) change his or her physical activity behavior. The *something* is in our case a more complex concept, namely *communication*. In short, **we tailor communication to the user**.

We first need to define 'communication' in the context of physical activity coaching. Communication (from the Latin word "communis", meaning to share) is a process of sharing information between two or more participants. In our context, we are talking mostly about a computer agent sharing information with a human user in order to e.g. motivate, or inform. Every communication instance can be seen as having four distinct properties: **timing**, **intention**, **content**, and **representation**. As an example, consider the following hypothetical message from a physical activity coaching system to its user:

"*You haven't been active enough today. Maintaining a healthy level of physical activity can drastically reduce the chances of cardiovascular disease. In order to achieve your daily goal, you need to walk for at least another 18 minutes or perform 12 minutes of vigorous exercise.*"

In this example, the **timing** of this communication would be the moment at which the system would present it to the user. The communication instance has three different **intentions**: (1) to provide information on the user's current progress (i.e. feedback), (2) to inform about the benefits of physical activity, and (3) to give a suggestion on an activity to perform. The **content** of the communication is the factual information presented (e.g. the fact that you haven't been active enough), and the **representation** is in this case a rather long-winded natural language text.

Now consider the second example in Figure 2.3: a screenshot of the web portal of the commercially available Nike+ Fuelband activity coach system. Although very different from the natural language example above, in this form of communication we can also identify **timing** (in this case the timing is user initiated as the time when he or she chooses to visit the web portal), **intention** (to provide feedback), **content** (activity values and goals for seven days, expressed in Nike Fuel points, the total amount of Nike Fuel earned, and the number of days the goal was reached), and **representation** (a bar graph with goal lines).

Looking at these communication instances, the goal of any form of tailoring would be to increase the likelihood that the system successfully conveys its intention to the user by matching each of the communication properties in some way to the user and/or his context.

From the literature, we have identified seven tailoring concepts with varying levels of complexity: *Feedback* (**FB**), *Inter-Human Interaction* (**IHI**), *Adaptation* (**Ad**), *User Targeting* (**UT**), *Goal Setting* (**GS**), *Context Awareness* (**CA**), and *Self Learning* (**SL**). These concepts will be defined in Section 2.5.2 below. Then in Section 2.5.3 the connections between the various concepts will be defined and explained with examples.

Figure 2.3: Screenshot of the Nike+ Fuelband web portal displaying a seven-day overview of activity counts (expressed in Nike Fuel points) and daily goals. Source: http://www.pocket-lint.com/news/114706-7-days-with-nike-fuel-band (August 2013).

2.5.2 Tailoring concepts

The seven tailoring concepts and their relationships to the communication properties — timing, intention, content and representation — are explained below.

Feedback — Feedback involves presenting individuals with information about themselves, obtained during assessment or elsewhere. To give feedback is a strategy for achieving the intention of motivating the user to change a behavior (another strategy could be e.g. to inform about the advantages of physical activity). Broadly speaking, three forms of feedback can be distinguished: descriptive, comparative and evaluative. Descriptive feedback *reports what is known about the recipient based upon his or her data*, comparative feedback *contrasts what is known about the recipient with what is known about others* and evaluative feedback *makes interpretations or judgments based on what is known about the recipient* (Hawkins et al. 2008). From these three forms of feedback, we treat comparative feedback as part of the broader tailoring concept "inter-human interaction" described below. Feedback is a tailoring concept that can exist on it's own, i.e. a communication can consist of solely a feedback message.

Inter-human interaction — We define inter-human interaction as the support for any form of interaction with other real human beings. Inter-human interaction is for example any type of built in support to contact professionals or peers, share information about performance or progress to selected individuals or any built in support for professionals or peers to contact the user to provide support or advice on physical activity. Inter-human interaction can provide additional motivation (peer pressure) or can provide a feeling of safety in case of a connection with a healthcare professional.

Adaptation — Adaptation "attempts to direct messages to individuals' status on key theoretical determinants (knowledge, outcome expectations, normative beliefs, efficacy and/or skills) of the behavior of interest" (Hawkins et al. 2008). For example, someone that wants to be more active in order to lose weight would receive different messages from the system than someone who wants to be more active to prevent exacerbation of a disease like COPD, since the motivation in both cases is different. Similarly, a user in the precontemplation phase of the stages of change model could receive motivation as to why being active is good for you, while such information would not be appropriate for someone in the action phase.

User targeting — User targeting attempts to increase attention or motivation to process messages by conveying, explicitly or implicitly, that the communication is designed specifically for 'you' (Hawkins et al. 2008)[3]. When tailoring communication, user targeting is a technique that attempts to fit the representation to the individual. Hawkins et al. (2008) defines the three most common tactics as follows. *Identification* attempts to target the individual by identifying the recipient by name (see (Dijkstra 2014) in this issue), using pictures of the recipient, or recognizing the recipient's birthday. *Raising expectation* through mentioning explicitly that an advice was specifically designed for the recipient — a form of 'placebo tailoring'. And finally *contextualization* by e.g. matching the content of messages to the recipient's age, sex, culture, or other user parameters.

Goal Setting — According to the Goal-Setting Theory, people are more likely to change behavior the higher the specificity and (achievable) difficulty of a

[3]Note that Hawkins uses the term 'personalization', which we changed to avoid confusion.

goal (Locke & Latham 2002). Goal setting is a technique used to present the user with short-term, as well as long-term goals that can instill a feeling of progress over the course of an intervention or the day. Goal setting is a tailoring concept that can only be used in combination with feedback.

Context awareness — For context awareness we adopt the definition from Dey & Abowd (1999): "*A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task*". In the area of physical activity coaching, we deal not so much with tasks, but with 'needs' or 'goals'; but the critical part of this definition is the notion of *context*. We define context as any information, *non-critical to the application's main functioning*, that can be used to characterize the situation of a user or the system (i.e. not including user characteristics). Context awareness is a tailoring concept that can be used in various ways to tailor timing, content and/or representation of communication instances.

Self learning — Tailoring techniques such as adaptation and user targeting aim to adapt communication to a user, while context aware systems aim to adapt to a user in a particular context. But the 'user' is a very dynamic entity. Any application that employs tailoring techniques has an intrinsic user model, which is never perfectly accurate or complete. A self-learning application is able to update its internal model of the user by recording and learning from the various interactions the user has with the application. Within an intervention that aims to achieve behavior change, the user is almost by definition something that changes over time. The intervention could (and should) for example move the user forward through the stages of change. This means that an intervention or behavioral change tool that uses adaptation to tailor communication intent to a specific stage of change should change with the user throughout his use of the application. This ability of a tool to change with the user is defined as *self learning*. A system can be self learning in the way it employs other tailoring techniques such as adaptation, context awareness, goal setting or user targeting.

2.5.3 The tailoring model

Figure 2.4 shows the seven tailoring concepts and the ways in which they can interact with each other. The graph includes an extra node for *Motivational*

Messages (**MM**) that is not a tailoring concept in itself, but a common technique used in physical activity promotion applications that can be enhanced through various ways of tailoring. This node is encircled, depicting that it is a possible end-node in the graph. The graph should be read as follows. Any path through the graph represents a possible combination of one or more tailoring techniques that should end at one of the two end-nodes (**FB** or **MM**). For example, the path **CA**→**FB** represents "context aware feedback", e.g. tailoring the timing of feedback based on the user's location, the time, or weather as in (op den Akker et al. 2010). Designers of physical activity promotion systems can use this to find various paths of tailoring. By starting at an end node (**FB** or **MM**), following edges back through the graph will lead to ever more complex forms of tailoring. The self learning (**SL**) nodes in the graph are special in that self learning can only be used in combination with other tailoring concepts.

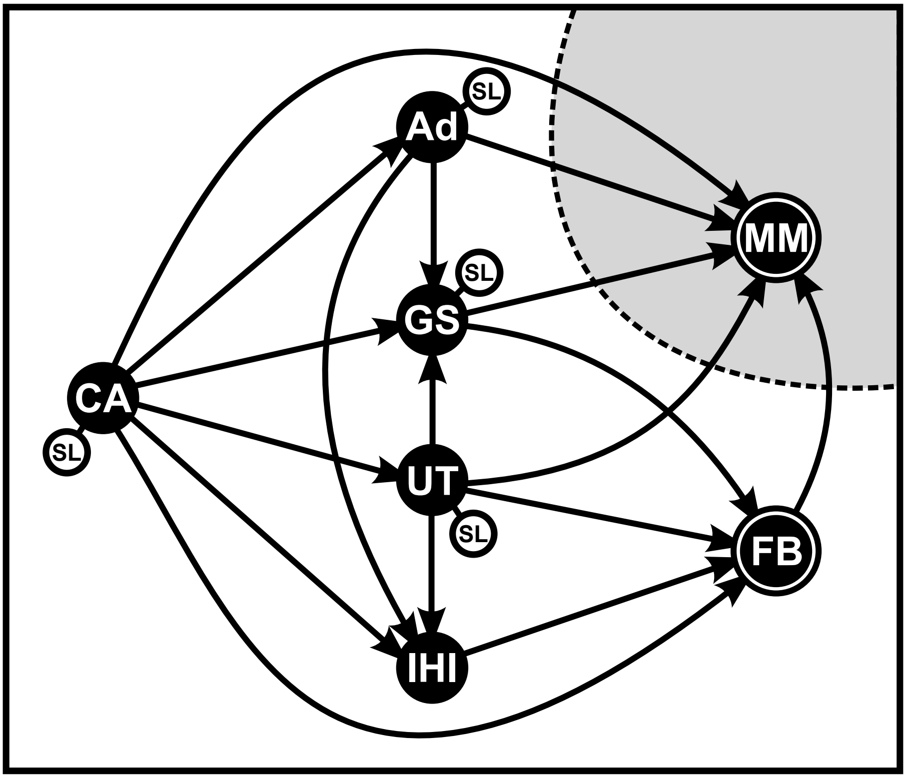


Figure 2.4: The tailoring model, showing the various tailoring concepts and how they can be combined to form various motivational communication through feedback (**FB**) or motivational messages (**MM**) — which is not a tailoring concept in itself. The other non-end nodes represent context awareness (**CA**), goal setting (**GS**), inter-human interaction (**IHI**), Adaptation (**Ad**), and user targeting (**UT**). Self learning (**SL**) can be used to further augment various other tailoring concepts. Starting at one of the end nodes, designers of physical activity coaching tools can work their way back through the graph to find ways in which to tailor their communication to the user.

The following list gives examples for each of the links defined in Figure 2.4. For those links directed to an end-node the relation to the communication model properties of *timing*, *intention*, *content*, and *representation* is given — indicating which of these communication properties it can affect.

CA→FB Tailor the feedback based on the user's context. E.g. use a different modality based on the user's location such as audio feedback in private settings, and text-based feedback in public settings (*timing*, *content*, *representation*).

GS→FB Improve the effectiveness of feedback by also providing daily, weekly or monthly goals, e.g.: "You did 8,472 out of 10,000 steps today!" (*content*).

UT→FB Change the wording of a feedback message by adding user targeting, e.g.: "Mr. Johnson, you have done 7,000 steps so far today!" (*representation*).

IHI→FB Receiving group based feedback (*timing*, *intention*, *content*, *representation*).

FB→MM Add feedback to a general motivational message, e.g. "*You have done 7,384 steps today. Keep it up!*".

CA→MM Provide a general motivational message based on user context, e.g. when the weather is nice: "It's perfect weather to go for a walk!" (*timing*, *content*, *representation*).

GS→MM Simply stating the daily goal, or informing about an automated change in goals (*content*).

IHI→MM Receiving a (motivational) message from a user in a peer-group (or friend) through the application (*timing*, *intention*, *content*, *representation*).

Ad→MM If the subject is in the precontemplation stage of change, inform him about the benefits of physical activity (*intention*, *content*).

UT→MM Provide the subject with a general motivational message, enhanced with user targeting, e.g.: "Mrs. Johnson, maybe you can take a lunch walk today" (*representation*).

CA→GS Tailor a user's goals for different days of the week, based on e.g. the user's agenda.

CA→UT Combine user information (likes, dislikes) with current context, e.g. recommending a cycling route nearby if the user likes cycling.

Ad→GS Set goals based on psychological constructs, e.g. set an easily attainable goal for someone with low Self Efficacy.

Ad→IHI Place a subject in a group of peers based on similarity in psychological constructs (e.g. same Stage of Change).

UT→GS Set a specific user goal based on a user parameter such as BMI.

UT→IHI Place a subject in a group of peers based on similarity in user preferences, e.g. likes, dislikes, age, gender.

CA→IHI Place a subject in a group of peers based on his current location (e.g. users in the same city).

2.6 Survey of real-time, tailored coaching systems

A total of 15 papers were included in the survey. We have analyzed each of the described systems regarding the use of the seven tailoring concepts defined previously. Table 2.6 lists all the included papers, and indicates whether the application makes use of these tailoring techniques. The table is split into three parts, the first part contains real time applications that deal with daily life activities, the second part with real time applications that are more related to exercises, and the third part contains two excluded papers that serve as useful examples throughout the analysis below. The columns indicate the use of Feedback (**FB**), Inter-Human-Interaction (**IHI**), User targeting (**UT**), Adaptation (**Ad**), Goal Setting (**GS**), Context Awareness (**CA**), and Self Learning (**SL**).

In the seven subsections below we will explain for each of these aspects of tailoring how it is currently applied in the literature, and how these methods are justified by underlying theories.

2.6.1 Feedback

The use of feedback is the most obvious form of tailoring, and it is applied in all of the real-time physical activity coaching system we have reviewed. Because it is such a common motivational tool, it is worth to have a more detailed look at its internal mechanisms and relevant (sub) properties. Feedback on physical activity is simply presenting the measured amount of activity performed to the user. It is almost impossible to imagine a real-time coaching system that does not

Article	Tailoring FB	IHI	UT	Ad	GS	CA	SL
Real-Time Systems — Daily Life Activities							
(Bickmore et al. 2009)	✓	-	-	-	-	-	-
(Bielik et al. 2012)	✓	✓	✓	-	✓	✓	✓
(Lin et al. 2012)	✓	✓	-	-	✓	-	✓
(Burns et al. 2012)	✓	✓	-	-	✓	-	-
(Fujiki et al. 2008)	✓	✓	-	-	✓	-	-
(King et al. 2013)	✓	✓	-	-	✓	-	-
(Consolvo et al. 2008)	✓	-	-	-	✓	-	-
(Qian et al. 2010)	✓	-	-	-	-	-	-
(Qian et al. 2011)	✓	-	-	-	-	-	-
Real-Time Systems — Excercise Based							
(Mulas et al. 2011)	✓	✓	-	-	✓	-	-
(Buttussi et al. 2006)	✓	-	-	-	-	-	-
(Buttussi & Chittaro 2008)	✓	-	✓	-	-	-	-
(Ståhl et al. 2008)	✓	-	-	-	✓	✓	-
(de Oliveira & Oliver 2008)	✓	✓	-	-	-	-	-
(Oliver & Flores-Mangas 2006)	✓	-	-	-	✓	-	-
Non Real-Time Systems — Excluded							
(Arteaga et al. 2010)	✓	✓	-	✓	-	-	-
(Saini & Lacroix 2009)	✓	✓	-	-	✓	-	-

Table 2.6: Overview of papers included in the literature study. The first group of nine articles deal with Daily Life (**DL**) activities, the second group of six are mostly exercise based systems. The final two papers are excluded because the systems are not real-time (**RT**), they do however serve as useful examples in the discussions below. The last columns mark the use of feedback (**FB**), inter-human interaction (**IHI**), user targeting (**UT**), adaptation (**Ad**), goal setting (**GS**), context awareness (**CA**), and self learning (**SL**).

employ this form of tailoring, but the ways of doing so are diverse. We will discuss the ways in which feedback communication instances can vary in their **timing**, **content**, and **representation**.

Timing

We define the timing of feedback as the moment at which users are able to request their current performance level, or are given a report of their performance by the system. The timing aspect of feedback is such a distinctive and widely varying parameter, that this survey already focuses on a subset of these type of systems — those that are able to provide feedback in real-time. An example of a system that, based on this criteria, was excluded from this survey is the Philips DirectLife system described in Saini & Lacroix (2009). The system requires the user to physically dock an activity sensor to a PC in order to upload data to a central server before being able to receive feedback through regular emails or by visiting a website.

The timing of feedback can vary between e.g. receiving weekly emails, all the way through continuous feedback as for example in the music-based personal trainers, MPTrain (Oliver & Flores-Mangas 2006) and TripleBeat (de Oliveira & Oliver 2008) that automatically adjust the music that is playing while exercising in order to match your current gait and heart rate goals. This type of music-based feedback is however meant for exercise-based physical activity coaching and is not suitable for use in daily life. The UbiFit Garden described by Consolvo et al. (2008) is designed for coaching on daily life activities, and presents a ubiquitous method of providing feedback by using a glanceable mobile phone wallpaper to provide subtle feedback to the user whenever and wherever the phone is used. This way, whenever the user wants to receive feedback he is able to see his own performance quickly on the display, and even if the user has to perform a different task on his phone, he is subtly reminded of his activity progress.

The initiative is however with the user, which means that opportune moments for behavior influencing may go by unused. The literature provides two ways of overcoming this issue. One is to simply provide feedback that is always visible as in the ActivMON system presented by Burns et al. (2012). The ActivMON is a wrist worn sensor incorporating a multicolor LED that slowly changes color from red to green as the user reaches his daily activity goal. The second and more obvious strategy is to let the system decide when to provide feedback. For example, Qian et al. (Qian et al. 2010, Qian et al. 2011) employ continuous feedback where the initiative is with the system itself. The system, developed for a Nokia N95 mobile phone, tracks the user's pace (steps/minute) and produces vibrations to indicate the user to walk faster or slower than their own baseline, or comfortable speed. Similarly, but in a more realistic activity coaching setting, the system developed by Bickmore et al. (2009) provides system-initiated positive reinforcement after the detection of a 10 minute bout of physical activity. Their system supports a flexible way of dealing with users who fail to respond, as well as a flexible way of defining a protocol for timing of user notification based on various state transitions and events.

With this last feature the authors touch upon a question that is in our view critical, and currently unanswered in the literature: what would this ideal protocol be? When examining the timing of feedback, there seem to be three dimensions that play an important role. Initially there is the choice whether the **initiative** for providing feedback lies with the user of the system, or with the system itself. Subsequently, when the system takes the initiative, the question remains what the right **moment** and **frequency** of feedback is. Lastly there is the issue of making sure that the user actually notices and responds to feedback at the moment it is provided.

Content

Besides the question of when to give feedback, there is the question of what to communicate to the user. Arteaga et al. (2010) performed detailed focus-group based user studies for the design of their system to motivate teenagers' physical activity, and one of the aspects included in their studies is personality theory. Based on the idea of similarity attraction, proposed by Byrne & Nelson (1965), the authors state that the set of motivational phrases included in the system should be matched to the users' personality traits. After presenting a set of motivational phrases to five focus group participants, some phrases were however unanimously preferred or disliked, regardless of personality. Whether or not this finding contradicts the idea that phrases should be matched to personality is unclear. It does show however, that carefully choosing the content of a feedback communication matters. There are other examples in the literature of natural language feedback messages in which special attention is given to phrasing. The wording of the feedback presented by the MOPET system by Buttussi et al. (2006) was deliberately chosen to be gentle ("You are walking regularly. If you are not tired, try to increase your speed."), because, as the authors state, the more aggressive inciting of the Philips Virtual Coach (Ijsselsteijn et al. 2004) did not work as well as expected. More complex approaches of matching motivational messages to the user have been proposed, as for example in Di Tullio & Grasso (2012), who propose an abstract argumentation framework — the values system — for automatically generating motivational arguments based on a matching between system- and user's beliefs. The system is based on a logical inference of arguments to influence the conversational partner's values (beliefs) system.

It is clear that in the design of some systems, the importance of content is recognized. However, in most of the included papers, either no special attention is given to the content, or the content is visual, music based or tactile (feedback representation will be discussed below). As such it is hard to draw firm conclusions regarding content. Intuitively it seems relevant, and Byrne and Nelson's similarity attraction suggests that content of feedback is ideally matched to individuals. However, further research is needed in order to provide clear recommendations on how to tackle content in a tailored physical activity coach.

Representation

Natural language text is not the only form of representation in which physical activity coaching can take place. The chosen representation for interacting with the user is the most diverse aspect of feedback in the included articles in this survey. It ranges from spoken dialogue (Bickmore et al. 2009, Mulas et al. 2011, Buttussi et al. 2006, Buttussi & Chittaro 2008, Arteaga et al. 2010) and music (de Oliveira & Oliver 2008, Oliver & Flores-Mangas 2006) to more graphical displays (Fujiki et al. 2008, Consolvo et al. 2008) or even tactile (Qian et al. 2010, Qian et al. 2011) or using simple ambient light (Burns et al. 2012). The choice of representation seems to be mostly based on the expected use of the application as well as its target audience. For example, all of the systems that are focused on exercise based physical activity employ either spoken dialogue or music as feedback modality, as the user would be too occupied with performing exercise to use the visual display of a coaching device. As it is difficult to give detailed feedback on performance through the audio channel, these modalities are always combined with a more graphical display, such as a graphical representation of a virtual race in e.g. Mulas et al. (2011).

The concept of exercise based feedback through a multimodal graphical/audio interface is not exactly new, and there are many free applications for smartphones available on the market, such as RunKeeper[4] or SportyPal[5], that offer exactly this. Those applications that aim to support physical activity during daily life show a more diversified use of feedback representation. Qian et al. (Qian et al. 2010, Qian et al. 2011) describe the development of a mobile phone based application that can provide haptic feedback on walking behavior for older adults. The application uses the phone's internal accelerometer to count steps, and provides vibration signals for real-time feedback to the user: move faster, or move slower. The author's state that vibration signals are an ideal modality for feedback because it reduces the need to consult the visual interface whilst in motion. There is however an obvious limitation in richness for this type of feedback. Similar problem of richness are obvious in the ActivMON system envisioned by Burns et al. (2012), who developed an ambient display that is worn on the wrist that uses a light-emitting diode to indicate the user's activity, as well as the activity of peers in the user's social group.

The effectiveness of using such simple modalities for providing feedback is unclear, as it has not been evaluated in much detail. The intensity of coaching is low, and we expect that at least some people require much more in terms of persuasion and motivation. As Bickmore et al. (2009) suggest, "...perhaps the most effective technologies are those which come closest to the 'gold standard' of one-on-one, face-to-face counseling with an expert health provider". They, and other authors (Bielik et al. 2012, Fujiki et al. 2008, Buttussi et al. 2006, Buttussi & Chittaro 2008, Ståhl et al. 2008, Arteaga et al. 2010) employ embodied agents to interact with the user in order to achieve a heightened level of social bonding and trust between the user and the system. A downside of presenting a human avatar

[4] http://www.runkeeper.com/
[5] http://www.sportypal.com/

to the user is that, as Bickmore et al. (2009) conclude, the perception of agent reliability is a pre-requisite for effective health outcomes as well as for relational bonding.

In summary

We described how feedback, or presenting individuals with information about themselves, is implemented in greatly varying ways across the different studies. We described the complex issue of feedback through the three properties of a communication instance: its **timing**, **content** and **representation**, and showed the diversity of use in each aspect. All of the included papers describe systems that are in the conceptual, or prototype phase, and have undergone only small scale evaluations at best. This, and the fact that for example *timing* of feedback is usually only a very small part of a larger complex activity coaching system, it is hard to draw firm conclusions on the effectiveness of the various employed strategies.

2.6.2 Inter-human interaction

The second tailoring technique is in a way quite different from feedback, as with inter-human interaction, it is not the system that does the coaching directly, but merely enables the user's peers to coach him or her. The inter-human interaction strategy is used in seven out of the 15 included papers. In the virtual running application 'Everywhere Run' by Mulas et al. (2011) users are guided through a running workout by a coaching application on their smartphone. The system allows users to download a workout plan from other users or from a paid personal (real life) coach. A system that can connect with professional healthcare personnel or coaches can provide a feeling of safety and comfort to the user, as he can rely on the professionally trained coach to design a workout schedule that matches the user's capabilities. Another IHI technique in a mobile phone based exercise coach is the inclusion of virtual competition as presented by Fujiki et al. (2008) and de Oliveira & Oliver (2008). The coaching system described by de Oliveira & Oliver (2008) supports virtual competitions in which the user can run against a fictional runner, the user himself on previous runs, or real runners that have previously run with the system. The competition is set up in such a way that it does not award faster running, or burning more calories, but awards those runners that achieve their own exercise goals in a healthy manner. The system allows users to choose their own competitor from a database of registered users, or to let the system automatically select a competitor based on performance similarity metrics. An evaluation study with 10 subjects showed no significant increase in performance between the system with competition and an earlier version of the system that did not include the competition aspect. Although from a subjective evaluation, five out of 10 subjects indicated that the competition element was the main reason for enjoying the system.

Other methods of supporting IHI are presented by Bielik et al. (2012), who envision embedding social network capabilities in their coaching system, allowing users of the system to connect, share, and communicate with each other through the application. Sharing your physical activity results, and particularly seeing the results of other users could work in a motivating way, similar to the virtual competitions described above. Such techniques are already widespread in commercial fitness tracking applications such as RunKeeper or SportyPal. Social influence can be a strong motivating element, and being able to connect to your friends or family through an activity coaching system could increase the connectedness that you experience with the system and foster long-term enjoyment in using the system.

2.6.3 Adaptation

Adaptation, or the process of tailoring message content to individuals' status on key theoretical determinants, was not found to be applied in any of the real-time physical activity coaching systems included in the study. We did however find one article describing the application of adaptation techniques in a physical activity coach by Arteaga et al. (2010), but the system did not adhere to the "real-time" criteria. Nevertheless, the system shows an interesting application of adaptation, which could have been implemented in real time. Arteaga et al. describe a smartphone application that uses physical activity games to motivate teenagers to be more active in order to prevent obesity. A core component of the system is that it assesses the user's personality profile through the use of a questionnaire that has to be filled out when starting the application. The personality profile defines the individual on the level of personality traits such as 'extraversion', or 'openness', for which the authors showed correlation to the results of a self-developed questionnaire consisting of physical activity-, personality-, and mobile phone related questions. The system then uses the personality traits to define a set of games that are specifically relevant to the user's personality. The system was evaluated in a small trial with five users, in which there was no comparison with a system that did not use the personality trait based adaptation step, so it is unfortunately not possible to comment on the effects of the authors' approach.

Adaptation seems to be a useful tailoring technique. The example given in the definition (Section 2.5) about not telling a user in the action stage of change what the benefits of physical activity are seems obviously useful and technically easy to implement. Unfortunately, the current state of the art of real time physical activity coaching systems is such, that these methods are simply not being employed.

2.6.4 Goal Setting

Goal setting is a tailoring form that can mostly be seen as an extension of feedback. Where feedback presents users with information about their performance (e.g. doing 7,000 steps), goal setting adds a certain value to the presented information (e.g. reaching 7,000 out of 10,000 daily steps). According to the Goal-Setting Theory (Locke & Latham 2002), people are more likely to change behavior the higher the specificity and (achievable) difficulty of a goal. Many of the included articles describe a form of goal setting for their applications (nine out of 15). Broadly speaking, two levels of complexity within goal setting strategies can be identified: simple numerical methods, and methods based on more complex tasks or schedules.

Similar to the example given, many of the applications that focus on daily life activities include a simple numerical goal, representing e.g. the amount of steps to take, or minutes of daily activity to perform. The variations in these applications relate to (1) the associated time scale of the goals — Bielik et al. (2012) allows users to set both short-term and long-term goals — and (2) the representation of the goals — Consolvo et al. (2008) shows a large butterfly on the glanceable display when the goal is met.

Another more complex way of goal setting is used in the exercise based systems described in (Oliver & Flores-Mangas 2006, Ståhl et al. 2008, Mulas et al. 2011).

These applications rely on the definition of an exercise schedule either by the user himself (Oliver & Flores-Mangas 2006) or by an expert through a connection with an inter-human interaction component (Mulas et al. 2011). Such schedules can be based on e.g. time spent in a specific heart-rate zone, user's running speed/duration, the performance of specific activities such as shopping (Ståhl et al. 2008), or a combination thereof. Currently none of the reviewed papers include a form of automatic generation of complex goals in the form of task-schedules.

2.6.5 User targeting

User targeting is the simplest form of tailoring in terms of technical complexity, but it is a technique not often used in the literature. Out of the 15 included papers, only two employ some form of user targeting. Buttussi & Chittaro (2008) require the user of their fitness coach to enter personal information — weight, height, gender and age — into the system before first use. From this information, using a so called auto-test, the maximum oxygen uptake is calculated based on a physical test and the user's gender and age. During use of the system, the user is given information on the amount of calories burnt which is calculated for the individual based on his/her weight. The user's gender and age are also taken into account when deciding when to give alerts or speed/intensity advices while performing exercises. This type of user targeting falls under the category of 'contextualization'. The second paper that describes some form of user targeting is by Bielik et al. (2012). In order to provide tailored motivation, the authors describe the use of a *user model* containing preferred activities for each individual. The authors state that these preferred activities are derived automatically over time (see Section 2.6.7: Self learning), but the details are unclear and it does not appear to be implemented. The idea however is that each individual receives recommendations about which physical activities could be performed in order to reach a daily goal, based on this set of preferred activities. This is another, clearly more advanced form of contextualization, that adapts the content of messages based on a user model. The authors also claim to take gender and age into account when determining an appropriate amount of activity, but it is again unclear how exactly this is done. Although it seems that these two methods of user targeting remain in the conceptual phase, the authors did perform an evaluation with another form of user targeting. Users, in this case children aged 12-13, would start the activity coach by choosing a nickname and were then able to customize a personal avatar that would be used later on for providing feedback. This technique, called *identification*, worked very intuitively for the children as it is common practice in online social services and games. Although its effects on physical activity performance was not evaluated, for the younger target populations, the ability to name and customize a personal avatar seems to be a logical and natural feature that could improve the connectedness between the user and the digital coach.

An overall commonality between the two author's usages of user targeting is in the implementation of an internal user model to keep track of application relevant features of the user. Figure 2.5 shows an example of a high level *user model* architecture that covers the examples from Buttussi & Chittaro (2008) and Bielik et al. (2012). The model consists of three high level types of information: **facts**, **preferences**, and **derived properties**. Facts are properties queried directly from the user such as age, gender or height; preferences contains slightly more fluid properties such as a chosen nickname, or visual features of an avatar, that are also explicitly given by the user, but could be changed over time. Derived properties contain dynamic user properties that are derived from the facts and preferences, as well as from continuous user input. This type of information includes e.g. a fitness level that is calculated after a specific self-test has been performed, but it also includes e.g. a statistically learned preference for a certain category of outdoor activities.

These various properties of the user model can then be used in varying degrees of transparency to the user. If e.g. calories burnt per hour of walking are calculated based on the user's weight, this is a form of user targeting that improves the quality of feedback that is given to the user. But it is also very unlikely that the user will actually know that this form of user targeting is taking place (low transparency). This is basically the opposite of raising expectation or 'placebo-tailoring' (see Section 2.5) in which the system would mention explicitly that advice is specifically targeted to the user. An example of such tailoring, also known as *name mentioning* is discussed in Dijkstra (2014) in this issue, where it is shown that the effects of name mentioning in a smoking cessation application is dependent on the person's individual stance on health in general (health value). Figure 2.6 shows some examples of user targeting techniques and transparencies.

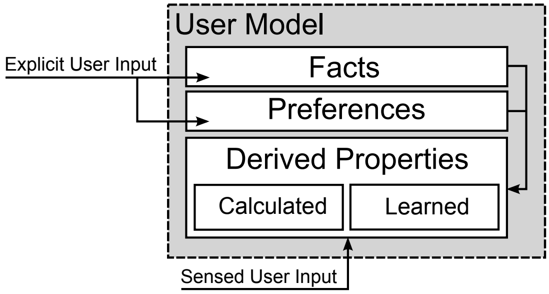


Figure 2.5: An example high level architecture of a user model containing facts, preferences (obtained explicitly from the user) and derived properties (learned, or calculated from facts, preferences and or sensed user input).

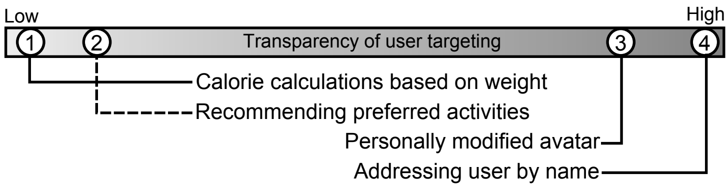


Figure 2.6: Four examples of user targeting techniques with varying degrees of visibility (or transparency) to the user.

2.6.6 Context awareness

Context Awareness is a tailoring technique used in two out of the 15 included studies. Ståhl et al. (2008) developed a fitness companion prototype system consisting of a mobile application and a home system. The mobile companion is developed as a spoken dialogue system, where the focus lies on spoken interaction with a virtual embodied agent (a 2D representation of a Nabaztag rabbit) that can be used during physical exercise such as running or walking. The mobile application is able to download a daily plan that is generated by the user's home system, consisting of various tasks like shopping or physical exercise activities. The companion is context aware in the sense that it uses location and time to suggest a suitable task for the user to perform. It is unspecified how this is implemented exactly, but one can imagine that the system does not suggest to the user to go shopping at a time when supermarkets would be closed, or suggest a user to go vacuum cleaning when the user is not at home.

The system described by Bielik et al. (2012) also applies context-awareness in their contextual information recommendation algorithms. The system uses, besides the User Model as described in Section 2.6.5, a Domain Model, containing information about e.g. the day of week, the month, and the weather. As the article mostly seem to describe a conceptual system, again not much details about the actual use of this information is given. The example that the authors give is not related to activity coaching, but to nutritional advice, where the system could recommend different drinks on *colder as opposed to warmer days or different meals at noon or in the evening*. But it is not hard to imagine the usage of the same type of contextual information for physical activity recommendations, as for example a recommendation to perform outdoor activities is less suitable when it is raining outside. Similarly, the system could define a daily activity goal based on the day of the week, allowing for example office workers to be less active during workdays, and make up for it on the weekends.

Note that in the discussion on context-awareness we do not consider the 'navigational cues' and 'location aware exercise demonstrations' given by the MOPET system (Buttussi et al. 2006, Buttussi & Chittaro 2008) as context aware features, as the navigation of users through fitness trails is the core function of their system. Similarly the paper by Bickmore et al. (2009) is excluded; although the title of the paper contains the term "Context Awareness", the only information that is sensed and used in the feedback is the user's physical activity, which we consider to be the **core** information, not the context.

2.6.7 Self learning

Self Learning is a tailoring technique where the system does not merely adapt to the individual, but automatically adapts to the changing individual in changing contexts of use. Out of all the included papers, only Bielik et al. (2012) mention self-learning techniques. The authors describe the application of a User Model, containing user activity preferences, which is built incrementally as users use the system. They also describe the use of educational quizzes that are meant to teach children the benefits of a healthy lifestyle. The results of the quizzes are stored in the User Model, and on subsequent selections of quizzes, questions in those areas in which the user has low results are selected with higher probability. As mentioned before, the work described by Bielik et al. seems to be in a conceptual phase, as their system evaluation describes covers only a small set of features described in the paper. The self-learning aspect is not included in those evaluations.

Self learning techniques in real-time physical activity coaching systems are much more sophisticated than the other possible tailoring techniques and as such, these techniques are not so ubiquitously used in the literature yet. Self-Learning can be seen as an upgrade for any other type of tailoring mechanism. For example, as a first step of tailoring a physical activity coach to its users, a designer can add context-aware activity recommendations: e.g. advising the user to go for a stroll through a nearby park, or go to the vegetable market when there is one currently taking place. Then, as an improvement to such a system, the system could keep track of the user's responses to given advice, and learn for example that this particular user does not like to simply go for a walk, but prefers to do only purposeful activities, and act accordingly.

2.7 Discussion

Although the fields of physical activity monitoring and tailoring are both extensive and richly illustrated with example systems, the intersection between the two fields where real time tailoring is employed (the focus of this article) yields a somewhat sparser set of examples. We identified only 15 articles which met the scope and search constraints defined in Section 2.3 in which real time tailoring for physical activity is discussed, covering between them a total of 12 different applications. From these articles, four describe only the use of feedback — the most obvious form of tailoring — and another four describe the additional use of one other tailoring concept (e.g. *feedback* and *inter-human interaction*, see Table 2.6). The only paper that describes the use of six of the seven tailoring concepts is mostly conceptual in nature (Bielik et al. 2012). This indicates that tailoring in this domain is a novel but emerging research area. As this research area matures, we hope to see more applications emerging that incorporate a variety of tailoring concepts based on sound theoretical foundations into the application design.

There exists a considerable body of theoretical literature on behavior change theories such as the Stages of Change (Prochaska & Velicer 1997), Theory of Planned Behavior (Ajzen 1991), Goal-Setting Theory (Locke & Latham 2002) as well as literature on e.g. persuasive technologies (Fogg 2003) and tailoring (Hawkins et al. 2008). Such theories provide the basis for our definitions of the tailoring concepts in Section 2.5. Throughout the survey we have given examples of how these theories can be applied to tailor a real-time physical activity coaching system. However, in our study of the 12 applications detailed in Section 2.4, we found that many of the articles failed to clearly specify a theoretical foundation for specific design decisions. For example, the automatic goal setting functionality described in Burns et al. (2011) is a textbook example of implementing the Goal-Setting Theory, but no motivation or reference regarding that theory is given. Instead, the motivation for application design is often based on similar applications in the domain. The citing of similar applications, and indeed eachother among the reviewed articles, seems to be more common than refering to theory. We believe that this issue arises due to the highly multidisciplinary nature of the field, as the development of a tailored motivational health tool requires the input from technicians, medical staff as well as people with a background in psychology.

In principal there is no issue in developing tailored behavior change applications based on previously demonstrated methods instead of established theory. The problem is that the effectiveness of the methods referred to are often not proven by the evaluation studies performed. The tailoring concepts that are implemented in the various systems can in general be described as intuitive, and from a technological point of view, the solutions are often not very complex. For example, *user targeted* feedback is provided in Buttussi & Chittaro (2008) by using a simple formula, taking into account user's characteristics; or activity goals provided in Burns et al. (2011) are given by averaging the activity of previous days and multiplying by a factor of 1.05. Although such forms of tailoring seem intuitive, their effectiveness is not proven. A common method of evaluation is to simply test the effectiveness of the application as a whole, making it difficult to draw conclusions regarding the individual tailoring methods implemented. As such, there is a need to demonstrate the effectiveness of tailoring in a more structured and controlled manner.

As the ultimate goal of physical activity coaching systems is generally to promote long-term behavior change, it is clear that evaluations are not a trivial component of the research in this field. Demonstrating the effectiveness of a coaching application would require longitudinal studies, and given the large differences between individuals (the premise for research in tailoring) also requires a large numbers of participants. Whenever such large-scale, long-term trials are feasible within the scope of the research, it would be prudent to perform exhaustive analysis on the available data. Given the high costs of such trials, it is necessary to ensure that every aspect of the intervention is regarded, including e.g. an analysis of the user's active engagement with the system, as discussed elsewhere in this issue (Bouvier et al. 2014). Although the focus of this work is on physical activity coaching applications, we believe that the problem pertains to a broader context. Often the *tailoring component* is just a piece of the puzzle in the development of an innovative ICT service, while the focus of evaluation is on the service as a whole. We believe that this is a principal burden for progress in the field of tailoring. Therefore it is essential to develop innovative evaluation strategies that can quickly and decisively demonstrate the effectiveness of isolated tailoring approaches for which the result can be translated to the domain in which the tailoring is applied.

The staged approach to evaluation posed by DeChant et al. (1996) seems a promising framework for evaluation that can aid in overcoming these issues. The framework, developed in the context of evaluating telemedicine applications, consists of four stages of evaluation, from small scale experimental studies that demonstrate technical efficacy, to larger trials that demonstrate overall system validity. Translated to tailoring research, the goal would be to first perform small scale, experimental lab-studies to evaluate the effectiveness of isolated tailoring concepts. The goal here would be to demonstrate that users recognize and enjoy the tailored interaction and feel that it actually matches their personal needs or preferences. Then, the system can be evaluated on higher level constructs. In small scale pilot studies, evaluation should focus on whether or not the tailored system increases compliance, and whether or not this effect lasts after a period of use. Only then, the highest level outcome effects (e.g. increase in physical activity) should be evaluated in larger scale trials involving more participants.

Future work in the field of tailored physical activity coaching systems should aim to address the major topics discussed here. In order to arrive at innovative methods of tailoring, multidisciplinary research teams should make use of the knowledge available in theoretical literature instead of referring to existing applications in the field. Many different paths of tailoring are currently unexplored. The tailoring model in Figure 2.4 points out many possible combinations of future tailoring approaches, and the discussion in Section 2.6 gives examples on all of the various tailoring concepts that can hopefully serve as a source of inspiration for future designers of these systems. There is especially room for the application of more advanced forms of tailoring such as context-awareness and self learning. Modern smartphones are able to collect a wealth of contextual information such as location, weather data and digital agenda's. They also contain the processing power necessary to perform powerful analysis using data mining or machine learning algorithms on the device itself. This combination gives the potential for a smartphone application to learn exactly how to deliver communication to the user in various contexts of use. Finally, innovative evaluation methods should be studied that can be used to quickly and cost-effectively demonstrate the merits of individual tailoring approaches.

2.8 Conclusion

In this work we have addressed the field of tailoring for real-time physical activity coaching systems. The survey was used to create a conceptual framework of tailoring. We identified seven key concepts: **feedback**, **inter-human interaction**, **adaptation**, **user targeting**, **goal setting**, **context awareness**, and **self learning**. Each of these concepts relate to each other, as well as to different properties of a *communication instance*. Earlier work that aimed to define the field of tailoring (Hawkins et al. 2008) has provided the basis for our framework, but lacked in covering emerging concepts like context awareness and self learning. The model we have developed (Figure 2.4) can be used for designers of activity coaching systems. By following different paths through the graph, different combinations of tailoring can be explored. The analysis of the included applications based on the various tailoring concepts (Section 2.6) provides useful examples of how these concepts are currently being applied. We addressed the state of the art and point out future research directions as (1) applying available theoretical constructs to develop new forms of tailoring, (2) exploring new combinations of tailoring, with a focus on more advanced tailoring processes, and (3) developing innovative, fast and effective methods of evaluating the merits of individual tailoring approaches in the context of ICT services as a whole.

Bibliography

Adams, J. & White, M. (2005), 'Why don't stage-based activity promotion interventions work?', *Health Education Research* **20**(2), 237–243.

Ahtinen, A., Isomursu, M., Mukhtar, M., Mäntyjärvi, J., Häkkilä, J. & Blom, J. (2009), Designing social features for mobile and ubiquitous wellness applications, *in* 'Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia', Cambridge, United Kingdom.

Ajzen, I. (1991), 'The theory of planned behavior', *Organizational Behavior and Human Decision Processes* **50**(2), 179–211.

Arteaga, S. M., Kudeki, M., Woodworth, A. & Kurniawan, S. (2010), Mobile system to motivate teenagers' physical activity, *in* 'Proceedings of the 9th International Conference on Interaction Design and Children', Barcelona, Spain.

Bailey, R. C., Olson, J., Pepper, S. L., Porszasz, J., Barstow, T. J. & Cooper, D. M. (1995), 'The level and tempo of children's physical activities: an observational study', *Medicine and Science in Sports and Exercise* **27**(7), 1033–1041.

Bandura, A. (1986), *Social foundations of thought and action: A social cognitive theory*, Prentice Hall.

Benyon, D., Hansen, P. & Webb, N. (2008), Evaluating Human-Computer Conversation in Companions, *in* 'Proceedings of the Fourth International Workshop on Human-Computer Conversation', Bellagio, Italy.

Bickmore, T. W., Caruso, L., Clough-Gorr, K. & Heeren, T. (2005), '"It's just like you talk to a friend" - Relational Agents for Older Adults', *Interacting with Computers, Special Issue on HCI and the Older Population* **17**(6), 711–735.

Bickmore, T. W., Mauer, D. & Brown, T. (2009), 'Context Awareness in a Handheld Exercise Agent', *Pervasive and Mobile Computing* **5**(3), 226–235.

Bickmore, T. W. & Picard, R. W. (2005), ‘Establishing and Maintaining Long-Term Human-Computer Relationships’, *ACM Transactions on Computer-Human Interaction* **12**(2), 293–327.

Bickmore, T. W., Schulman, D. & Sidner, C. (2013), ‘Automated interventions for multiple health behaviors using conversational agents’, *Patient Education and Counseling* **92**(2), 142–148.

Bielik, P., Tomlein, M., Krátky, P., Mitrík, v., Barla, M. & Bieliková, M. (2012), Move2Play: an innovative approach to encouraging people to be more physically active, *in* ‘Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium’, Miami, Florida, USA, pp. 61–70.

Bouten, C. V., Koekkoek, K. T. M., Verduin, M., Kodde, R. & Janssen, J. D. (1997), ‘A Triaxial Accelerometer and Portable Data Processing Unit for the Assessment of Daily Physical Activity’, *IEEE Transactions on Biomedical Engineering* **44**(3), 136–147.

Bouvier, P., Sehaba, K. & Lavoué, E. (2014), ‘A trace-based approach to identifying user’s engagement and qualifying their engaged-behaviours in interactive systems’, *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change (online)* .

Burns, P., Lueg, C. & Berkovsky, S. (2011), ActivMON: A Wearable Ambient Activity Display, *in* ‘Proceedings of the 6th Conference on Mobile und Ubiquitäre Informationssysteme (MMS2011)’, Kaiserslautern, Germany, pp. 11–24.

Burns, P., Lueg, C. & Berkovsky, S. (2012), Activmon: encouraging physical activity through ambient social awareness, *in* ‘Proceedings of the 2012 ACM annual Conference on Human Factors in Computing Systems’, Austin, Texas, USA, pp. 2363–2368.

Burns, P., Lueg, C. & Berkovsky, S. (2013), Colours That Move You: Persuasive Ambient Activity Displays, *in* ‘Lecture Notes in Computer Science, Proceedings of the 8th International Conference on Persuasive Technology’, Vol. 7822 of *Lecture Notes in Computer Science*, Sydney, Australia, pp. 27–32.

Buttussi, F. & Chittaro, L. (2008), ‘MOPET: A Context-Aware and User-Adaptive Wearable System for Fitness Training’, *Artificial Intelligence in Medicine* **42**(2), 153–163.

Buttussi, F., Chittaro, L. & Nadalutti, D. (2006), Bringing mobile guides and fitness activities together: a solution based on an embodied virtual trainer, *in* ‘Proceedings of the 8th conference on Human-computer interaction with mobile devices and services’, Espoo, Finland, pp. 29–36.

Buttussi, F., Chittaro, L. & Nadalutti, D. (2009), Filtering Fitness Trail Content Generated by Mobile Users, *in* ‘Proceedings of the 17th International Conference on User

Modelling Adaptation and Personalization (UMAP 2009)', Trento, Italy, pp. 441–446.

Byrne, D. & Nelson, D. (1965), 'Attraction as a linear function of proportion of positive reinforcements', *Journal of Personality and Social Psychology* **36**(6), 659–663.

Campbell, T., Ngo, B. & Fogarty, J. (2008), Game design principles in everyday fitness applications, *in* 'Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work', San Diego, California, USA, p. 249.

Choudhury, T., Consolvo, S., Harrison, B., Hightower, J., LaMarca, A., LeGrand, L., Rahimi, A., Rea, A., Bordello, G., Hemingway, B., Klasnja, P., Koscher, K., Landay, J. A., Lester, J., Wyatt, D. & Haehnel, D. (2008), 'The Mobile Sensing Platform: An Embedded Activity Recognition System', *IEEE Pervasive Computing* **7**(2), 32–41.

Colineau, N. & Paris, C. (2011), 'Motivating reflection about health within the family: the use of goal setting and tailored feedback', *User Modeling and User-Adapted Interaction* **21**(4-5), 341–376.

Consolvo, S., Everitt, K., Smith, I. & Landay, J. A. (2006), Design requirements for technologies that encourage physical activity, *in* 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems', Montréal, Québec, Canada, pp. 457–466.

Consolvo, S., Klasnja, P., McDonald, D. W. & Landay, J. a. (2009), Goal-setting considerations for persuasive technologies that encourage physical activity, *in* 'Proceedings of the 4th International Conference on Persuasive Technology', Claremont, California, USA.

Consolvo, S., McDonald, D. W. & Landay, J. A. (2009), Theory-Driven Design Strategies for Technologies that Support Behavior Change in Everyday Life, *in* 'Proceedings of the 27th International Conference on Human Factors in Computing Systems', Boston, Massachusetts, USA, pp. 405–414.

Consolvo, S., Mcdonald, D. W., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., Klasnja, P., Lamarca, A., Legrand, L., Libby, R., Smith, I. & Landay, J. A. (2008), Activity Sensing in the Wild: A Field Trial of UbiFit Garden, *in* 'Proceedings of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems', Firenze, Italy, pp. 1797–1806.

Cortellese, F., Nalin, M., Morandi, A., Sanna, A. & Grasso, F. (2009), Static and Dynamic Population Clustering in Personality Diagnosis for Personalized eHealth Services, *in* 'Proceedings of the 3rd International Workshop on Personalisation for e-Health, AIME2009', Verona, Italy, pp. 1–9.

de Oliveira, R. & Oliver, N. (2008), TripleBeat: Enhancing Exercise Performance with Persuasion, *in* 'Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services', Amsterdam, the Netherlands, pp. 255–264.

DeChant, H. K., Tohme, W. G., Mun, S. K., Hayes, Wendelin, S. & Schulman, K. A. (1996), 'Health Systems Evaluation of Telemedicine: A Staged Approach', *Telemedicine Journal: the Official Journal of the American Telemedicine Association* **2**(4), 303–312.

Dey, A. K. & Abowd, G. D. (1999), Towards a Better Understanding of Context and Context-Awareness, *in* 'Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing', Karlsruhe, Germany.

Di Tullio, E. & Grasso, F. (2012), A Model for a Motivational System Grounded on Value Based Abstract Argumentation Frameworks, *in* 'Electronic Healthcare - Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Volume 91', Malaga, Spain, pp. 43–50.

Dijkstra, A. (2014), 'The persuasive effects of personalization through name mentioning in a smoking cessation message', *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change (online)* .

Ekeland, A. G., Bowes, A. & Flottorp, S. (2010), 'Effectiveness of telemedicine: A systematic review of reviews', *International Journal of Medical Informatics* **79**(11), 736–771.

Enwald, H. P. K. & Huotari, M.-L. A. (2010), 'Preventing the Obesity Epidemic by Second Generation Tailored Health Communication: An Interdisciplinary Review', *Journal of Medical Internet Research* **12**(2), e24.

Erriquez, E. & Grasso, F. (2008), Generation of Personalised Advisory Messages: An Ontology Based Approach, *in* 'Proceedings of the 21st IEEE International International Symposium on Computer-Based Medical Systems (CMBS)', Jyvaskyla, Finland, pp. 437–442.

Festinger, L. (1957), *A Theory of Cognitive Dissonance*, Stanford University Press.

Fischer, G. (2001), 'User Modeling in Human-Computer Interaction', *User Modeling and User-Adapted Interaction* **11**(1), 65–86.

Fogg, B. (2003), *Persuasive Technology: Using Computers to Change What We Think and Do*, Morgan Kaufmann.

Fry, J. P. & Neff, R. A. (2009), 'Periodic Prompts and Reminders in Health Promotion and Health Behavior Interventions: Systematic Review', *Journal of Medical Internet Research* **11**(2), e16.

Fujiki, Y., Kazakos, K., Puri, C., Buddharaju, P. & Pavlidis, I. (2008), 'NEAT-o-Games : Blending Physical Activity and Fun in the Daily Routine', *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment* **6**(2), 1–22.

Fujiki, Y., Kazakos, K., Puri, C., Pavlidis, I., Starren, J. & Levine, J. (2007), NEAT-o-Games: Ubiquitous Activity-Based Gaming, *in* 'CHI '07 Extended Abstracts on Human Factors in Computing Systems', San Jose, California, USA, pp. 2369–2374.

Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I.-M., Nieman, D. C. & Swain, D. P. (2011), 'American College of Sports Medicine position stand. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise', *Medicine & Science in Sports & Exercise* **43**(7), 1334–1359.

Gerber, S., Fry, M., Kay, J., Kummerfeld, B., Pink, G. & Wasinger, R. (2010), PersonisJ: mobile, client-side user modelling, *in* 'Proceedings of the 18th International Conference on User Modelling Adaptation and Personalization (UMAP 2010)', Vol. 6075, Big Island of Hawaii, USA, pp. 111–122.

Goffmann, E. (1959), *The Presentation of Self in Everyday Life*, Anchor Books.

Hawkins, R. P., Kreuter, M. W., Resnicow, K., Fishbein, M. & Dijkstra, A. (2008), 'Understanding tailoring in communicating about health', *Health Education Research* **23**(3), 454–466.

Ijsselsteijn, W., De Kort, Y., Westerink, J., De Jager, M. & Bonants, R. (2004), Fun and Sports: Enhancing the Home Fitness Experience, *in* 'Proceedings of the 3rd International Conference on Entertainment Computing (ICEC2004)', Vol. 3166, Eindhoven, the Netherlands, pp. 46–56.

Kennedy, C. M., Powell, J., Payne, T. H., Ainsworth, J., Boyd, A. & Buchan, I. (2012), 'Active Assistance Technology for Health-Related Behavior Change: An Interdisciplinary Review', *Journal of Medical Internet Research* **14**(3), e80.

King, A. C., Hekler, E. B., Grieco, L. A., Winter, S. J., Sheats, J. L., Buman, M. P., Banerjee, B., Robinson, T. N. & Cirimele, J. (2013), 'Harnessing Different Motivational Frames via Mobile Phones to Promote Daily Physical Activity and Reduce Sedentary Behavior in Aging Adults', *PloS one* **8**(4), e62613.

Klasnja, P., Consolvo, S., McDonald, D. W., Landay, J. A. & Pratt, W. (2009), Using Mobile & Personal Sensing Technologies to Support Health Behavior Change in Everyday Life: Lessons Learned, *in* 'AMIA Annual Symposium Proceedings', Vol. 2009, San Francisco, California, USA, pp. 338–342.

Kohl, H. W., Fulton, J. E. & Caspersen, C. J. (2000), 'Assessment of Physical Activity among Children and Adolescents: A Review and Synthesis', *Preventive Medicine* **31**(2), 54–76.

Kreuter, M. W., Strecher, V. J. & Glassman, B. (1999), 'One size does not fit all: the case for tailoring print materials', *Annals of Behavioral Medicine* **21**(4), 276–283.

Lane, N. D., Xu, Y., Lu, H., Hu, S., Choudhury, T., Campbell, A. T. & Zhao, F. (2011), Enabling Large-scale Human Activity Inference on Smartphones using Community Similarity Networks (CSN), *in* 'Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11', Beijing, China, pp. 355–364.

Lin, M., Lane, N. D., Mohammod, M., Yang, X., Lu, H., Cardone, G., Ali, S., Doryab, A., Berke, E., Campbell, A. T. & Choudhury, T. (2012), BeWell+: Multi-dimensional Wellbeing Monitoring with Community-guided User Feedback and Energy Optimization, *in* 'Proceedings of the Conference on Wireless Health - WH '12', San Diego, California, USA, pp. 1–8.

Locke, E. A. & Latham, G. P. (2002), 'Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey', *American Psychologist* **57**(9), 705–717.

Lombard, D. N., Lombard, T. N. & Winett, R. A. (1995), 'Walking to Meet Health Guidelines: The Effect of Prompting Frequency and Prompt Structure', *Health Psychology. Official Journal of the Division of Health Psychology American Psychological Association* **14**(2), 164–170.

Lustria, M. L. A., Cortese, J., Noar, S. M. & Glueckauf, R. L. (2008), 'Computer-tailored health interventions delivered over the Web: Review and analysis of key components', *Patient Education and Counseling* **74**(2), 156–173.

Lutes, L. D. & Steinbaugh, E. K. (2010), 'Theoretical models for pedometer use in physical activity interventions', *Physical Therapy Reviews* **15**(3), 143–153.

Maitland, J., Sherwood, S., Barkhuus, L., Anderson, I., Hall, M., Brown, B., Chalmers, M. & Muller, H. (2006), Increasing the Awareness of Daily Activity Levels with Pervasive Computing, *in* 'Proceedings of the 1st International Conference on Pervasive Computing Technologies for Healthcare', Innsbruck, Austria, pp. 1–9.

McKenzie, T. L. (2002), The use of direct observation to assess physical activity, *in* G. Welk, ed., 'Physical Activity Assessments for Health Related Research', Human Kinetics Publishers, Inc., pp. 179–195.

Mokka, S., Väätänen, A., Heinilä, J. & Välkkynen, P. (2003), Fitness Computer Game with a Bodily User Interface, *in* 'Proceedings of the 2nd International Conference on Entertainment Computing', ICEC '03, Pittsburgh, Pennsylvania, USA, pp. 1–3.

Morandi, A. & Serafin, R. (2007), A personalized motivation strategy for physical activity promotion in diabetic subjects, *in* 'Proceedings of the 2nd Workshop on Personalisation for E-Health', Corfu, Greece, pp. 51–55.

Mulas, F., Carta, S., Pilloni, P. & Manca, M. (2011), Everywhere Run: a virtual personal trainer for supporting people in their running activity, *in* 'Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology - ACE '11', Lisbon, Portugal.

Mulas, F., Pilloni, P. & Carta, S. (2012), Everywhere Race!: A Social Mobile Platform for Sport Engagement and Motivation, *in* 'Proceedings of the 2nd International Conference on Social Eco-Informatics', Venice, Italy, pp. 63–69.

Noar, S. M., Benac, C. N. & Harris, M. S. (2007), 'Does tailoring matter? Meta-Analytic Review of Tailored Print Health Behavior Change Interventions', *Psychological Bulletin* **133**(4), 673–693.

O'Brien, S. & Mueller, F. F. (2007), Jogging the Distance, *in* 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07', San Jose, California, USA, pp. 523–526.

Oinas-Kukkonen, H. & Harjumaa, M. (2009), 'Persuasive Systems Design: Key Issues, Process Model, and System Features', *Communications of the Association for Information Systems* **24**(1), 485–500.

Oliver, N. & Flores-Mangas, F. (2006), MPTrain: A Mobile, Music and Physiology-Based Personal Trainer, *in* 'Proceedings of the 8th Conference on Human Computer Interaction with Mobile Devices and Services', Espoo, Finland, pp. 21–28.

Oliver, N. & Kreger-Stickles, L. (2006*a*), Enhancing Exercise Performance through Real-time Physiological Monitoring and Music: A User Study, *in* 'Pervasive Health Conference and Workshops', Innsbruck, Austria, pp. 1–10.

Oliver, N. & Kreger-stickles, L. (2006*b*), PAPA: Physiology and Purpose-Aware Automatic Playlist Generation, *in* 'Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR2006)', Victoria, Canada, pp. 250–253.

op den Akker, H., Jones, V. M. & Hermens, H. J. (2010), Predicting Feedback Compliance in a Teletreatment Application, *in* 'Proceedings of ISABEL 2010: the 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies', Rome, Italy.

Orji, R., Vassileva, J. & Mandryk, R. L. (2014), 'Modeling the Efficacy of Persuasive Strategies for Different Gamer Types in Serious Games for Health', *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change (online)* .

Plasqui, G., Bonomi, A. G. & Westerterp, K. R. (2013), 'Daily physical activity assessment with accelerometers: new insights and validation studies', *Obesity Reviews* **14**(6), 451–462.

Prochaska, J. O. & DiClemente, C. (1986), 'Toward a Comprehensive Model of Change', *Applied Clinical Psychology* **13**(1), 3–27.

Prochaska, J. O. & Velicer, W. (1997), 'The Transtheoretical Model of Health Behavior Change', *American Journal of Health Promotion* **12**(1), 38–48.

Qian, H., Kuber, R. & Sears, A. (2010), Maintaining Levels of Activity using a Haptic Personal Training Application, *in* 'CHI '10 Extended Abstracts on Human Factors in Computing Systems', Atlanta, Georgia, USA, pp. 3217–3222.

Qian, H., Kuber, R. & Sears, A. (2011), 'Towards developing perceivable tactile feedback for mobile devices', *International Journal of Human-Computer Studies* **69**(11), 705–719.

Qian, H., Kuber, R., Sears, A. & Murphy, E. (2011), 'Maintaining and modifying pace through tactile and multimodal feedback', *Interacting with Computers* **23**(3), 214–225.

Saini, P. & Lacroix, J. (2009), Self-setting of Physical Activity Goals and Effects on Perceived Difficulty, Importance and Competence, *in* 'Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09', Claremont, California, USA, p. 1.

Sallis, J. F. & Saelens, B. E. (2000), 'Assessment of physical activity by self-report: status, limitations, and future directions', *Research Quarterly for Exercise and Sport* **71**(4), S1–S14.

Saris, W. H. M. & Binkhorst, R. A. (1977), 'The Use of Pedometer and Actometer in Studying Daily Physical Activity in Man. Part I: Reliability of Pedometer and Actometer', *European Journal of Applied Physiology and Occupational Physiology* **37**(3), 219–228.

Short, C. E., James, E. L., Plotnikoff, R. C. & Girgis, A. (2011), 'Efficacy of tailored-print interventions to promote physical activity: a systematic review of randomised trials.', *The International Journal of Behavioral Nutrition and Physical Activity* **8**(1), 113.

Sohn, M. & Lee, J. (2007), UP Health: Ubiquitously Persuasive Health Promotion with an Instant Messaging System, *in* 'CHI '07 Extended Abstracts on Human Factors in Computing Systems', San Jose, California, USA, pp. 2663–2668.

Speakman, J. R. (1998), 'The history and theory of the doubly labeled water technique', *The American Journal of Clinical Nutrition* **68**(4), 932S–938S.

Stå hl, O., Gambäck, B., Hansen, P., Turunen, M. & Hakulinen, J. (2008), A Mobile Fitness Companion, *in* 'Proceedings of the Fourth International Workshop on Human-Computer Conversation', Bellagio, Italy.

Thaler, R. H. & Sunstein, C. R. (2008), *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale University Press.

Turunen, M., Hakulinen, J., Stå hl, O., Gambäck, B., Hansen, P., Rodríguez Gancedo, M. C., de la Cámara, R. S., Smith, C., Charlton, D. & Cavazza, M. (2011), 'Multimodal and mobile conversational Health and Fitness Companions', *Computer Speech and Language* **25**(2), 192–209.

Wang, Y., Burgener, D., Kuzmanovic, A. & Macia-Fernandez, G. (2011), Understanding the Network and User-Targeting Properties of Web Advertising Networks, *in* 'Proceedings of the 31st International Conference on Distributed Computing Systems', Minneapolis, Minnesota, USA, pp. 613–622.

Yim, J. & Graham, T. C. N. (2007), Using games to increase exercise motivation, *in* 'Proceedings of the 2007 Conference on Future Play', Toronto, Canada, pp. 166–173.

2

Chapter 3

Reaching Kairos: Adaptive Prediction of the Opportune Moment for Stimulating Physical Activity

Harm op den Akker, Valerie M. Jones and Hermie J. Hermens (Submitted).

Harm op den Akker, Valerie M. Jones and Hermie J. Hermens, in *Proc ISABEL2010*.

Abstract

Achieving and maintaining a healthy lifestyle, particularly in terms of daily-life physical activity, is receiving growing scientific, societal and commercial interest. Sensors which can be used to accurately assess a user's activity behavior throughout the day have become smaller and cheaper and can connect wirelessly to increasingly powerful smartphones. This combination of emerging technologies enables the development of ubiquitous and intelligent coaching applications that aim to guide the user in achieving sustainable behavior change. These applications can accompany the user throughout the day and have the potential to provide tailored coaching whenever most appropriate. To identify the most appropriate *timing* for feedback is the problem addressed in this article. We describe the development and evaluation of a physical activity coaching application which automatically identifies the most opportune moment for providing coaching to the individual user. The smartphone application constantly evaluates the context of the user and uses machine learning to classify *situations* as suitable moments for delivery of motivational coaching. The first phase — the cold start phase, in which no data for the current user is known — uses a κ-nearest-neighbor classifier ($\kappa = 25$) trained on historical data. In the second phase, real-time prediction and self-learning takes place by means of a Support Vector Machine implementation which is automatically generated on the smartphone and which is re-trained periodically. The system was evaluated by means of a three-month longitudinal trial with 10 Chronic Obstructive Pulmonary Disease (COPD) patients. Results show a significant increase in compliance to opportune motivational cues compared to earlier trials involving COPD patients. In the period between receiving the 41^{st} and 105^{th} cue, user compliance to the system is on average 62.70% versus 56.00% for a historical COPD dataset (individual p-values < 0.05).

3.1 Introduction

Due to the aging population, the prevalence of chronic diseases is increasing worldwide. The growing demand on healthcare services calls for cost-effective treatments that reduce the demands on healthcare professionals. Provision of eHealth and telemedicine services, in particular technology mediated services which stimulate and support patient's self-efficacy, is a fast growing field of research (Ekeland et al. 2010) and is widely regarded as a promising paradigm to reduce the burden on the healthcare system. An important factor in the prevention and treatment of chronic diseases and healthy aging support is the maintenance of a healthy lifestyle in terms of regular physical activity. Lee et al. investigated the effects of physical inactivity on a number of major non-communicable diseases worldwide and estimate that physical inactivity causes 9% of premature mortality and that if completely eliminated would increase the world's life expectancy by 0.68 years (Lee et al. 2012). Deaths from the major noncommunicable diseases are projected to rise from 36 million in 2008 to 52 million in 2030 (World Health Organization 2013).

One of the major noncommunicable diseases in the western world is Chronic Obstructive Pulmonary Disease (COPD), a respiratory condition characterized by a non-reversible, progressive limitation of airflow in the lungs. Treatment of COPD is focused on breaking through patterns of sedentary behavior (fear induced movement avoidance) by stimulating patients to perform regular physical activity, while teaching them to recognize limits incurred by the disease (Cooper 2009). Optimal disease management for COPD includes a system of patient self-management and promotion of a healthy, active lifestyle (Spruit et al. 2013). Current treatment consists predominantly of prescribed physical exercises by physiotherapists. These treatment programs fail to induce sustained changes in physical activity behavior, mainly due to the lack of incorporation into everyday life. Telemedicine or eHealth has the potential to accommodate sustainable behavior change by monitoring and providing intensive coaching on physical activity during day-to-day activities of not only COPD patients, but everyone who wants to maintain a healthy active lifestyle.

Over the past years, a telemedicine service to promote sustainable physical activity behavior change was designed, implemented and evaluated in several different studies, e.g. (Dekker-van Weering et al. 2012, Tabak et al. 2014). Patients were given a 3D-accelerometer based sensor to assess daily activity patterns, combined with a smartphone for providing continuous visual feedback in the form of a graph. By comparing the patient's daily activity, measured in IMA[1]-counts per minute (Bouten et al. 1996), to a predefined reference activity pattern, the patient is automatically given motivational cues based on his performance at regular intervals throughout the day (see Figure 3.1). In this version of the system, the patients were set to receive cues at fixed times (e.g. once every hour), and the messages were selected randomly from a list.

The evaluation of this system has shown its effectiveness in stimulating physical activity (Dekker-van Weering et al. 2012, Tabak et al. 2014), but there are two major issues that need to be solved: (1) the increase in physical activity tends to not persist long term, and (2) the response to the motivational cues has a large variance across subjects and studies. The differing responses to the feedback mechanisms employed can be attributed to the high degree of heterogeneity within the target populations. It is widely believed that personalization — or tailoring — can help in increasing the effectiveness of technologies that aim to achieve behavior change.

In this article we describe the design, implementation and evaluation of a tailored motivational cue component which was added as an additional module to an existing remote monitoring and coaching service. The system, which we call *Kairos* is the module in the smartphone application that learns how to find the opportune moments for providing motivational cues to the user. This is done by constantly predicting whether or not a user will comply to a motivational cue by learning from the response of the user to past cues (in terms of activity). In this way, Kairos is tailored to each individual user and adapts its reasoning over time as the system is being used. In order to evaluate the performance of the system, we focus here on a physical activity behavior change application targeted specifically at COPD patients. However, regular physical activity is beneficial for everyone. The current advice is that all adults should perform physical exercise corresponding to thirty

[1]IMA: time Integral of the Modulus of Accelerometer output (Bouten et al. 1996).

Figure 3.1: Screenshots of the activity coaching application. On the left: cumulative physical activity plotted over a reference activity line indicating the user's goal between 08:00 and 22:00. On the right: example of an encouraging motivational cue.

minutes of moderate-intensity cardio-respiratory exercise training for at least five days per week (Garber et al. 2011). As such, the practical use of the application is aimed at a broader audience.

The article is structured as follows. Section 3.2 summarizes related work in the field of adaptive feedback on physical activity. The design and implementation of the Kairos module is presented in the three subsequent sections. Section 3.3 explains the general principles that are applied for predicting motivational cue compliance and contains a high-level overview of the functionality of the Kairos module. Sections 3.4 and 3.5 explain the two different phases of operation for the module, the cold start phase and the adaptive phase, respectively. Section 3.6

describes the evaluations of the system that were performed with COPD patients, focusing on the performance and the adaptive nature of the system. A general discussion of the work is presented in Section 3.7 and conclusions and future recommendations in Section 3.8.

3.2 Related work

In 2008, Hawkins et al. defined tailoring as "*any of a number of methods for creating communications individualized for their receivers, with the expectation that this individualization will lead to larger intended effects of these communications*" (Hawkins et al. 2008). The authors break down tailoring into three separate strategies: 'personalization', 'feedback', and 'content matching', covering mostly the tactics used in earlier forms of health content delivery such as using tailored print (Skinner et al. 1999, Noar et al. 2007, Boudreau et al. 2011). Tailoring is in general considered to be a promising tool for increasing the effectiveness of telemedicine or eHealth applications (Brug et al. 1999, Kroeze et al. 2006, Sohl & Moyer 2007). A study by Bull et al., (Bull et al. 1999) showed that patients who received tailored information on their physical activity behavior were most likely to improve, compared to patients receiving more general information. Similar results were seen in a study by Campbell et al. (Campbell et al. 1994), where computer-tailored feedback was twice as effective in improving patient's dietary habits, compared to general feedback. Although tailoring seems intuitively useful and has been shown to be effective, it is not yet widely employed in real-time physical activity coaching systems.

Recent advances in smartphones and wearable sensor technology have enabled the use of rich, real-time coaching mechanisms. An increase in processing power and available sensors on mobile devices opened up possibilities of detailed sensing and reasoning on contextual data for individual users, giving rise to more advanced tailoring possibilities. In (op den Akker, Jones & Hermens 2014) we provide an updated definition and model of tailoring that includes more advanced methods such as 'context awareness' and 'self learning'. The tailoring model provided in (op den Akker, Jones & Hermens 2014) is based on a survey of the field of tailored, real-time physical activity coaching and includes a discussion on 12 different applications.

The timing of communicating motivational coaching messages is an issue that has not been given much attention. For those applications in which the system takes the initiative in communicating feedback to the user, the moment of delivery is often chosen based on a single contextual factor: the current physical activity level. For example, in the haptic personal trainer (Qian et al. 2011), the system warns the user with vibrations to indicate that he or she should walk faster or slower than an individual baseline, or comfortable speed. Similarly, the Handheld Exercise Agent by Bickmore et al. (Bickmore et al. 2009) provides system initiated positive reinforcement after the system detected that the user has walked for 10 consecutive minutes or more. These applications use the context of the user (their physical activity behavior) to determine the moment of feedback delivery, and so, tailor their communication. However, the level of tailoring is very low, as the "contextual" data used is limited to a single type of information, and is solely based on thresholds.

For additional background we direct the reader to our extensive survey on tailoring in real-time physical activity coaching applications in (op den Akker, Jones & Hermens 2014). Related reviews are presented by e.g. Fry and Neff on periodic prompts for encouraging physical activity (Fry & Neff 2009), or by Kennedy et al. on active assistance technology (Kennedy et al. 2012).

3.3 Methods

The overall objective of this work is to increase long-term adherence to physical activity interventions delivered through an ambulatory activity monitoring and feedback application running on a smartphone. In the work reported here we focus specifically on the timing of motivational cues. These text-based cues help to increase the user's awareness of his own activity and contain suggestions for tasks to perform in order to balance activity levels. These suggestions can either be *encouraging* (when the user is under-performing, the system ***encourages*** physical activity), *discouraging* (when the user is over-performing, the system ***discourages*** activity), or *neutral* (neither encouraging nor discouraging activity, in case the user is within 10% of his current goal).

Previous studies have demonstrated the added value of providing such motivational cues in addition to providing real-time feedback presented as a continuous graph. In (Dekker-van Weering et al. 2012) the authors analyzed the activity levels of patients suffering from chronic low back pain in 30 minute intervals before and after a motivational cue. They found that patients indeed decreased their activity level significantly after a discouraging message (-17%, $p < 0.001$) and increased significantly after an encouraging message (+29%, $p < 0.049$). A more detailed analysis was done in (Tabak et al. 2014) focusing on COPD patients. The authors found that the highest patient response to discouraging cues was 30 minutes after the message (-29%, $p < 0.001$), while for encouraging cues, this was 5 minutes after the motivational cue (+23%, $p = 0.005$). The neutral messages provoked no significant response, as can be expected.

In order to address the overall objective of increasing long-term adherence, this work investigates whether *compliance* to individual motivational cues can be increased by selecting an opportune moment for delivery of the cue as opposed to time-based cues (i.e. delivery at regular intervals). With *opportune* we mean here the moment at which the expected compliance to the cue is largest (this moment is defined by a contextual feature vector, see Table 3.3). Our definition of compliance to a motivational cue is given in equation 3.1, where $\Delta_{pre}(c)$ is the amount of activity performed in the 30-minute interval before the cue and $\Delta_{post}(c)$ the activity in the 30-minute interval after the cue. In short, the user complies to cue c, if the intention of the cue corresponds to the observed effects.

$$\begin{aligned}&\text{The client is compliant to motivational cue c if:}\\&\Delta_{pre}(c) < \Delta_{post}(c) \text{ , if c is of type encouraging}\\&\Delta_{pre}(c) > \Delta_{post}(c) \text{ , if c is of type discouraging}\end{aligned} \tag{3.1}$$

The solution presented here is based on a machine learning approach, where the system automatically predicts the compliance for hypothetical cueing moments. Earlier work on this topic presented in (op den Akker et al. 2010) demonstrated the theoretical feasibility of this approach; showing that compliance can be predicted with an average accuracy of 86.03% ($\pm 4.09\%$) over 38 subjects using a simple set of contextual features. We first describe the overall functionality and architecture of the activity monitoring and coaching application (Section 3.3.1). Section 3.3.2 describes the application flow for generating motivational cues, and section 3.3.3 provides the details of the initialization and training phase in which classifier selection and self-learning takes place.

3.3.1 Overall system architecture

The Kairos feedback timing prediction module is implemented in an existing activity monitoring and coaching platform. The coaching platform consists of sensors, smartphones, a server and web portals. The sensor, a ProMove-3D activity tracker (Bosch et al. 2009) developed by Inertia Technology, uses a 3D accelerometer to track the user's movement throughout daily life. The raw accelerometer values are converted on the sensor-node into IMA values, a unit that is shown to correlate to energy expenditure (Bouten et al. 1996). The data is sent to the smartphone over Bluetooth every 10 seconds to allow real-time coaching on the phone. The smartphone application uses the mobile network or WiFi to synchronize data to a central server. The server takes care of data storage, user accounts and access control, and provides APIs for delivering data to various web portals. Details regarding the requirements, design, development and evaluation of the platform are given in (op den Akker et al. 2012).

This work is focused on improving the coaching delivered through the smartphone. This is achieved by implementing the Kairos system as a stand-alone module, integrated into our smartphone framework. Figure 3.2 shows the modular architecture of this Android-based smartphone application. Each module pictured provides autonomous functionality to the system, and runs in its own Java thread. Communication between modules can occur through *data streams*, for continuous data (e.g. from *BluetoothModule* to *ProMove3DReader*), or through message passing via a central *Hub* module (not pictured). The arrows in Figure 3.2 represent the most important communication between the modules. The BluetoothModule handles general Bluetooth connections with external devices. The ProMove3DReader implements the communication protocol for the ProMove3D activity. IMA values coming from the sensor are stored in the IMADataModule, that also contains *reference* values. These reference values are the target/goal

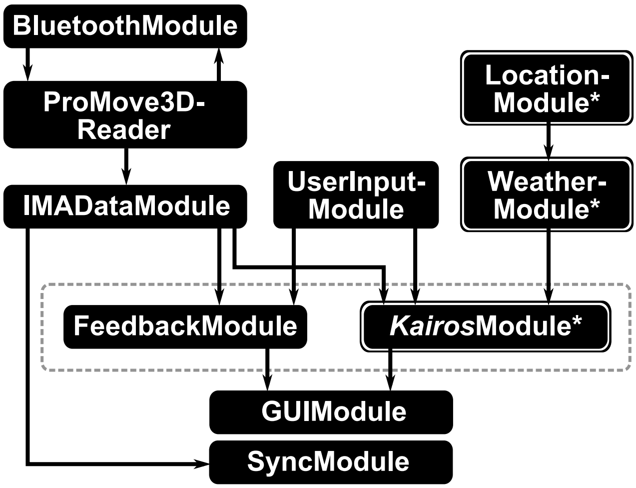


Figure 3.2: Modular architecture of the Android-based smartphone application for activity monitoring and coaching. In this work, the FeedbackModule is replaced by the *Kairos*Module. Additionally, the LocationModule and WeatherModule are added to the system.

values representing how active the user should be. In the Kairos evaluation, these reference values are set based on the baseline measurements of the patient, and re-adjusted weekly by a physiotherapist (Tabak, op den Akker, Vollenbroek-Hutten & Hermens 2014). The UserInputModule can be configured to trigger interactions from the system to the user, such as showing a motivational cue (or feedback message), or presenting a questionnaire to the user. The FeedbackModule, once triggered by a request, will pick a motivational cue, and pass it to the GUIModule to display on the user interface.

In this work, the FeedbackModule is replaced by the KairosModule to generate motivational cues in a more intelligent way. In this new version of the system, the UserInputModule is set to trigger feedback once every minute (this mechanism is kept in place for compatibility reasons). However, the KairosModule will for each feedback request determine itself whether or not to honour it. Two additional new modules are added to the system: the LocationModule provides broad location information (based on the cellular network) in order for the WeatherModule to retrieve location-based weather information as input for the KairosModule.

3.3.2 Generating motivational cues

Figure 3.3 shows an abstract component view of the Kairos module in which the numbered arrows indicate the sequential flow through the various steps of operation.

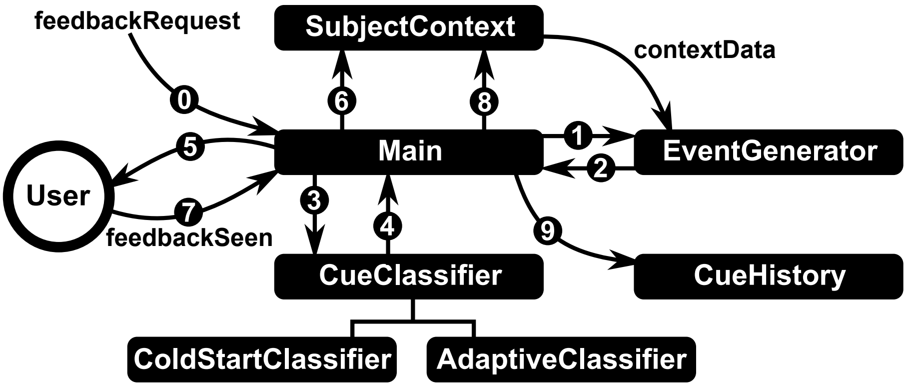


Figure 3.3: Flowchart representing the process of motivational cue generation in the smartphone application.

The Main process receives requests to send a cue from the application's event scheduler every minute (0). A cue-request-message contains information on the user's deviation between the current and the desired physical activity level. This information is passed on to the EventGenerator (1) that generates the "hypothetical" cue-event that could be given at that moment in time.

Example: *the user's activity level is currently 16% below his predefined reference line. Based on this information the EventGenerator generates an encouraging message, where the content is picked randomly from a set of predefined messages which have the same intention but are expressed differently in order to reduce habituation effects.*

The cue-event is enriched with a number of contextual parameters that it retrieves from the SubjectContext, and is then sent back to the Main process (2). Subsequently, the CueClassifier — which can either be an AdaptiveClassifier or a ColdStartClassifier — is used to predict the compliance to the hypothetical cue-event (3). The predicted compliance (`true` or `false`) is sent back to the Main process (4). If the predicted compliance is false, the process stops. If the predicted compliance is true, the motivational cue is sent to the application's Graphical User Interface for presentation to the user (5).

Example: *the message enriched by context information makes up a feature vector, such as (message: "Please go for a walk", time: "13:42:44", weather: "sunny", temperature: "18"). The classifier is trained on this type of data to predict whether or not the user would comply to this message. Based on the classifier's decision, the user will or will not receive the message on the smartphone screen.*

The cue-event including its contextual parameters and predicted compliance value is sent to the SubjectContext (6) for further processing. As soon as the user of the application reads the motivational message, a notification is sent to the Main process (7) and passed on to the SubjectContext (8). Thirty minutes after the message was read by the user, the cumulative activity of the user of this past half hour is retrieved and used for calculating whether or not the user actually complied with the cue (according to Equation 3.1). All information is then stored in the CueHistory (9), which writes the data to a file on the smartphone.

Example: *the system generates the message "Please go for a walk". After some time the user notices it, reads the message, and confirms by pressing OK. Thirty minutes after this moment, the user's response to the message is evaluated and stored for future training of the classifiers.*

Using this algorithm it is theoretically possible for the user to receive a motivational cue every minute. As this is not a desired behavior for the system, a manual check is included to prevent messages from being generated within 30 minutes of each other. The duration of this interval is set to 30 minutes because this is the length of the interval we use for calculating compliance.

3.3.3 Initialization and classifier selection

Each time the smartphone application is started — in the experimental setup, every morning of a measurement day — a number of initialization steps have to take place before the module can start with the generation of motivational cues as described above. As depicted in Figure 3.3, the system can either use an AdaptiveClassifier or a ColdStartClassifier, depending on whether enough user specific data has been collected. Furthermore, if enough data is available, the real-time classifier has to be retrained periodically. Algorithm 1 below describes this procedure in pseudo code.

All the motivational cues that have been presented to the user, and for which actual compliance has been calculated are loaded from the disk (line 1). If the total number of available cues is smaller than 25 (this value is explained in Section 3.5) the pre-generated cold start classifier is loaded from disk, and initialization is complete (line 3-5). If enough data is available, the system attempts to load a previously stored version of the real-time classifier from the disk (line 6). If this is the first time the threshold of 25 is reached, and the classifier does not exist yet, a new classifier is created and stored to disk (line 7-9).

Every time the application initializes with the real-time classifier, the classifier is retrained using all available data. This can be done in two ways. A light retrain retrains the classifier keeping the feature set unchanged, while a heavy retrain also starts a new feature selection phase (both are explained in more detail in Section 3.5). Each time a heavy retrain takes place, a counter is incremented and stored with the classifier (line 13). The counter is retrieved (line 10), and if the number of available motivational cues exceeds a new multiple of 25, a new heavy retrain is initialized (line 11-12). Otherwise, a light retrain is performed on the classifier (line

```
1  cueHistory ← loadCueHistory();
2  totalMotivationalCues ← cueHistory.size();
3  if totalMotivationalCues < 25 then
4  |   classifier ← loadColdStartClassifier();
5  else
6  |   classifier ← loadAdaptiveClassifier();
7  |   if classifier does not exist then
8  |   |   classifier ← new AdaptiveClassifier();
9  |   end
10 |   heavyRetrainCount ← classifier.getHeavyRetrainCount();
11 |   if totalMotivationalCues ≥ ((heavyRetrainCount +1) ×25) then
12 |   |   classifier.doHeavyRetrain(cueHistory);
13 |   |   classifier.setHeavyRetrainCount(heavyRetrainCount +1);
14 |   else
15 |   |   classifier.doLightRetrain(cueHistory);
16 |   end
17 end
```

Algorithm 1: Initialization procedure of the KairosModule.

14), and initialization completes. In short, every time 25 new instances become available, the feature selection procedure is run again.

3.3.4 Datasets

The core components of the Kairos system described above are the two classifiers: the cold start classifier (Section 3.4) and the real-time classifier (Section 3.5). Sections 3.4 and 3.5 describe the research done on developing these two classifiers based on activity and feedback data that was collected between October 2008 and June 2011. Five different datasets were used for the prediction of feedback compliance: two datasets from monitoring of chronic fatigue syndrome (CFS) patients, one dataset from monitoring of chronic low back pain (CLBP) patients, one dataset from chronic obstructive pulmonary disease (COPD) patients, and one from a study of obese subjects and healthy controls:

CFS RCT1 — The CFS RCT1 dataset consists of activity data from 75 patients suffering from chronic fatigue syndrome, divided into an intervention group — receiving motivational cues — and a control group. Data was collected during a randomized controlled trial between November 2008 and January 2011 in the Netherlands. Overall results of the study are described in (Evering 2013).

CFS RCT2 — The CFS RCT2 dataset consists of activity data and motivational cue event logs from 28 patients suffering from CFS (Evering 2013). In this randomized controlled trial, patients were divided into two groups receiving an activity target based on two different methods. The data was collected between January 2010 and April 2011.

CLBP — The CLBP dataset contains activity and motivational cue data from 17 patients suffering from chronic low back pain. This dataset does not contain a control group and every subject received motivational cues. The data was collected between October 2008 and December 2009 and is described in more detail in (Dekker-van Weering et al. 2012).

COPD RCT — The COPD RCT dataset consists of activity and motivational cue data from 16 patients suffering from chronic obstructive pulmonary disease. The dataset contains data from the intervention arm of a randomized controlled trial and is collected between October 2010 and June 2011. The RCT is described in (Tabak et al. 2013) and details on the intervention arm specifically are given in (Tabak et al. 2014).

Obesity — The Obesity dataset contains activity- and feedback data from obese subjects as well as activity data from healthy subjects. The dataset contains a total of 40 subjects and was collected between October 2008 and January 2009 as part of a student project.

An overview of the datasets is given in Table 3.1. Besides describing different patient populations, all datasets were gathered using the same measurement device and smartphone application. Differences in the data collection procedure relate to

	General		*Motivational Cues*					
Dataset	**N**	**Days**	**N**	**Days**	**Enc**	**Neu**	**Dis**	**Tot**
CFS RCT1	75	1444	32	446	572	1620	495	2687
CFS RCT2	28	516	25	251	425	832	183	1440
CLBP	17	322	17	153	101	1055	615	1771
COPD RCT	16	255	15	135	483	761	244	1488
Obesity	40	308	17	100	748	183	75	1006
Totals	**176**	**2845**	**106**	**1085**	**2329**	**4451**	**1612**	**8392**

Table 3.1: General information on the five datasets used for data analysis; including number of subjects (**N**), and total measurement days (**Days**). Specific information regarding motivational cues includes encouraging- (**Enc**), neutral- (**Neu**), discouraging- (**Dis**), and total motivational cues (**Tot**).

the frequency of feedback (as either once per hour or once every two hours) as well as in the details of setting the reference activity pattern.

3.4 Cold start phase

Whenever a new user starts using the activity coaching application with predictive cue timing, the system does not yet have any data that defines the user's behavior. In machine learning, this is generaly known as the *cold start* issue. In this particular case, the system needs to generate a user-specific classifier — the real-time classifier (Section 3.5) — based on the responses to previously given motivational cues. Since these are not yet available, the system has to fall back on a less tailored approach. The simple solution would be to fall back on the old method of presenting feedback once every hour until enough data has been gathered. In the Kairos system we use a more advanced approach based on data from other previous users of the activity coaching system. As explained in Section 3.3.3, whenever less than 25 motivational cues with known compliance are available for the current user, the application employs the ColdStartClassifier. This classifier was trained on the available data from the five datasets described in Section 3.3.4. In this section the details regarding the creation of this classifier are explained.

3.4.1 Data

Not all of the 8,392 motivational cues that were generated in the five studies of Section 3.3.4 are usable for training the cold start classifier. In order to arrive at the final dataset for training and evaluation, the following filters were applied to the data:

Remove neutral cues (-4,451) — it is not possible to define *compliance* for neutral messages (e.g. "You're doing fine!"), as these cues are not delivered with the intention of changing the user's behavior. Although neutral messages might have an effect on activity behavior, it is not logically clear whether or not to classify either an activity increase or decrease as *compliant*, therefore all 4,451 neutral messages are removed.

Remove time-overlapping cues (-903) — the timestamp of the motivational cues in which we are interested is the moment that the user observed the cue on the smartphone's display. In the system used for gathering this data, whenever a user did not acknowledge a cue (by pressing "OK"), the next cue-screen would be overlayed on top of the previous one. Subsequently, when the user *did* observe the cues on the smartphone's screen, he would acknowledge a series of messages in short succession. In this case, all but the last acknowledged cue are filtered out.

Remove cues with unknown compliance (-269) — in order to calculate the compliance using Equation 3.1, at least 90% of the measured activity data is required in the 30 minute intervals before and after the motivational cue. Measured data is not always available due to e.g. sensor failure, or a cue being generated at the beginning or end of the day.

After applying the above filters, a total of 2,769 motivational cues usable for classification remain. Table 3.2 shows an overview of the data per dataset, including the average compliance over all cues.

3.4.2 Features

The goal of the classifier is to predict compliance for a potential motivational cue based on some contextual data known at real-time while running the application.

Dataset	N	Enc	Dis	Tot	Compliance
CFS RCT1	32	463	410	873	66.21%
CFS RCT2	25	327	153	480	63.75%
CLBP	17	58	465	523	61.95%
COPD RCT	15	234	131	365	58.08%
Obesity	17	469	59	528	47.54%
Totals	**106**	**1551**	**1218**	**2769**	**60.35%**

Table 3.2: Overview of the motivational cue data usable for classification; including number of subjects (**N**), encouraging- (**Enc**), discouraging- (**Dis**), total motivational cues (**Tot**) and average compliance.

This means that when defining the feature-vector for the cues, we are limited to information that can be gathered by the smartphone in real-time. Because we are using 'historical' data, not gathered primarily for the purpose of this classification challenge, we are also limited to the data that was actually recorded during those previous studies. The set of potential features defined for the cold start classifier is given in Table 3.3. To simplify the definitions, we assume that time `t` is the time at which the motivational cue was generated, or a potential motivational cue is generated. So, e.g. time `t - 30` corresponds to the moment thirty minutes before the generation of the cue.

3.4.3 Feature- and classifier selection experiments

For the creation of the cold start classifier we use the WEKA platform (Witten & Frank 2000, Hall et al. 2009). This toolkit contains implementations of many different machine learning algorithms, making it convenient for quickly exploring different classifier types. The toolkit is written in Java which also allows for easy integration in an Android application. We performed a series of machine learning experiments in order to (1) determine the most suitable classifier, and (2) determine the set of features that yields the best performance. We define performance in this case as classifier accuracy, i.e. the number of correctly classified instances divided by the total number of instances in the dataset (2,769). The baseline accuracy of the classifier, equal to the average compliance over the whole dataset,

Feature	Description
Activity Related Features (3)	
averageCompliance	The average compliance of all previous cues.
distanceFromReference	Percentage of deviation from the reference line at time `t`.
approachingReference	`true` if the distance from reference at time `t` is less than the distance at time `t-30`, `false` otherwise.
User Related Features (3)	
diagnosis	The diagnosis as either `cfs`, `clbp`, `copd`, or `obesity`.
age	The age of the user in years.
gender	The gender of the user as either `male`, or `female`.
Time Related Features (5)	
dayOfWeek	One of {`mon,tue,wed,thu,fri,sat,sun`}.
weekDay	`true` for {mon,tue,wed,thu,fri}, `false` otherwise.
dayPart	`morning` (<12:00), `evening` (>18:00), or `afternoon`.
hourOfDay	Time of day rounded down to the nearest full hour.
dayOfUsage	Number of days the user has been using the system.
Message Related Features (12)	
cueType	Either `encouraging` or `discouraging`.
cueGoOutside	`true` if cue tells you to go outside, `false` otherwise.
cueIsAQuestion	`true` if cue is phrased as a question, `false` otherwise.
cueSuggestIdle	`true` if cue suggests doing nothing, `false` otherwise.
totalCuesToday	Total amount of cues this day.
encouragingCuesToday	Total amount of encouraging cues this day.
discouragingCuesToday	Total amount of discouraging cues this day.
sameCueTodayCount	Total times same cue was given this day.
sameCueOverallCount	Total times same cue was given overall for this user.
sameTypeCueTodayCount	Total times same cue-type (enc/dis) was given this day.
sameTypeCueOverallCount	Total times same cue-type (enc/dis) was given overall.
Weather Related Features (2)	
maximumTemperature	Max. outside temperature today on the user's location.
minimumTemperature	Min. outside temperature today on the user's location.

Table 3.3: Overview and definitions of the 24 features used for the cold start classifier.

is 60.35% (see Table 3.2).

We employ genetic algorithms to greatly speed up the process of finding good feature sets. Genetic algorithms are search algorithms based on concepts from (biological) evolution to effectively explore large search spaces (Goldberg 1989). We aim to find an 'optimal' set of features from a possible set of 22 (i.e. a search space of $2^{22} - 1 = 4,194,303$ possible feature sets). Since the search space is too large to fully explore, our genetic approach allows us to do a guided exploration.

In our particular implementation we use chromosomes, which are 22 character long binary strings, that represent the subsets of features. Each position in the string maps to a specific feature, where a '1' means to include the feature and a '0' to exclude it from the set. The process of feature selection is then as follows. We first generate an initial population of 100 random chromosomes. For each of these chromosomes we first calculate its fitness. For this, we filter the feature set according to the mapping in the chromosome bit-string. Then, using the reduced feature set we train a certain classifier and evaluate it using 10×10-fold cross validation. The average accuracy of these cross validation runs is the chromosome's assigned fitness value. Then, a new set of 100 chromosomes is generated by selecting pairs of 2 chromosomes (using tournament selection (Miller & Goldberg 1995) with a tournament size of 2), and using these 2 chromosomes as 'parents', to generate new 'child' chromosomes through crossover ($P = 0.7$) and mutation ($P = 0.001$). This new population is evaluated, and the process is repeated until 95% convergence is reached (i.e. 95% of the chromosomes in the population are the same). This process results in a list of chromosomes that share the highest found fitness. From this list, we select the one with the highest amount of included features (or choose one randomly if there are more than one) as the final 'optimal' feature set.

Using this workflow we evaluated 60 different classifiers (including 34 variations of the nearest-neighbor algorithm) for which implementations were available in the WEKA toolkit. Figure 3.4 shows the results for a selection of the used classifiers. We show the improvement over baseline ($\frac{1671}{2769} = 60.35\%$) as a percentage by dividing the number of additionally correctly classified instances for the classifier by the remaining number of instances (i.e. $(n - 1671)/(2769 - 1671)$). The results in black are obtained using the realistic feature set as described in Table 3.3. We repeated the experiments with an enhanced feature set, with a total of 46 features, including features that can not be used in the real-time application, such as retrospectively added weather information. This was done to show potential improvement in accuracy, and the results are plotted in the background in gray.

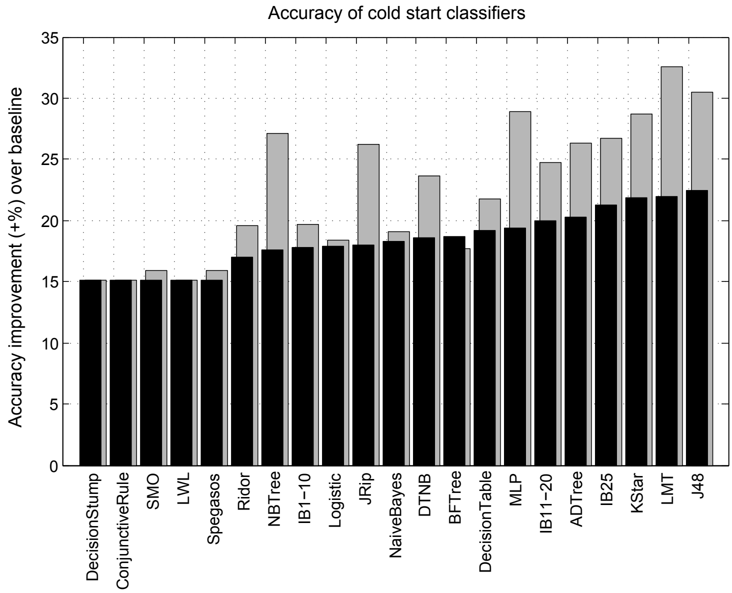


Figure 3.4: Percentage of accuracy improvement over baseline for training cold start classifiers for motivational cue prediction. In black, the results are shown when using the 'real' feature set from Table 3.3; in gray the possible results when using additional features.

Since the results of the best classifiers in Figure 3.4 do not differ significantly, we selected the simple IB25, or 25-nearest-neighbor algorithm as our final cold start classifier type. This classifier compares the feature vector of unseen cues to cues with known compliance, and averages the score of the 25 closest cues to produce the final prediction. We chose this classifier for its low complexity as well as fast performance. The classifier has a final accuracy score of 68.77%, a 21% improvement over baseline.

3.5 Adaptive phase

In the adaptive phase of the application, the feature selection and training of classifiers occurs on the smartphone using data from the individual user only. First we performed simulation experiments on the historical data. This was done to show the feasibility of predicting cue-compliance based on the data from a single user, but also to select which classifier would be most suitable to be used in the application. These experiments are described here together with details regarding the actual implementation of the system.

3.5.1 Data

From the dataset used to train the cold start classifier (Section 3.4), we now only use those subjects that have at least 25 motivational cue instances available (the adaptive phase will only start after 25 cues have been collected). Removing those subjects with less than 25 cues leaves us with a total of 50 subjects, and a total of 2,024 motivational cues, with an average compliance of 62.40%.

3.5.2 Features

The feature set used for showing the potential performance of adaptive classifiers slightly differs from the feature set used for the cold start classifier, defined in Table 3.3. The *user related features* are not usable, because they would be the same for every instance in the dataset for individual users. Since these experiments are done to show the feasibility of the approach, we do not pose restrictions on processing time, or real-time availability of features, but instead use all the data that is available in our historical datasets. With this in mind we added four weather features that were extracted from the Royal Netherlands Meteorological Institute's database: mean temperature, precipitation-duration and -sum, and a measure of the cloud-cover at the approximate location of the user (close to the place of residence). In addition, we added 13 new features extracted from the activity pattern as performed by the user in the hour before the cue, as well as 4 additional features extracted from the cue itself and the history of given cues. A total of 42 features were used to generate the data for the 50 subjects for these experiments.

3.5.3 Feature- and classifier selection experiments

The experiment setup and configuration was the same as described for the generation of the cold start classifier in Section 3.4.3. For each of the 50 subjects we performed the genetic feature selection experiments for the 60 different classifier types. For each classifier type we summed the total number of correctly classified instances over all subjects. We then calculated the improvement over baseline ($\frac{1263}{2024} = 62.40\%$) as a percentage by dividing the number of additionally correctly classified instances for the classifier by the remaining number of instances (i.e. $(n - 1263)/(2024 - 1263)$). The results of these experiments are summarized in Figure 3.5, showing for a selection of the classifiers the average improvement over baseline.

From the results in Figure 3.5 it can be seen that there is a potential for good accuracy when predicting cue compliance using the data of one subject only. The best results are obtained when using the MultiLayerPerceptron (MLP), or neural network classifier with 72% improvement over baseline (accuracy of 89.38%). However, because the best three classifiers (Ridor, Logistic and MLP) are relatively complex and have high training times, we chose to use the *SPegasos* classifier (Shalev-Shwartz et al. 2007). The SPegasos, or *Primal Estimated sub-GrAdient SOlver for SVM* classifier uses an optimized iterative algorithm for finding support vectors, and thus is an implementation of a support vector machine. The classifier shows a good trade-off between performance (87.80%, a 68% improvement over baseline) and training time, and as such is selected for implementing into the smartphone application.

3.6 Evaluation

The Kairos implementation presented here was evaluated in the context of the European AAL project IS-ACTIVE[2]. The evaluation was set up as a single-case

[2] http://www.is-active.eu/

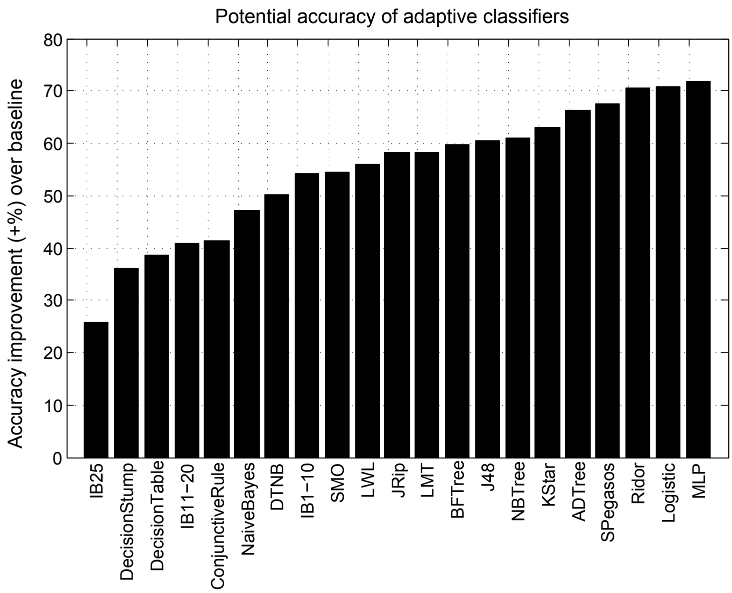


Figure 3.5: Results for feasibility of adaptive prediction of cue compliance as average percentage improvement over baseline for 50 subjects for various classifier types.

experimental study with the main purpose of investigating long-term changes in activity behavior of COPD patients (general results of the intervention are presented in (Tabak, op den Akker, Vollenbroek-Hutten & Hermens 2014)). In this article we describe the study design and evaluation results from the perspective of the work on automatic prediction of motivational cue timing.

3.6.1 Study design

A total of 10 patients (4 female) with a clinical diagnosis of COPD participated in the longitudinal study and were recruited from a rehabilitation center in the Netherlands. The study consisted of four phases: (1) baseline, (2) cold start intervention, (3) adaptive intervention, and (4) follow-up. The baseline measurement consisted of a one-week "blind" measurement of physical activity. During this period, patients wore the activity sensor and used the smartphone application for data collection only. After the baseline period, a mean daily activity level for each patient was calculated. Based on this mean, the reference line was defined by spreading the mean daily activity according to the daily distribution of healthy individuals: 40% in the morning, 30% in the afternoon and 30% in the evening. These values were then raised by 10% to provide a challenging, but achievable personal goal (Locke & Latham 2002).

The intervention phases (2 & 3) lasted for a total of three months. During the intervention, the patients where able to see a graph of their current cumulative daily physical activity on the smartphone application projected over their reference (goal) line. Also, the patients started to receive motivational cues, first provided by the cold start classifier (Section 3.4) and later by the adaptive classifier (Section 3.5). Whenever the patient's current activity is more than 10% below- or above the reference line, they are eligible for receiving encouraging and discouraging cues respectively. Additionally, during the intervention phases, the patient's reference line was updated to reflect changes in mean daily physical activity. This process was performed manually by a physiotherapist using a synchronized web portal service. The final phase of the study was a one-week follow-up measurement of daily physical activity, taking place three months after the end of the intervention.

3.6.2 Results: cue compliance

Results from a clinical point of view are described in (Tabak, op den Akker, Vollenbroek-Hutten & Hermens 2014) and show that the intervention is effective for some patients, but not for all. We focus here on the evaluation of the cold start and adaptive classifiers, as well as the patient's compliance to the motivational cues generated; and compare these results to our earlier studies. Table 3.4 shows for each patient in the trial the total number of days for which data is available ($\mathbf{D}_{tot}$), the number of intervention days ($\mathbf{D}_{cue}$), the encouraging ($\mathbf{C}_{enc}$), discouraging ($\mathbf{C}_{dis}$) and total ($\mathbf{C}_{tot}$) amount of cues generated and the mean compliance per individual as total number of complied cues divided by the total number of cues for which compliance could be calculated.

Subj.	$\mathbf{D}_{tot}$	$\mathbf{D}_{cue}$	$\mathbf{C}_{enc}$	$\mathbf{C}_{dis}$	$\mathbf{C}_{tot}$	**Compliance**	
isa07	66	52	57	137	194	116/178	65.17%
isa08	52	43	59	162	221	132/211	62.56%
isa09	32	27	21	6	27	15/25	60.00%
isa10	55	42	61	5	66	29/60	48.33%
isa11	80	67	145	160	305	183/270	67.78%
isa12	82	72	67	70	137	79/130	60.77%
isa13	88	74	251	56	307	151/280	53.93%
isa14	48	22	16	64	80	47/77	61.04%
isa15	17	11	30	51	81	36/69	52.17%
isa16	40	32	11	27	38	24/37	64.86%
Totals:	**560**	**442**	**718**	**738**	**1456**	**812/1337**	**60.73%**

Table 3.4: Basic statistics on the motivational cues provided in the evaluation. For each subject we show the total number of measured days ($\mathbf{D}_{tot}$), total number of days in the intervention period ($\mathbf{D}_{cue}$), Total encouraging ($\mathbf{C}_{enc}$), discouraging ($\mathbf{C}_{dis}$) and combined ($\mathbf{C}_{tot}$) cues and the compliance to those combined cues.

The average compliance of 60.73% over 1337 cues is not significantly better than the average compliance of 60.35% over 2769 items from all previous datasets (Table 3.2), using z-test ($p = 0.4052$). Compared to the 58.08% of the historical COPD RCT dataset, this difference of 2.65% is also not significant ($p = 0.1788$). However, the patients in our trial used the application for a relatively long period of time, receiving on average 134 usable cues per patient, compared to only 26 for the historical datasets. As compliance decreases over time, the comparison of final mean compliance is not completely fair. Figure 3.6 shows this compliance over time by plotting the average compliance for all patients after receiving n messages for all historical data, the historical COPD dataset, and for the Kairos data. The graph shows the 95% confidence intervals for the Kairos dataset and the historical COPD dataset.

The graph in Figure 3.6 is divided into four segments, describing four phases of use of the coaching system. After having received 25 cues, average compliance has peaked for the historical datasets. We hypothesize that this initial high compliance is due to the novelty effect of receiving coaching from the system. Between the period of having received 26 to 55 cues, after the novelty effect has worn off, average compliance drops especially for the COPD RCT dataset and to a lesser extent for the full Historical dataset. Then, a stabilizing phase occurs (56 to 105 cues) in which the average compliance slowly converges to a certain equilibrium. Between 41 and 55 cues, Kairos performs significantly better than the total historical dataset ($p < 0.05$). In the period between 41 to 105 cues, the Kairos classifier performs significantly better than the historical COPD dataset with an average compliance of 62.70% for the Kairos classifier versus 56.00% for the historical data ($p < 0.05$). From the point of having received 106 or more messages, average compliance remains stable.

3.6.3 Results: classifiers

We look here at the performance of the adaptive classifiers during real-time use. Table 3.5 shows for each subject the number of intervention days ($\mathbf{D}_{cue}$), the final number of cues available for classification (**Cues**), the total number of light retrains ($\mathbf{T}_{light}$) and heavy retrains ($\mathbf{T}_{heavy}$) of the classifier, and the accuracy achieved by the last trained classifier (**Accuracy**).

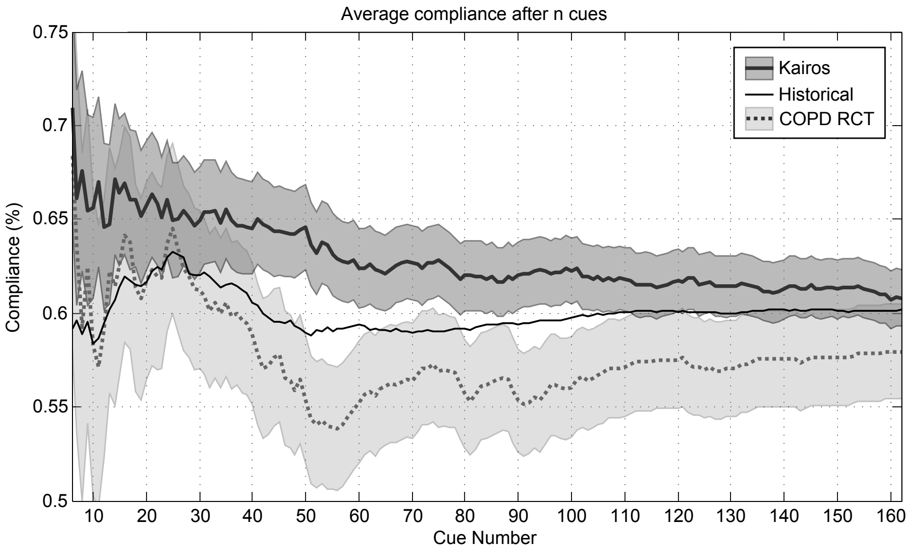


Figure 3.6: Average compliance in percentage over time after receiving *n* cues for the Kairos dataset (top), the complete historical dataset (middle) and the historical COPD RCT dataset (bottom).

As explained in Section 3.3.3, classifier training occurs at every application start-up. As the study protocol dictates the user to switch off the phone at the end of every day (for recharging), one would expect the number of training events (769) to match approximately the number of intervention days (442). Nevertheless, a large difference can be observed here. This difference of 327 training events can be caused by the user deliberately restarting the application during the day, but is much more likely to indicate regular application crashes, especially for patient isa08 and isa11.

From Table 3.5 it can be observed that final classifier performance is in general higher for subjects that have fewer useable cues. We aim to clarify this observation

Subject	$\mathbf{D}_{cue}$	Cues	$\mathbf{T}_{light}$	$\mathbf{T}_{heavy}$	Accuracy
isa07	52	178	66	6	67.24%
isa08	43	211	142	8	69.31%
isa09	27	25	34	1	96.00%
isa10	42	60	58	2	93.33%
isa11	67	270	181	10	69.63%
isa12	72	130	88	5	80.00%
isa13	74	280	75	11	68.93%
isa14	22	77	27	3	87.01%
isa15	11	69	12	2	84.06%
isa16	32	37	37	1	89.19%
Totals:	**442**	**1337**	**720**	**49**	**78.31%**

Table 3.5: Overview of received cues and number of light- ($\mathbf{T}_{light}$) and heavy-retrains ($\mathbf{T}_{heavy}$) for the classifiers from the Kairos evaluation for every patient.

in Figure 3.7, by showing the average adaptive classifier performance over time. The abscissa shows increasing numbers of cues available for the classifier, while the ordinate indicates classifier accuracy given the number of available cues. The values on the ordinate are presented as *the percentage of improvement over baseline*. Some values are interpolated to provide a continuous graph, and the shaded area shows the standard deviation of the averages over the 10 subjects. The black circles plotted in the graph show results of the same classifiers (SPegasos with Hinge-loss function) trained on the historical datasets.

The graph shows a trend that more available data for the classifier causes a drop in performance. Given few instances, the classifiers manage to achieve relatively high accuracies of 70-90% improvement over baseline. However this performance quickly drops to between 20 and 40% improvement over baseline (given more than 100 features), showing the inherent complexity of the task of predicting human behavior.

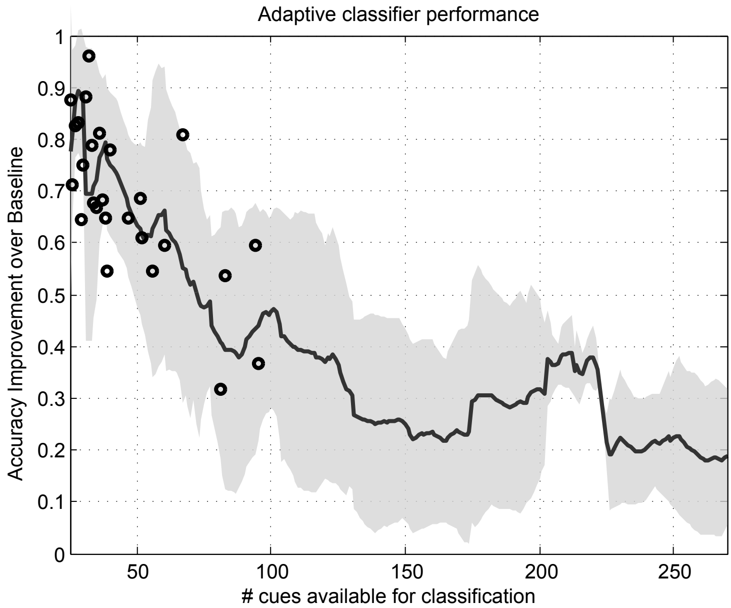


Figure 3.7: Average real-time performance (accuracy improvement over baseline) for the adaptive classifiers for increasing numbers of available cues.

Finally, in Figure 3.8 we show the training duration of the light- and heavy retrains of the classifiers in real-time. The abscissa shows the number of cues available for classification and the ordinate shows the duration of the training process in minutes. All experiments were run on HTC Desire S phones with 768MB of RAM and a 1GHz Qualcomm MSM8255 Snapdragon processor, running only the activity coaching application on the standard Android OS.

As expected, the data shows an exponential increase in training time when using more instances. For the light retrains (without feature selection) the training time is short and will remain below 10 minutes for up to 1000 instances. For the heavy retrains (with feature selection), the training time is already approaching

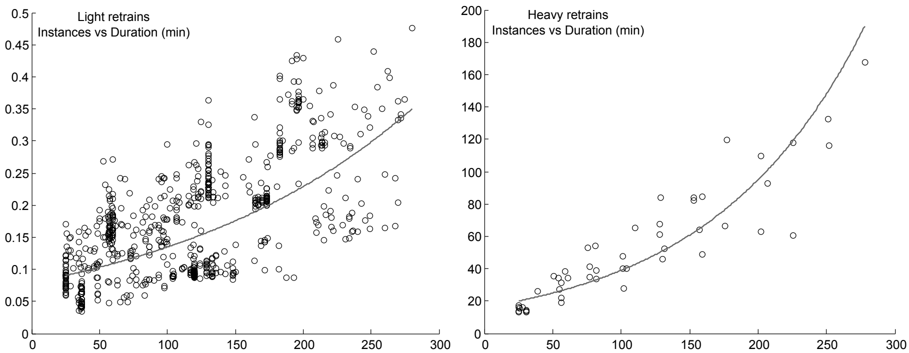


Figure 3.8: Training duration in minutes for increasing numbers of available cues for the light-retrains (left) and heavy-retrains (right).

the limits of what is desirable with a training time of 2 hours and 47 minutes for 278 instances.

3.7 Discussion

The aim of the work presented here was to increase long-term adherence to the physical activity coaching system, as well as to improve the response rate to motivational cues. To achieve this, we aimed to show the feasibility of applying a real-time self-learning system to tailor the timing of motivational cue generation to individual users. Our initial experiments in developing the cold start classifier, and showing the potential of adaptive classification, show promising results for the task of predicting compliance to individual motivational cues. Our selected support vector machine implementation for adaptive classification shows a potential accuracy of 87.80%, a 68% improvement over baseline. The cold start classifier — as can be expected — has lower accuracy (68.77%, a 21% improvement over baseline), but still has a potential to provide better timing of motivational cues than fixed timed cues. Also, we showed that the theoretical performance in the cold start phase can be further increased by using additional features, more complex models, and additional contextual information — e.g. using *logistic model* *trees* (LMT) shows a potential accuracy of 73.25% (see Figure 3.4).

From a technical point of view the solution to determine cue timing presented here is relatively complex, but the longitudinal trial performed showed that the system worked without interfering with the overall system's performance. There was no noticeable drop in battery life of the smartphone, and no technical issues that were caused by the use of the Kairos module. For two patients, the application was resetting multiple times throughout the course of the intervention, however there was no indication that these crashes were caused by the new module.

We have shown the potential benefits and the technical successful implementation of the Kairos module. The results regarding motivational cue compliance from the longitudinal trial were moderately positive. Looking at overall compliance to individual cues, the adaptive cue timing from the Kairos module does result in better compliance than the fixed timing method of providing cues in our previous study of COPD patients. The comparison is however not completely fair due to the limitation in the study of not having a 100% comparable control group. The historical datasets were recorded using a different version of the smartphone application. The principals are the same, but the new Android version has an aesthetically more appealing user interface. This may have an effect on overall user compliance that can not be measured using the current comparison, although it must be said that any possible effects should be expected to be minor.

In terms of compliance, the Kairos system consistently outperforms regular timed motivational cues. When comparing the Kairos data to results from the historical COPD dataset, compliance is significantly higher between receiving the 41^{st} and 105^{th} cue. However, compliance to individual cues is still only around 65% leaving many cues ignored by the user. Fortunately there are still several points of improvement that can be implemented for the Kairos system — whereas the fixed timing method of providing cues seems more difficult to improve. Although one can vary the amount of cues generated by changing the time intervals (e.g. from once per hour to once every 30 minutes or 2 hours), there is no logical explanation as to why this would provide more opportune feedback. The Kairos system however can still be improved in several ways. First, the cold start classification phase can be improved by training a better classifier, using more data, and additional features, as shown in Section 3.4. In the adaptive phase, there is still a difference between the potential accuracies the classifiers can achieve and those achieved at runtime of the device. As the training and feature selection phase are processor power intensive, the run-time application uses a configuration that will cause the process to terminate within a reasonable amount of time. This configuration — related to the tuning of the genetic algorithms — leads to a reduced amount of feature sets to be explored during real-time use, limiting the potential viability of the generated classifier. Using newer, more powerful smartphones can already lessen the need to simplify this feature selection phase and potentially improve the classifier's performance.

Additionally, a user's expected behavior depends greatly on the context in which the user is told to perform an action. For example, someone who is in a work meeting and is being told to go for a walk, will be unable to comply to this message. Having such contextual knowledge available on the smartphone would benefit the prediction task, and is currently no longer beyond the state of art. In (Klaassen et al. 2013), the authors show that users indeed have varying preferences for which messages they want to receive in certain situations. Applications such as Google Location History[3], or VTT's Lifeliner[4] can already automatically provide semantic information about the user's current whereabouts and are ready to be integrated into a system like Kairos.

3.8 Conclusion

The Kairos system for automatic generation of opportune motivational cues is shown to work in a longitudinal trial of 10 Chronic Obstructive Pulmonary Disease patients. No significant technical issues have been reported, and the system correctly operated through the cold start and adaptive phases. Compliance to individual cues is significantly higher than that of a previously measured COPD patient population between the 41^{st} and 105^{th} generated cue. However, these results are not completely comparable as the historical dataset was collected using a different smartphone application. Overall the Kairos system provides a clear

[3]maps.google.com/locationhistory
[4]cauit.erve.vtt.fi/lifeliner

improvement over a fixed timing method of presenting motivational cues. State of the art technology in the field of smartphone devices, context aware sensing, and reasoning algorithms enable a number of further possible improvements to the Kairos system.

Bibliography

Bickmore, T. W., Mauer, D. & Brown, T. (2009), 'Context Awareness in a Handheld Exercise Agent', *Pervasive and Mobile Computing* **5**(3), 226–235.

Bosch, S., Marin-Perianu, M., Marin-Perianu, R., Havinga, P. J. & Hermens, H. J. (2009), Keep on Moving! Activity Monitoring and Stimulation Using Wireless Sensor Networks, *in* 'Proceedings of the 4th European Conference on Smart Sensing and Context (EuroSCC 2009)', Guildford, United Kingdom, pp. 11–23.

Boudreau, F., Godin, G. & Poirier, P. (2011), 'Effectiveness of a computer-tailored print-based physical activity intervention among French Canadians with type 2 diabetes in a real-life setting', *Health Education Research* **26**(4), 573–585.

Bouten, C. V., Verboeket-Van De Venne, W. P., Westerterp, K. R., Verduin, M. & Janssen, J. D. (1996), 'Daily physical activity assessment: comparison between movement registration and doubly labeled water', *Journal of Applied Physiology* **81**(2), 1019–1026.

Brug, J., Campbell, M. K. & Van Assema, P. (1999), 'The application and impact of computer-generated personalized nutrition education: A review of the literature', *Patient Education and Counseling* **36**(2), 145–156.

Bull, F. C., Kreuter, M. W. & Scharff, D. P. (1999), 'Effects of tailored, personalized and general health messages on physical activity', *Patient Education and Counseling* **36**(2), 181–92.

Campbell, M. K., DeVellis, B. M., Strecher, V. J., Ammerman, A. S., DeVellis, R. F. & Sandler, R. S. (1994), 'Improving Dietary Behavior: The Effectiveness of Tailored Messages in Primary Care Settings', *American Journal of Public Health* **84**(5), 783–787.

Cooper, C. B. (2009), 'Airflow obstruction and exercise', *Respiratory Medicine* **103**(3), 325–334.

Dekker-van Weering, M. G. H., Vollenbroek-Hutten, M. M. R. & Hermens, H. J. (2012), 'Do Personalized Feedback Messages about Activity Patterns Stimulate Patients with Chronic Low Back Pain to Change their Activity Behavior on a Short Term Notice?', *Applied Psychophysiology and Biofeedback* **37**(2), 81–9.

Ekeland, A. G., Bowes, A. & Flottorp, S. (2010), 'Effectiveness of telemedicine: A systematic review of reviews', *International Journal of Medical Informatics* **79**(11), 736–771.

Evering, R. M. H. (2013), Ambulatory Feedback at Daily Physical Activity Patterns: A Treatment for the Chronic Fatigue Syndrome in the Home Environment?, PhD thesis, University of Twente.

Fry, J. P. & Neff, R. A. (2009), 'Periodic Prompts and Reminders in Health Promotion and Health Behavior Interventions: Systematic Review', *Journal of Medical Internet Research* **11**(2), e16.

Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I.-M., Nieman, D. C. & Swain, D. P. (2011), 'American College of Sports Medicine position stand. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise', *Medicine & Science in Sports & Exercise* **43**(7), 1334–1359.

Goldberg, D. E. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley Longman, Inc.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. & Witten, I. H. (2009), 'The WEKA Data Mining Software: An Update', *ACM SIGKDD Explorations Newsletter* **11**(1), 10–18.

Hawkins, R. P., Kreuter, M. W., Resnicow, K., Fishbein, M. & Dijkstra, A. (2008), 'Understanding tailoring in communicating about health', *Health Education Research* **23**(3), 454–466.

Kennedy, C. M., Powell, J., Payne, T. H., Ainsworth, J., Boyd, A. & Buchan, I. (2012), 'Active Assistance Technology for Health-Related Behavior Change: An Interdisciplinary Review', *Journal of Medical Internet Research* **14**(3), e80.

Klaassen, R., op den Akker, R., Lavrysen, T. & van Wissen, S. (2013), 'User preferences for multi-device context-aware feedback in a digital coaching system', *Journal on Multimodal User Interfaces* **7**(3), 247–267.

Kroeze, W., Werkman, A. & Brug, J. (2006), 'A Systematic Review of Randomized Trials on the Effectiveness of Computer-Tailored Education on Physical Activity and Dietary Behaviors', *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine* **31**(3), 205–23.

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N. & Katzmarzyk, P. T. (2012), 'Impact of Physical Inactivity on the World's Major Non-Communicable Diseases', *Lancet* **380**(9838), 219–229.

Locke, E. A. & Latham, G. P. (2002), 'Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey', *American Psychologist* **57**(9), 705–717.

Miller, B. L. & Goldberg, D. E. (1995), 'Genetic Algorithms, Tournament Selection, and the Effects of Noise', *Complex Systems* **9**(3), 193–212.

Noar, S. M., Benac, C. N. & Harris, M. S. (2007), 'Does tailoring matter? Meta-Analytic Review of Tailored Print Health Behavior Change Interventions', *Psychological Bulletin* **133**(4), 673–693.

op den Akker, H., Jones, V. M. & Hermens, H. J. (2010), Predicting Feedback Compliance in a Teletreatment Application, *in* 'Proceedings of ISABEL 2010: the 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies', Rome, Italy.

op den Akker, H., Jones, V. M. & Hermens, H. J. (2014), 'Tailoring Real-Time Physical Activity Coaching Systems: A Literature Survey and Model', *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change* .

op den Akker, H., Tabak, M., Marin-Perianu, M., Huis In't Veld, R. M., Jones, V. M., Hofs, D., Tönis, T. M., van Schooten, B. W., Vollenbroek-Hutten, M. M. & Hermens, H. J. (2012), Development and Evaluation of a Sensor-Based System for Remote Monitoring and Treatment of Chronic Diseases - the Continuous Care & Coaching Platform, *in* 'Proceedings of the Sixth International Symposium on e-Health Services and Technologies (EHST 2012)', Geneva, Switzerland, pp. 19–27.

Qian, H., Kuber, R., Sears, A. & Murphy, E. (2011), 'Maintaining and modifying pace through tactile and multimodal feedback', *Interacting with Computers* **23**(3), 214–225.

Shalev-Shwartz, S., Singer, Y. & Srebro, N. (2007), Pegasos: Primal Estimated sub-GrAdient SOlver for SVM, *in* 'Proceedings of the 24th International Conference on Machine Learning', Corvallis, Oregon, USA, pp. 807–814.

Skinner, C. & Campbell, M. (1999), 'How effective is tailored print communication?', *Annals of Behavioral Medicine* **21**(4), 290–298.

Sohl, S. J. & Moyer, A. (2007), 'Tailored interventions to promote mammography screening: A meta-analytic review', *Preventive Medicine* **45**(4), 252–261.

Spruit, M. A., Singh, S. J., Garvey, C., ZuWallack, R., Nici, L., Rochester, C., Hill, K., Holland, A. E., Lareau, S. C., Man, W. D.-C., Pitta, F., Sewell, L., Raskin, J., Bourbeau, J., Crouch, R., Franssen, F. M. E., Casaburi, R., Vercoulen, J. H., Vogiatzis, I., Gosselink, R., Clini, E. M., Effing, T. W., Maltais, F., van der Palen, J., Troosters, T., Janssen, D. J. A., Collins, E., Garcia-Aymerich, J., Brooks, D., Fahy, B. F., Puhan, M. A., Hoogendoorn, M., Garrod, R., Schols, A. M. W. J., Carlin, B., Benzo, R., Meek, P., Morgan, M., Rutten-van Mölken, M. P., Ries, A. L., Make, B., Goldstein, R. S., Dowson, C. A., Brozek, J. L., Donner, C. F. & Wouters, E. F. M. (2013), 'An official American Thoracic Society/European Respiratory Society Statement: Key Concepts and Advances in Pulmonary Rehabilitation', *American Journal of Respiratory and Critical Care Medicine* **188**(8), e13–e64.

Tabak, M., op den Akker, H. & Hermens, H. J. (2014), 'Motivational cues as real-time feedback for changing daily activity behavior of patients with COPD', *Patient Education and Counseling* **94**(3), 372–378.

Tabak, M., op den Akker, H., Vollenbroek-Hutten, M. M. & Hermens, H. J. (2014), 'Improving Long-Term Activity Behaviour of Individual Patients with COPD using an Ambulant Activity Coach', *Applied Psychophysiology and Biofeedback (submitted)* .

Tabak, M., Vollenbroek-Hutten, M. M., van der Valk, P. D., van der Palen, J. & Hermens, H. J. (2013), 'A telerehabilitation intervention for patients with Chronic Obstructive Pulmonary Disease: a randomized controlled pilot trial.', *Clinical Rehabilitation (online)* .

Witten, I. H. & Frank, E. (2000), *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*, Morgan Kaufmann Publishers.

World Health Organization (2013), 'The World Health Report 2013 - Research for Universal Health Coverage'.

3

Chapter 4

Tailored Motivational Message Generation: A Model and Practical Framework for Real-Time Physical Activity Coaching

Harm op den Akker, Miriam Cabrita, Rieks op den Akker, Valerie M. Jones and Hermie J. Hermens (Submitted).

Harm op den Akker, Laura S. Moualed, Valerie M. Jones and Hermie J. Hermens, in *Proc ISABEL2011*. DOI:10.1145/2093698.2093845

Wilko Wieringa, Harm op den Akker, Valerie M. Jones, Rieks op den Akker and Hermie J. Hermens, in *Proc AIME2011*. DOI: 10.1007/978-3-642-22218-4_7

Chapter 4

Abstract

This paper presents a well-founded, comprehensive and practical framework for automatically generating tailored effective physical activity promotion messages in a real-time setting. Basic aspects of motivational message communication are time, intention, content and presentation. Tailoring of messages to the individual user may involve all aspects of communication. A linear modular system is presented for generating tailored motivational messages. It is explained how properties of user and user context are taken into account in each of the modules of the system and how they affect the linguistic presentation of the generated messages. The model of motivational messages presented is based on an analysis of existing literature as well as the analysis of a corpus of motivational messages used in previous studies. The model extends existing 'ontology-based' approaches to message generation for real-time digital coaching systems found in the literature. Practical examples are given on how simple tailoring rules can be implemented throughout the various stages of the framework. Such examples should guide further research by clarifying the notion of what it means to use e.g. *user targeting* to tailor a message. Future work is pointed out in generalizing the present model and framework beyond the physical activity promotion domain, defining efficient ways of evaluating individual tailoring components, and improving potential effectiveness through the creation of accurate and complete user- and context models.

4.1 Introduction

Due to the ageing population, the prevalence of chronic diseases is increasing worldwide. The growing demand on healthcare services calls for cost-effective treatments that reduce the demands on healthcare professionals. Provision of eHealth and telemedicine services, in particular technology mediated services which stimulate and support patient's self-efficacy, is a fast growing field of research (Ekeland et al. 2010, Klasnja & Pratt 2012) and is widely regarded as a promising paradigm to reduce the burden on the healthcare system. An important factor in the prevention and treatment of chronic diseases and healthy ageing support is the maintenance of a healthy lifestyle in terms of regular physical activity. Lee et al. investigated the effects of physical inactivity on a number of major non-communicable diseases worldwide and estimate that physical inactivity causes 9% of premature mortality and that if completely eliminated would increase the world's life expectancy by 0.68 years (Lee et al. 2012). Deaths from the major noncommunicable diseases are projected to rise from 36 million in 2008 to 52 million in 2030 (World Health Organization 2013).

An obvious strategy to stimulate users to reach their physical activity goal is to provide reminders when the user is deviating from a predefined desired level of activity. This approach has already been shown to provide added value over merely measuring daily activity levels and giving the user the option of seeing their progress in real time, both for Chronic Low Back Pain patients (Dekker-van Weering et al. 2012) and for Chronic Obstructive Pulmonary Disease patients (Tabak et al. 2014). However, there is a need identified in these studies as well as in related research (Erriquez & Grasso 2008, Cortellese et al. 2009) to provide a more sophisticated and *tailored* delivery of messages.

In 2008, Hawkins et al. defined tailoring as "*any of a number of methods for creating communications individualized for their receivers, with the expectation that this individualization will lead to larger intended effects of these communications*" (Hawkins et al. 2008). Within the scope of motivational messages for physical activity, the intended effect, overall, is to elicit sustainable behavior change. Tailoring is generally considered to be a useful method of increasing the effectiveness of health interventions, through the delivery of tailored print information in the early days (Skinner et al. 1999, Noar et al. 2007, Boudreau et al. 2011), but more recently also in telemedicine or eHealth applications (Campbell et al. 1994, Bull et al. 1999). Tailoring in the field of physical activity promotion is an area that is still on the rise as shown by the survey presented in (op den Akker, Jones & Hermens 2014), and the question of how to tailor *messages* promoting physical activity is far from solved.

The goal of this work is twofold. The first goal is to define a Model of Motivational Messages encompassing all the concepts relevant to the topic of motivational message generation. Second, and based on the model, we aim to develop a well-founded and practical framework for automatically generating effective physical activity promotion messages in a real-time setting. We base our work on existing literature on the topic, as well as years of experience in developing and evaluating such systems in our own lab. The work presented here is not a literature review, as related literature was already systematically analyzed in e.g. (Latimer et al. 2010). We summarize the results of this review, and related research on tailoring and message generation, in order to identify concepts that play a role in the effect of messaging on physical activity (Section 4.2). Based on the analysis of a large variety of motivational messages we structure these identified concepts in a theoretical Model of Motivational Messages (Section 4.3). Based on this model we present a practical framework for automatic generation of messages, taking into account the message *timing*, *intention*, *content* and *representation* (Section 4.4). We discuss how the model and framework can improve message effectiveness, and allows for a more structured evaluation of message based coaching in Section 4.5 and finish with conclusions in Section 4.6.

4.2 Background

As input for our model of motivational messaging and the derived framework for automatically generating such messages, we look at the literature describing theory, as well as the state of art in real-time motivational messaging for physical activity. In this section we aim to identify which concepts play a role in the generation of motivational messages. The concepts that are bracketed and bold-highlighted throughout the text are identified concepts that will be summarized at the end of this section in Table 4.1.

4.2.1 Offline motivation

A study by Berry & Latimer-Cheung (Berry & Latimer-Cheung 2013) analyzed the impact of public and private health-promotion message campaigns and provides a set of recommendations for such messaging based on the literature regarding social marketing and physical activity promotion. Some of their recommendations — building a brand, work together to create campaigns — are less relevant for the current topic, but others apply to the topic of personal message generation. The first such recommendation is to create messages that grab people's attention, which the author's say is achieved through tailoring and message framing. Tailoring provides messages that are more relevant to the recipient (**tailoring**). According to the elaboration likelihood model (Petty & Cacioppo 1986, Petty & Wegener 1999) this causes the recipient to process the message more thoroughly leading to a higher likelihood of changing thoughts and behaviors. Message framing is another technique for grabbing the recipient's attention (Rothman & Salovey 1997). Messages can be gain-framed (emphasizing the benefits of activity) or loss-framed (emphasizing the consequences of inactivity), and the authors argue that gain-framed messages are the most persuasive (**message framing**). A second relevant recommendation from the article is to base messages on well-founded theory. The authors recommend message composition to be based on behavioral change theories such as the Theory of Planned Behavior (Ajzen 1991) and Social Cognitive Theory (Bandura 1986), citing evidence that messages targeting affective attitudes, self-efficacy, and self-regulation offer promising message content (**theory-based**). A review by Michie et al. shows that message-based strategies based on theory are more likely to elicit behavior change in terms of physical activity (Michie et al. 2009).

The recommendations given in (Berry & Latimer-Cheung 2013) are supported by a systematic review on the construction of physical activity messages by Latimer et al. (Latimer et al. 2010). The authors systematically review studies that use three approaches for providing physical activity messages: tailoring, gain-framing, and targeting messages to affect a change in self-efficacy (**self-efficacy**) (Bandura 1986, Rodgers et al. 2008). A total of 22 studies were analyzed on the effects of the three mentioned approaches to message construction. In 12 studies, some form of message tailoring was applied, and in seven out of 12, this resulted in significantly greater physical activity compared to a control group. However, in six of the seven studies reporting significant effects on physical activity, the control group received no messaging at all. Only in (Marcus, Bock, Pinto, Forsyth, Roberts & Traficante 1998) significant improvements were reported compared to a control group receiving non-tailored messages. It should be noted that the study reported in (Marcus, Bock, Pinto, Forsyth, Roberts & Traficante 1998), published in 1998, used regular mail as medium for assessment and delivery of the intervention. In general, all of the considered studies that included some form of tailoring used regular mail (Cardinal & Sachs 1996, Marcus, Bock, Pinto, Forsyth, Roberts & Traficante 1998, Marcus, Emmons, Simkin-Silverman, Linnan, Taylor, Bock, Roberts, Rossi & Abrams 1998, Bull et al. 1999, Naylor et al. 1999, Blissmer & McAuley 2002, Marshall et al. 2003), email website/delivery (Hager et al. 2002, Napolitano et al. 2003, Spittaels et al. 2007) or a comparison between such methods (Marcus, Lewis, Williams, Dunsiger, Jakicic, Whiteley, Albrecht, Napolitano, Bock, Tate, Sciamanna & Parisi 2007, Marcus, Napolitano, King, Lewis, Whiteley, Albrecht, Parisi, Bock, Pinto, Sciamanna, Jakicic & Papandonatos 2007) while none of them used more advanced methods such as real time messaging through mobile phones. The authors mention that "*the general pattern of findings also suggests that more frequent doses of information may enhance the effects of tailored messages*" (Latimer et al. 2010), suggesting that real-time messaging through mobile phones may further increase the positive effects on physical activity. Finally, the authors conclude that if tailoring is used, the Transtheoretical Model's stages of change (Prochaska & Velicer 1997) should be considered as target for tailoring.

Six studies in the review were evaluated on the effects of framing messages (gain-framed vs. loss-framed) on physical activity as well as on user's intentions. Three out of four studies that assessed the effects of framed messages on physical activity reported positive effects regarding gain-framed messages. Furthermore, two out of six studies reported a positive main effect of gain-framed messages on intention, and three out of six reported moderate effects, where gain-framing would have positive effects on intention under certain conditions. Overall the authors cautiously recommend the use of gain-framed messages rather than loss-framed messages, although the effects of mixed-framed messages are as of yet inconsistent. The third aspect evaluated in the review is the use of messages that target a change in self-efficacy. On this topic, only four articles were included, and the findings of these studies were considered to be mixed, showing no real evidence for a systematic effect. The authors do state, however, that constructing messages based on theory holds promises on effect outcomes and should be considered.

4.2.2 Real-time message generation

The literature described so far is targeted at an offline delivery of messages, i.e. delivered through regular mail, or telephone, or using a website that respondents were prompted to visit e.g. weekly. One of the suggestions from the review in (Latimer et al. 2010) is that real-time delivery of messages through mobile phones may improve the effects of the messaging on physical activity behavior, compared to the slower modes of delivery discussed earlier. The purpose of the framework for message generation we present in Section 4.4 is to be used in a smartphone-based activity coaching application, as used in our previous studies (Dekker-van Weering et al. 2012, Tabak et al. 2014, op den Akker, Tabak, Jones & Hermens 2014). Therefore, we look specifically at the literature targeting mobile phone based delivery of motivational messages.

Erriquez & Grasso (Erriquez & Grasso 2008) propose an ontology-based approach for generating personalized messages that are sent to the user's mobile phone throughout the day. The work is based on an earlier publication by Morandi & Serafin (Morandi & Serafin 2007) in which a personalized motivation strategy is presented to stimulate physical activity in diabetic subjects. The strategy defined is personalized based on a number of aspects. First: the stage of change of the user — assessed through a questionnaire on readiness to change health behavior and in particular through an eHealth application (**stage of change**). Second: personal preferences and habits regarding physical activity are assessed through a questionnaire as well (**personal preferences**). The third aspect is the setting of a personal exercise goal, based on a treadmill walking test, the user is placed on a specific point in a progressive curve towards a healthy physical activity level (**goal setting**) (Locke & Latham 2002). The fourth aspect of personalization is through the use of information from a personal diary. The user can keep track of four aspects of daily life satisfaction related to work, social relationships, the weather and mood (**mood**). Finally, messages are personalized on the day of the week (e.g. on friday suggesting activities for the weekend), as well as the location and weather — e.g. suggesting outside activities in case of nice weather (**context**). The starting point of the ontology in (Erriquez & Grasso 2008) is a decomposition of messages into five components: (1) a *comment* to provide feedback on performance, (2) an *argument* to say that physical activity is good for the user, (3) an *aid to introspection* relating performance with the user's diary values, (4) a *suggestion* on how to improve performance, and (5) an *encouragement* to conclude the message. Starting with these 5 ontological classes the discourse ontology is created. For each of these classes, subclasses are defined that can model the various types of tailoring that can be applied. For example, the *argument* message component has subclasses for messages focusing on *negative consequences*, or *positive consequences* (i.e. message-framing (Rothman & Salovey 1997)).

A similar ontological approach to message generation is taken in our earlier work in (Wieringa et al. 2011). Here the focus is on motivational messages that suggest an activity in order to *encourage* more physical activity, or *discourage* activity. The ontological structure is based on the type of activities the system can suggest, where e.g. encouraging advice can be subdivided into *inside* or *outside* categories, and further splits are made based on the type of activity as e.g. *exercise*, or *household*. Context-aware message generation is achieved by pruning the message tree in real time based on context data such as weather or whether or not the user has a job. The proposed system includes a mechanism for self-learning (**self-learning**). Each time a message is generated, the response to that message (i.e. compliance — or whether or not the user actually performed more activity after reading the message (op den Akker et al. 2010)) is stored in the appropriate node of the message ontology. This information is then used to guide the selection in each split of the ontology, increasing the likelihood of paths that led to successful messages in the past.

4.2.3 Summary of concepts

Based on the literature described here we have identified 10 concepts that play a role in the generation of motivational messages. These concepts, highlighted throughout the text above are summarized and shortly explained in Table 4.1. A general trend that can clearly be observed throughout the literature is the advocacy of tailoring. In (op den Akker, Jones & Hermens 2014) we surveyed the literature and constructed a model of tailoring for real-time physical activity coaching applications. We link the 10 concepts related to motivational messaging to the tailoring model and continue to use the terminology as defined in (op den Akker, Jones & Hermens 2014) in the development of the model of motivational messages in Section 4.3.

4.3 A Model of Motivational Messages

In this work we aim to provide a practical framework for generating motivational messages, encompassing all of the concepts identified in the literature. In order to develop a structured and logical framework we first define our domain model of motivational messages. The model presented here is based on a number of previous modeling attempts from the literature as well as an analysis of our own set of motivational messages used in various eHealth/telemedicine interventions.

4.3.1 High level model

In (op den Akker, Jones & Hermens 2014) we provide a working definition of tailoring, noting that we always tailor *something* to *someone*. The *someone* in our scenario is the user of a physical activity promotion tool, and the *something* is the system's communication to the user. Thus, **communication is tailored to the user**. This communication occurs in instances, and a motivational message is a prime example of such a communication instance. In (op den Akker, Jones & Hermens 2014) we formalized an idea already presented in (op den Akker et al. 2011) that every communication instance can be seen as having four distinct properties: **timing**, **intention**, **content**, and **representation**. These four components form the backbone of the model of motivational message — depicted in Figure 4.1 — and each will be explained shortly based on the following example:

> "*You haven't been active enough today. Maintaining a healthy level of physical activity can drastically reduce the chances of cardiovascular disease. In order to achieve your daily goal, you need to walk for at least another 18 minutes or perform 12 minutes of vigorous exercise.*"

Timing — The **timing** of a motivational message is the moment at which the system presents the message to the user without request (system-initiated) or

Concept	Description
tailoring	The process of creating communications individualized to the receivers (Hawkins et al. 2008).
message framing	Specifically focusing the message on attaining desirable outcomes (gain frame) or undesirable outcomes (loss frame) (Rothman & Salovey 1997).
theory-based	Actively base message composition on behavioral change theories, such as the theory of planned behavior (Ajzen 1991) and social cognitive theory (Bandura 1986).
self-efficacy	Belief in one's own capabilities towards performing a task (e.g. changing health behavior) (Bandura 1986, Rodgers et al. 2008).
stage of change	The attitude towards changing a behavior as categorized into five stages: precontemplation, contemplation, preparation, action and maintenance (Prochaska & Velicer 1997).
personal preferences	User's likes and dislikes, e.g. in terms of preferred physical activities to perform (Morandi & Serafin 2007).
goal-setting	Individuals are more likely to change the higher the specificity and (achievable) difficulty of a goal; taking into account e.g. self-efficacy and feedback (Locke & Latham 2002).
mood	Diffuse affective states not resulting from specific events but more likely associated with general views at a point in time (Biddle & Mutrie 2008).
context	Using context to provide relevant information to the user (Dey & Abowd 1999), e.g. providing suggestions based on location or weather.
self-learning	Automatically improve decision making processes by using new data available from interactions with the user (op den Akker, Jones & Hermens 2014).

Table 4.1: Overview of identified relevant concepts related to motivational message generation.

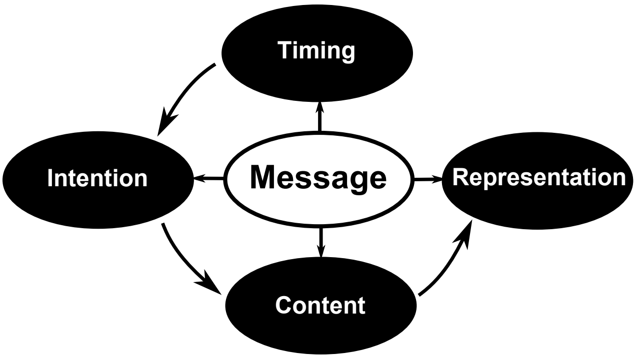


Figure 4.1: The backbone of the model of motivational messages.

the moment at which the user requests information from the system (user-initiated). The timing of motivational messages is extensively discussed, and a system for generating messages at opportune moments is presented in (op den Akker, Tabak, Jones & Hermens 2014).

Intention — In (op den Akker, Jones & Hermens 2014) we identified three different **intention**s in the message: (1) to provide information on the user's current progress (i.e. feedback), (2) to inform about the benefits of physical activity, and (3) to give a suggestion on an activity to perform. In our model of motivational messages we view these three intentions (feedback, argument, and suggestion) as secondary intentions. With that in mind, we define the example message's primary intention as simply to '*encourage*' physical activity. This may seem obvious, as for most physical activity promotion tools, encouraging activity is the only goal. However, for various chronic disease patient populations such as Chronic Obstructive Pulmonary Disease (Tabak et al. 2014), Chronic Fatigue Syndrome (Evering 2013) or Chronic Low Back Pain (Dekker-van Weering et al. 2012), a "secondary" goal is identified to balance activity levels over the day. This means that sometimes — e.g. during the mornings — a message intention can be to *discourage* activity. Finally, if the user is performing well, the message intention could be to simply state that the user should keep up the good work.

We call this a *neutral* intention.

Content — The **content** of a motivational message is the information that the message wants to convey to the user. In the example message, the content is the fact that the user has not been active enough, a factual statement about the advantages of physical activity, and the suggestion to walk or perform vigorous exercise. The content of messages is the primary topic in the related literature discussed in Section 4.2, and the framework as presented in (Erriquez & Grasso 2008) deals also exclusively with the content of messages.

Representation — The **representation** deals with the 'outer form' of the message. In our example the representation, at its highest level, is natural language text. Natural language text messages is the focus of this work, but it should be kept in mind that there are various other possible ways of presenting the intention and content of a message to the user, such as through the visualization of a flowering garden (Consolvo et al. 2008), or music (Oliver & Flores-Mangas 2006, de Oliveira & Oliver 2008).

The distinction between timing, intention, content and representation forms the backbone of the model in Figure 4.1. In the sections below we will now explain the details of each of these high level concepts.

4.3.2 Modeling Timing

The issue of message timing is an important aspect in the process of real-time motivational messaging to encourage behavior change, as presenting a message to an individual at a moment when he is able and willing to take action can be of major influence to the message effectiveness. However, within the scope of this work — targeting the *generation* of messages — we can only assume that we have to generate a message *now*. The timing of motivational messages is extensively discussed, and a system for generating messages at opportune moments is presented in (op den Akker, Tabak, Jones & Hermens 2014). For the sake of completeness of the message model, we include timing, and the high level distinction between *user initiated* messaging (when the user specifically requests a message), and automatically generated *system initiated* messaging (when the system decides to initiate a communication).

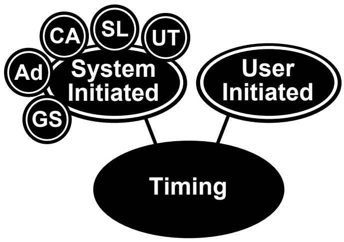


Figure 4.2: The timing branch of the motivational message model showing two high level types of message timing and the four most relevant types of tailoring that can be applied to system initiated timing: goal setting (**GS**), adaptation (**Ad**), context awareness (**CA**), self learning (**SL**), and user targeting (**UT**).

The timing branch of the model in Figure 4.2 shows the most relevant tailoring aspects associated to the selection of message timing. Goal setting (**GS**), context awareness (**CA**) and self learning (**SL**) are used in the Kairos system for automatic prediction of message timing presented in (op den Akker, Tabak, Jones & Hermens 2014). Adaptation (**Ad**) could also be used, as it is likely that user's in e.g. different stages of change or with different levels of self-efficacy need fewer or more frequent reminders regarding their physical activity behavior. User targeting (**UT**) can relate to e.g. using the information in a personal digital agenda to decide not to disturb the user during an important business meeting.

4.3.3 Modeling Intention

As explained in the definition of *intention* above, we distinguish between primary and secondary message intention. We define the primary intention of a *motivational message* to be a choice between *encouraging* physical activity (in case the user is not active enough), *discouraging* physical activity (in case the user is too active), or to provide a *neutral* comment (in case the user is doing fine). This high level separation forms the first split in the **intention** branch of our message ontology in Figure 4.3. For the secondary message intention, we have identified four possible intentions that align with the model proposed by Erriquez & Grasso (Erriquez & Grasso 2008): **reinforcement** (similar to 'encouragement' in

(Erriquez & Grasso 2008)), to provide *feedback* (called 'comment' in (Erriquez & Grasso 2008)), to give an *argument* and/or to provide a *suggestion*.

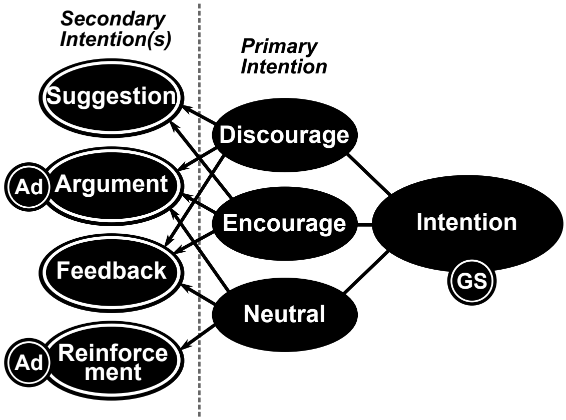


Figure 4.3: The intention branch of the motivational message model showing the choice between primary intentions and the selection of secondary intention(s) as well as the two most relevant types of tailoring that can be in this selection process: adaptation (**Ad**), and goal setting (**GS**).

Three of the four secondary intentions are present in the example sentence provided at the beginning of this section. The *feedback* intention relates to the purpose of giving the user an insight in his current performance. An *argument* intention can be used to educate the user about the benefits of physical activity (gain-framed) or the negative consequences of inactivity (loss-framed) and can attempt to convince the user of changing his view on physical activity. The *suggestion* intention can be applied to give the user a suggestion on how to achieve his desired level of physical activity. The fourth secondary intention, **reinforcement**, is an intention that is only applicable when the user is performing well, and is meant to reinforce the user's current behavior. It should be clear that the primary intentions defined here are mutually exclusive, as the intention of a message cannot be to e.g. *encourage* and *discourage* physical activity at the same time. On the other hand, a motivational message can contain any combination of at least one of the secondary intentions, forming a composite statement as the example given above.

The tailoring aspects that play a role in the definition of message intention are indicated in Figure 4.3 by the small circles: goal setting (**GS**) and adaptation (**Ad**). Without a target activity level, or goal, it is impossible to define the primary intention of a message as either neutral, encouraging or discouraging, as you need knowledge about the user's performance in relationship to some given goal. On the level of secondary intention the most relevant form of tailoring is *adaptation* (i.e. tailoring to key theoretical determinants (Hawkins et al. 2008, op den Akker, Jones & Hermens 2014)). This is because the decision of presenting the user with an argument could depend on e.g. the user's stage of change — as providing an argument for users in the preparation stage of change seems to be a promising tailoring strategy. Similarly, the decision to provide a **reinforcement** when the user is performing well may be most appropriate (but not exclusively) for users with low self-efficacy.

4.3.4 Modeling Content

The content of the message is the topic most prominently discussed in literature, and can also be considered the most complex one. Over the last years we have been working on a large number of physical activity motivation studies in which motivational messaging was included. Results have been published from studies targeting various chronic diseases such as Chronic Obstructive Pulmonary Disease (Tabak et al. 2014), Chronic Fatigue Syndrome (Evering 2013) or Chronic Low Back Pain (Dekker-van Weering et al. 2012). The message sets used in these studies were all comparable (with minor variations), and consisted of a total of 24 message (4 neutral, 11 encouraging and 9 discouraging), e.g.:

> "*Your activity level is sufficient. Keep it up!*" (neu)
> "*You have taken more rest. Take a walk around the block.*" (enc)
> "*You have become more active. Sit down for a while.*" (dis)

From these 24 messages, exemplified above, the composite nature of the message already becomes apparent. The first content-part of the message is always a statement regarding the user's current activity performance — better known as **feedback**. The **follow-up** to this differs based on the intention of the message. For neutral messages, the second part is always a short **reinforcement** message, indicating that the user is doing well and should continue to do so. For encouraging, as well as discouraging messages, the second part always contains a practical **suggestion** to either perform an active (e.g. walking) or a passive (e.g. reading) activity.

A big motivator for the work presented here is the fact that research has identified the need to better match the motivational messages to individual users (Noar et al. 2007). The same conclusion was drawn from our own research on the topic, and in various ongoing research the set of messages to be used for activity motivation is receiving increasing attention. From ongoing research we have analyzed two additional message sets, one targeting cancer recovery patients, and one targeting office workers. The message set for cancer recovery patients is created in the context of the Dutch FitterNaKanker (fit after cancer) project and consists of 237 different messages, split into three groups targeting people who primarily need to *balance* activity, people who need to primarily *increase* activity, and people who even need to *decrease* activity. The message set for office workers consists of 72 messages, and is created in the context of the Dutch COMMIT Swell project. Each message is categorized to target one of eight *persona's*, based on three characteristics: (1) the activity balance (as proper or improper), (2) intention to change (as *contemplation / preparation / action* or *precontemplation / maintenance* stages of change) and (3) self-efficacy (as low, or average-high) (Achterkamp et al. 2013). There is some overlap between the messages in these two sets. Before looking into the content of the messages it is interesting to note that both datasets use the notion of *persona's* for matching messages to certain user-types. This is considered a form of *user targeting* (**UT**) (op den Akker, Jones & Hermens 2014) and needs to be supported by our model and framework.

From the analysis of the messages we have identified one additional high-level message component: the **argument**. This type of message, exemplified below is common in the new message sets, and aims to provide a reason as to why the user should increase (or decrease) his physical activity. The example below consists of **feedback**, followed by an **argument**.

> "*You are somewhat less active. An increase in physical activity can make you feel more energetic.*"

The second compositional split for the message content branch is a subdivision of the **suggestion** component identified earlier. Consider the two examples messages given below. Both messages start with an **argument** component, highlighting some positive aspect about physical activity. However, the second part of the message differs fundamentally. In both cases a suggestion is presented to the user, but we classify the first example ("*are you using your bike for small errands?*") as a **lifestyle** suggestion, whereas the second example is considered an **activity** suggestion. The difference may seem small, but there is a significant difference in the way these two message components can be tailored to the user.

> "*The effect of small activity bouts is often underestimated. Are you using your bike for small errands?*"

> "*Physical exercise can lead to mental relaxation. Time to go for a bike ride?*"

In Figure 4.4 the content branch of the motivational message model is shown, again indicating the various forms of tailoring that apply to each individual component in the small circles. This branch of the message model shows four of the same high level components as identified in (Erriquez & Grasso 2008) (where we use the terminology **feedback** instead of 'comment', and **reinforcement** instead of 'encouragement'). The tailoring concepts that play the most important role in tailoring content are: goal setting (**GS**), user targeting (**UT**), adaptation (**Ad**) and context awareness (**CA**). Feedback content can be tailored through goal setting to include information about the goal — e.g. "*You have done 7,000 steps*", compared to "*You have done 7,000 out of 10,000 steps*". User targeting can be used to provide tailored feedback by using user characteristics such as age, gender or weight to present the information in a tailored way, e.g. calories burned. The argument component, providing a reason for the user to perform physical activities, should match the current beliefs about physical activity of the user (adaptation), but should also be relevant for the user's medical condition (user targeting). By using information such as self-efficacy level or stage of change, an argument can be

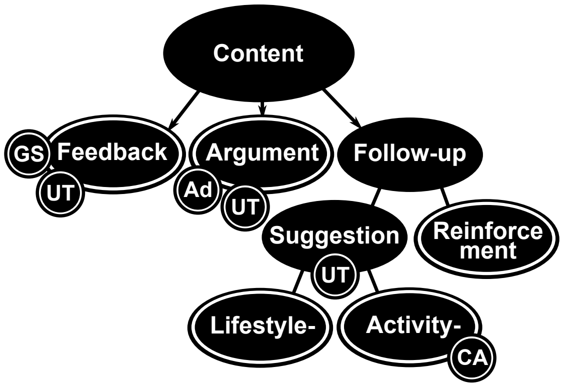


Figure 4.4: The content branch of the motivational message model showing the message content ontology and the four most relevant types of tailoring that can be applied to the various components: goal setting (**GS**), user targeting (**UT**), adaptation (**Ad**) and context awareness (**CA**).

selected that is most relevant to the individual user. User targeting also includes aspects such as personal preferences regarding daily life activities, as such, lifestyle- and activity suggested can both be tailored to suggest activities that match these preferences (such as biking, outdoor activities, etc.). As activity suggestions are intended to be followed immediately, context awareness plays a very important role. The system should, for example, not recommend any outdoor activities when it's raining, or suggest household chores when the user is at work.

4.3.5 Modeling Representation

The final branch of the motivational message model deals with the outer representation of the message. The representation defines how the message is communicated to the user, which can be done in a variety of different ways. The examples given earlier are the flower garden of the UbiFit Garden application (Consolvo et al. 2008), in which **feedback** is given to the user by showing different types and sizes of flower on the screen, enhanced with goal setting by presenting a large butterfly if a daily goal is met. Another example is the music based feedback of the MPTrain (Oliver & Flores-Mangas 2006) and TripleBeat (de Oliveira & Oliver 2008) systems that automatically select music of an appropriate beat/rhythm to stimulate exercising (jogging) at the proper pace. Another different modality is used in the Haptic Personal Trainer (Qian et al. 2010, Qian et al. 2011) where vibration signals are given to the user to promote a faster or slower walking pace. The model of motivational messages would not be complete without the concept of representation and the notion of different modalities for presenting information to the user. However, previous efforts in human computer interaction research to model the different available modalities exist (Bernsen 1993, Bernsen 1997, Obrenovic & Starcevic 2004, Jaimes & Sebe 2007), and a contribution to this field far exceeds the scope of this work.

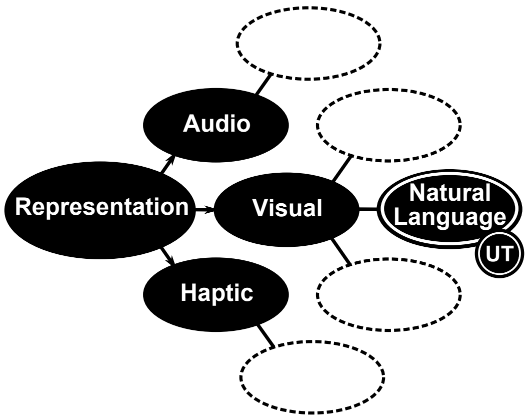


Figure 4.5: The representation branch of the motivational message model showing a small excerpt of an ontology of modalities. Due to the scope of this work, natural language text, and user targeting (**UT**) as a way of tailoring this, is the only positioned modality.

Figure 4.5 shows a simplified model of message modalities, indicating that many gaps exist. In this work, the focus is on the generation of natural language text messages, which offers a modality that allows us to present the nuances of content and intention as defined in the previous sections. The tailoring concept most relevant to the representation of motivational messages through natural language text is user targeting (**UT**). On the one hand, the 'identification' tactic (Hawkins et al. 2008) of identifying the recipient by name, is a typical example of tailoring message representation. On the other hand, tailoring the message tone (formal vs informal), style (suggestive vs imperative) or length (long vs short) to the user's personal preference, can be considered a more advanced form of user targeting.

4.3.6 The complete model

Finally, we present the complete model of motivational messages in Figure 4.6. There is an implied ordering in the model of the concepts of timing, intention, content and representation, indicated by the arrows between these concepts. This ordering represents a temporal relationship that implies that message components should logically be generated in sequence by the framework (presented in Section 4.4). The reason for this particular temporal ordering is as follows.

Reasoning backwards, the *representation* deals with the outer form of the message, and depends on a detailed description of the message *content* to be defined. The message *content* is built up from the several sub-components (feedback, argument, etc...), each of which match with a predefined *intention*. The message's primary intention of discouraging activity, encouraging activity, or providing neutral comment depends on the user's current level of physical activity and is therefore dependent on time. Another reason that the *timing* component starts the chain, is that in case of *user initiated* timing, the intentions of the message should match those requested by the user (e.g. if the user presses a button to request feedback, the generation of the message intention should take this request into account).

The temporal relationship between the four main components of the model is present strictly for practical reasons. The goal of providing this model is to provide a structured framework in which all of the relevant forms of tailoring can be placed. It should help to clarify what it means to tailor a message to, for example, the user's stage of change, and which part of the message generation process can be influenced by this. The following section will provide technical details regarding

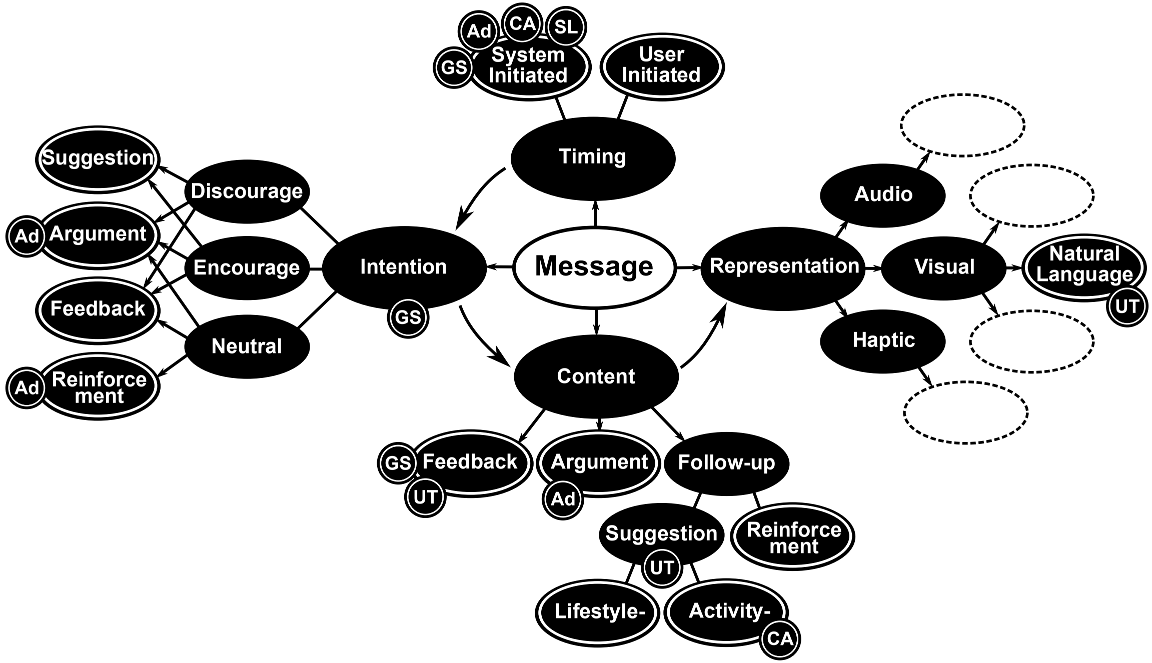


Figure 4.6: The complete model of motivational messages encompassing the four major components of timing, intention, content and representation.

the implementation of this model into a framework for message generation. It will provide examples of how several tailoring methods fit into this process.

4.4 A Practical Framework for Generating Motivational Messages

The value of the model of motivational messages comes from its practical application. The decomposition of a message into its four main components, as well as the temporal ordering between these components forms a logical and solid basis for the Message Generation Framework. The goal is to use this framework in our various studies on real-time physical activity interventions. The mobile platform developed in our lab runs on Google's Android operating system and is written in Java. Regarding the development of the Message Generation Framework we will show some relevant details of the Java implementation. By showing the one-to-one mapping of application blocks to the model defined in Section 4.3 we aim to demonstrate the model's practical value. After describing the overall architecture of the framework, we will provide examples of how the application of various tailoring methods can be addressed by implementing simple decision or selection rules in the process of generating the four message components.

First, an object representation of the model is implemented. Solid lines in the model (Figure 4.6) are implemented as abstractions, while arrows are implemented as property relationships between objects. The first two tiers of the object model are shown as a UML class diagram in Figure 4.7 representing the information as defined in the model, with some additional logging parameters in case of the *Timing* object.

The actual generation of the messages is handled in a matching organization of generator classes. The *MessageGenerator* class is the entry point to the framework, and calls in sequence the generators for the specific subcomponents. The *TimingGenerator* is triggered with an external request and generates a *Timing* object (see Section 4.4.1). Given as input the Timing object, the *IntentionGenerator* generates the message *Intention* object (Section 4.4.2). The generated Intention is passed on to the *ContentGenerator*, that generates the message *Content* (Section

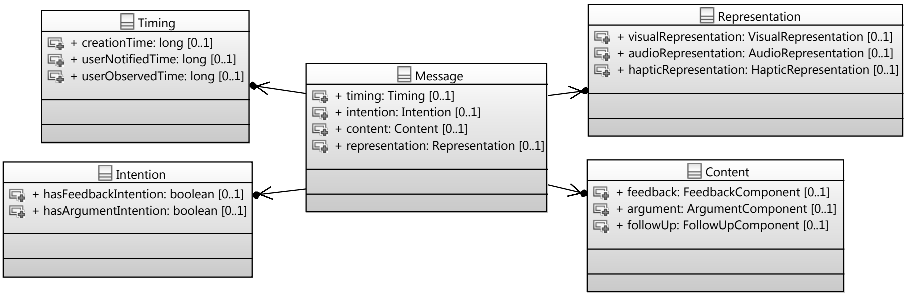


Figure 4.7: The first two tiers of the Java object model implementation of the model of motivational messages.

4.4.3). In the final step, the Content object is passed to the *RepresentationGenerator*, which selects a 'tangible' *Representation* of the message that matches the given content (Section 4.4.4). This functional part of the framework is represented as a UML sequence diagram in Figure 4.8.

4.4.1 Generating timing

During execution the *MessageGenerator* class is called with a request to generate a motivational message. This implies that the actual *timing* of the message is handled outside of the framework. An implementation of a system that automatically determines an optimal timing based on the tailoring methods as indicated in the model (Figure 4.6) is already presented in (op den Akker, Tabak, Jones & Hermens 2014) and could later be embedded into the current framework. For now, the *MessageGenerator* is called with the request to generate a *system initiated* message as determined by a system such as the Kairos system (op den Akker, Tabak, Jones & Hermens 2014), or to generate a *user initiated* message, if the user pressed some button on his device.

In the case of system initiated timing, the *TimingGenerator* simply instantiates the *Timing* object as being of type *SystemInitiatedTiming* and completes. The case of user initiated timing is slightly more complex. A user initiated message

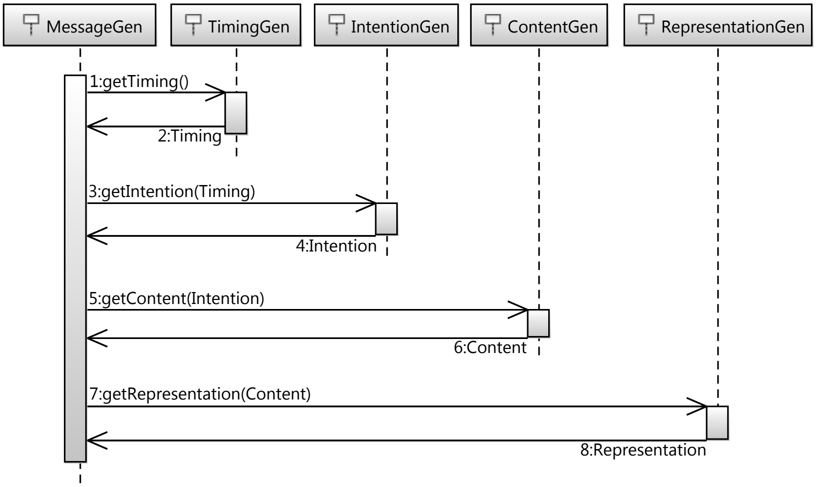


Figure 4.8: High level functional overview of the incremental, sequential process of generating the message components.

is triggered by the user through some interaction with the application's Graphical User Interface (GUI). This GUI interaction raises certain expectations towards its results, as a user who presses e.g. a '*suggestion*' button will expect to receive a suggestion message. These expectations are related to the message's intentions, which will be generated in the next step of the process. The *UserInitiatedTiming* object includes a number of parameters for storing these "requested intentions" (see Figure 4.9). The *IntentionGenerator* (Section 4.4.2) should then honor these requests while generating the actual *Intention* object.

4.4.2 Generating intention

The second step in the process of generating a motivational message deals with the message *Intention*. Defining the message intention is a two step process of choosing first the primary intention (discourage, encourage or neutral), and then choosing one or more applicable secondary intentions.

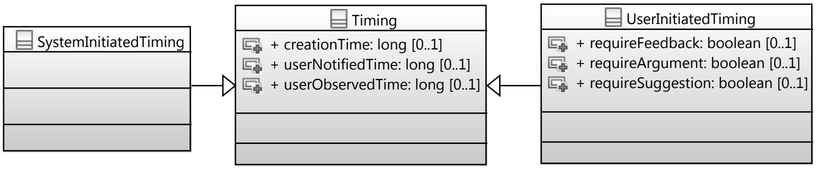


Figure 4.9: The Timing part of the message object model.

The primary intention of whether to *encourage* activity, *discourage* activity, or provide a *neutral* comment to reinforce the current behavior is based on the user's level of physical activity compared to his goal at the *time* of generating the message. It is important to notice that this time dependency is the reason why the *timing* selection occurs before the *intention* selection in the model and framework. The selection of the primary intention is described in Algorithm 2 and is based on the idea that the user is allowed a certain deviation from his current goal — `deviationMax` usually set to 10% of the current goal — before receiving encouraging or discouraging messages.

```
if Math.abs(currentActivity − currentGoal) <= deviationMax then
    intention ← new NeutralIntention();
else
    if currentActivity < currentGoal then
        intention ← new EncourageIntention();
    else
        intention ← new DiscourageIntention();
    end
end
```

Algorithm 2: Selection of primary intention.

For the secondary intentions, we can see in the model (Figure 4.6) that *adaptation* plays a role in selecting the *Reinforcement* component for neutral intentions, and the *Argument* component for the other intentions. Adaptation refers to the process of tailoring towards "*key theoretical determinants ... of the behavior of interest*" (Hawkins et al. 2008, op den Akker, Jones & Hermens 2014). In this particular case the hypothesis is that the decision to provide the user with an argument (e.g. "Being sufficiently physically activity will reduce the chance of cardiovascular disease.") should be dependent on the user's stage of change (Prochaska & Velicer 1997), as someone in e.g. the maintenance stage is likely to be already aware of such arguments, and may, if anything, only need infrequent reminding. Therefore, the probability of selecting an argument intention should be dependent on the *stage of change* user parameter. A similar dependency can be hypothesized in the selection of a reinforcement intention, based on the user's level of self-efficacy, as users with low self-efficacy would benefit more from positive reinforcement. The selection of the reinforcement parameter based on self-efficacy, as well as the time since receiving a previous reinforcement message is shown in Algorithm 3.

```
selfEfficacyFactor ← 1.0 − UserModel.getSelfEfficacy();
elapsedTime ← InteractionModel.getTimeSinceLastReinforcement();
if elapsedTime > 120 then
 |  timeFactor ← 1.0;
else
 |  timeFactor ← elapsedTime / 120;
end
probability ← selfEfficacyFactor × timeFactor;
if probability > Math.random() then
 |  intention.setReinforcementIntention(true);
end
```

Algorithm 3: Selection algorithm for the reinforcement intention dependent on the user's self efficacy level and the time elapsed since the last reinforcement was sent to the user.

Algorithm 3 shows an example of how the framework's decompositioning facilitates turning the complex conceptual problem of "*how to use adaptation to* *tailor a motivational message*" into a simple probabilistic decision rule. The example rule implemented here contains some explicit and some implicit assumptions that leave room for fine-tuning or open the possibility of auto-tuning using a self-learning system. These assumptions are, for example, the linear scaling of the `selfEfficacyFactor` or the linear scaling of the `timeFactor` over a 2 hour period (the time values represent minute values). Another point to note is the dependency on a `UserModel` to retrieve the self-efficacy level and a `InteractionModel` to retrieve information on when the user last received a previous reinforcement message. Dependencies on such models, including `ContextModels` (containing contextual information such as weather data) and `DomainModels` (containing information relevant to the current domain, e.g. physical activity levels) are present all throughout the framework and are necessary to form a clear separation between decision rules and information gathering/processing components. Finally, the intention branch of the object model is show in Figure 4.10.

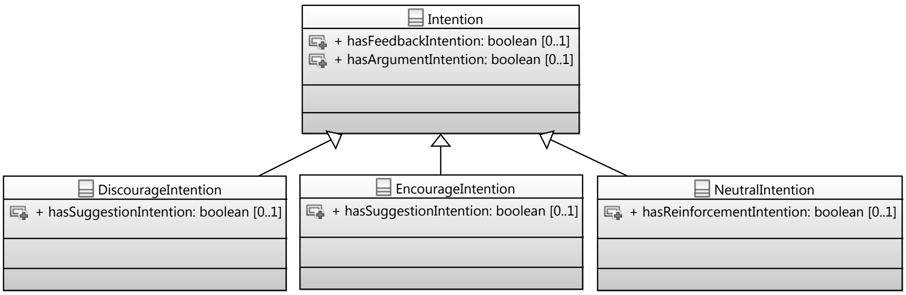


Figure 4.10: The Intention part of the message object model.

4.4.3 Generating content

In the content generation step of the process, detailed decisions should be made on what the system is going to say to the user to convey the given intentions. Similar to the intention generation step, this is a multi-step process where for each provided intention, the corresponding content component can be generated in sequence. Most of the available literature on the topic of message generation focuses on this part of the process. Providing a technical definition of a complete implementation of this framework is not the goal of the current work. Therefore we will illustrate the task of the content generator with an example and show how the work described elsewhere in the literature fits within the framework (Klaassen, op den Akker & op den Akker 2013, Wieringa et al. 2011).

Feedback

As an example we describe the generation of the *Feedback* content component. Looking at the model in Figure 4.6 there are two possible ways of tailoring: goal setting (**GS**) and user targeting (**UT**). The goal setting method of tailoring feedback content is a rather simple decision on whether or not to include information regarding the goal into the feedback component. Compare the following feedback component representations: "*You have done 7,385 steps*" with "*You have done 7,385 of the 10,000 daily steps*". It seems obvious that including the goal information may be a useful addition to the feedback, but it is possible that it is not necessary to be reminded of this all the time. We can therefore implement a probability rule based only on the time since last receiving goal-information in a previous motivational message, similar to Algorithm 3 (lines 2 – 7).

Providing user targeted feedback can be achieved by tailoring the unit (i.e. steps) to some user specific parameters obtained from the `UserModel`. For example, as many commercial fitness applications do, the progress can be expressed as the number of calories burned, based on the user's gender, age and weight parameters. This information could be further tailored if the user's favorite food is known from the `UserModel`. In this case the unit for progression can be e.g. *Pizzas*: "You have burned 3.5 pizzas so far, 1.7 more to reach your goal!" (see e.g. (Klaassen, op den Akker & op den Akker 2013)).

Activity suggestions

The suggestion component was also identified in (Erriquez & Grasso 2008), and the authors state that "*such suggestion has to be, again, personalized on the basis of habits, preferences and beliefs of the patient*". In our earlier work in (Wieringa et al. 2011) we describe an ontology of activity suggestions based on the type of activity considered. Activities are positioned into the ontology based on distinguishing between e.g. 'inside' and 'outside' activities, and activity types such as 'household', 'exercise', or 'recreational'. The selection of a relevant activity occurs by traversing the ontology from top to bottom. Context awareness (**CA**) is enforced by first pruning branches of the tree based on information from the `ContextModel`. For example, outside messages are disregarded if the `ContextModel.isWeatherGood()` functions returns `false`. User targeting (**UT**) is achieved in a similar fashion by pruning branches of the tree based on user parameters. For example, activities involving a bike will be pruned if `UserModel.userHasBike()` returns `false`.

In the second step of the selection algorithm, the decision to traverse a certain branch in the tree is based on stored compliance to previous messages. Every time a user responds well to a suggestion such as (`Outside` → `Household` → `Gardening`) this information is stored upwards along that branch of the tree, increasing the likelihood of future selection. The idea being that users who apparently like gardening, also like household activities, and also like outside activities. This mechanic of automatically updating the user's inferred preferences to certain activity suggestions effectively enables the self-learning (**SL**) of activity suggestion selection.

Object model

The content generator allows room for many more decision rules and selection algorithm implementations in the generation of each of its various components. A final example, featured prominently in the literature (Section 4.2), would be the decision to either gain-frame or loss-frame the *Argument* component. The final result of the content generator is a complete *Content* object (see Figure 4.11 that still only needs a 'tangible' representation before being communicated to the user.

4.4.4 Generating representation

The final step in the process is to generate an outer form, or representation of the motivational message. As is shown in the model (Figure 4.6), there exist many possible modalities for presenting the message to the user. Bernsen (Bernsen 1994) claims to provide a complete taxonomy of uni-modal output modalities and presents 48 possible combinations of modalities. It is clear that not all of

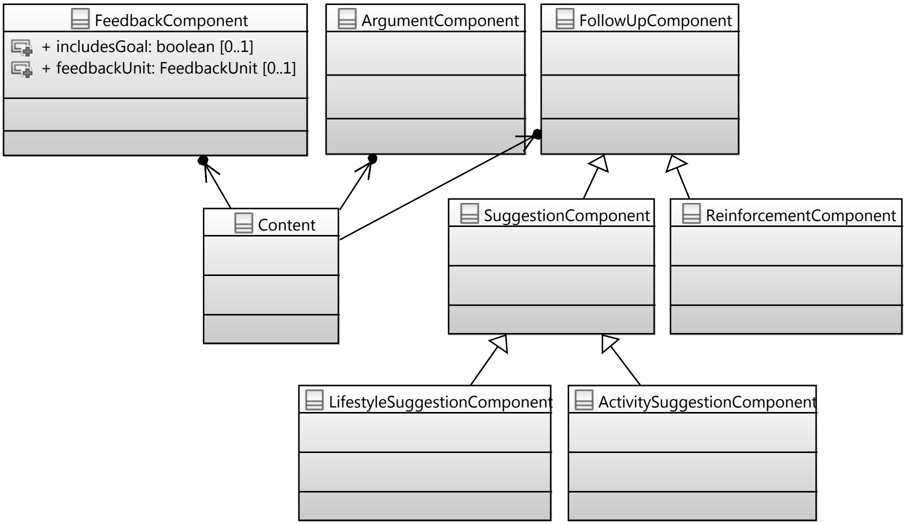


Figure 4.11: The Content part of the message object model.

these combinations will be suitable to present the type of information that can be generated by our framework of motivational messages. However, it should be stressed that natural language text is just one of many possibilities. We will however use the generation of natural language text as an example modality of representation to complete the definition of the framework.

By no means do we intend to advance the field of natural language generation, so as an example, we limit ourselves to a basic method of matching predefined utterances to the parameters provided by the *Message* and *Content* objects. To support various forms of tailoring and limit the number of required messages in our database, we suggest the use of slot-filling. As an example, we take the following instance of a *Message* object as input to the *PresentationGenerator*:

```
<Message>
  <Intention primary="Encourage">
    ...
  </Intention>
  ...
  <Content>
    <FeedbackComponent
      includesGoal="true"
      feedbackUnit="steps"
      activityLevel="7365"
      activityToGoal="2635"
      goalLevel="10000"/>
  </Content>
  ...
</Message>
```

The generator's goal would be to gather from a given message database, all the messages that are of type `Feedback`, have the `includesGoal` parameter set to `true` and include '`steps`' in the `feedbackUnit` parameter. If more than one message is available from the database, a selection can be made randomly. An example selection could be the following messages:

> So far you've done [level] [unit] out of the [goal].
> Progress: [level]/[goal] [unit].
> You need [togo] [unit] to reach your goal.

Selecting the last of the three messages examples and applying the slot-filling would result in the final representation of the feedback component: "*You need 2635 steps to reach your goal.*". The most straightforward implementation of the representation generator component would apply the same method to the matching of natural language text strings to the previously generated parameters of the motivational message. Additional user targeting (**UT**) can in this step be applied by using slot-filling to address the recipient of the message by name: "*Dear John, your progress is 7,365/10,000 steps.*". The effect of such name mentioning depends on the individual user. In case of smoking cessation name mentioning effects were shown to be dependent on the person's individual stance on health in general (health value) (Dijkstra 2014). As mentioned in the definition of the representation branch of the model (Section 4.3.5) another possible form of user targeting that can be applied in this step is the selection of text-messages based on tone and style. For example, some users might prefer a longer and formal messages style, e.g.: "*You have not yet achieved your planned goal of 10,000 steps for today. Another 2,635 steps are needed to complete your goal.*", while others might prefer something shorter and less formal, e.g.: "*Another 2,635 steps to go man!*". Determining which of the styles best fit the current user could be be achieved by applying self-learning algorithms.

4.5 Discussion

There is a considerable body of literature regarding motivational message generation. The literature provides a solid foundation for the work presented in this paper, as during the analysis we identified a large number of theoretical concepts that play a role in the effectiveness of such messages (see Table 4.1). What is lacking in the current state of art are practical tools that can help guide the design and development of real-time coaching applications. By structuring the identified concepts, and relating it to a well founded model of tailoring (op den Akker, Jones & Hermens 2014), we have provided a model of motivational messages. This model places all relevant concepts into its proper context and can show at a glance where each tailoring method can be used. The model itself can in practice help to guide research by clarifying the notion of what it means to use e.g. *user targeting* to tailor a message.

Perhaps more importantly, the defined structural decomposition and temporal relationships between the four major components of timing, intention, content and representation, allows for a sequential process of message generation. This method of divide and conquer significantly reduces the complexity of this task. Furthermore, each major component of the model is further deconstructed to a level on which reasoning about tailoring in real-time context of use becomes relatively simple and intuitive. This simplicity and intuitive nature of "decision rules" is demonstrated in several examples throughout the definition of the Message

Framework in Section 4.4. We present this framework as a one-to-one technical implementation of the model, to show the practical nature of the model, and provide an example of how it can be used to generate tailored motivational messages in a real-time application.

We aim to use the framework in future studies in physical activity promotion, and we hope that the modular architecture will facilitate the process of experimenting with different tailoring algorithms. The question of how to efficiently evaluate the effects of these different forms of tailoring remains unanswered. As the ultimate goal of physical activity interventions is to promote sustainable behavior change, such studies are in principle longitudinal in nature. Longitudinal studies allow for a technical and functional evaluation of self-learning functionalities which we believe holds perhaps the ultimate promise of tailoring. However, such studies are also expensive to execute in terms of time and resources, and are less suitable to quickly study the effect or fine-tune the parameters of simple tailoring rules.

We have aimed to provide a complete model and framework of motivational messages, but as messaging between computers and humans is at the core of the field of Human Computer Interaction, there will undoubtedly be aspects that are not taken into consideration. When talking about a message it seems obvious to take into account the message's *sender* and *receiver*. The message sender has a certain role and a relationship with the receiver that is likely to influence the timing, intention, content and representation of the message. The sender of the messages can be seen as the system, but there is always some "real" person or authority (the author) behind it. If a physical activity promotion application is provided by a healthcare institute, this *author* is likely to have a high credibility (authority) when presenting e.g. an argument message to a patient. On the other hand, if this application is purchased from a commercial party (e.g. FitBit), the credibility would be lower, and the decision to provide such messages should be taken accordingly.

The absence of the sender/receiver in the model can in our case be defended by the fact that we assume a certain sender and receiver role for our specific applications. As our research lab is part of a rehabilitation center, we consider the *author* of the messages to have a credible medical role and the *receiver* to be a patient suffering from some chronic disease. However — and this can be considered future work — we believe that our model, perhaps if augmented with the sender/receiver component, could translate well to other behavior change domains such as smoking cessation, dietary behavior, or even to encourage energy saving behavior for consumers.

There are however more obvious directions for future work, as the framework presents many gaps for implementing tailoring functionalities. Each "end-node" in the model where a tailoring method is attached, presents its own research question and direction for future research. Another area of future work can be identified from the example implementations of tailoring methods in Section 4.4. The selection of a *reinforcement intention* for example, shows the dependency of tailoring on a reliable, easy to use, and correct User Model. From a technical point of view the process of tailoring can be seen as defining a function that takes as input information from the user, defined in a User Model, Context Model, Interaction Model, and/or Domain Model, and provides an output that is in some way presented to the user. Besides finding an "optimal" function, it is clear that the correctness and detail of information in the given models are perhaps even more important in reaching an "optimal" output.

4.6 Conclusion

In this work we presented a model of motivational messages (Figure 4.6) based on the analysis of literature and a large collection of motivational messages used in physical activity promotion interventions in our lab. We presented the decomposition of a motivational message into four high level components: timing, intention, content and representation, each of which are further deconstructed and explained in detail. The result is a theoretical model of messages that takes into account the relevant concepts identified in literature, and provides a link to the seven key tailoring concepts identified in previous work (op den Akker, Jones & Hermens 2014). To highlight the practical use of the model we present a practical framework, that demonstrates the sequential process of generating each of the four message components. Examples are given on how simple and intuitive tailoring rules can be implemented along the way. Future challenges lie in efficiently evaluating the effect of the tailoring on short-term as well as long-term intervention outcomes.

Bibliography

Achterkamp, R., Cabrita, M., op den Akker, H., Hermens, H. J. & Vollenbroek-Hutten, M. M. (2013), Promoting a Healthy Lifestyle: Towards an Improved Personalized Feedback Approach, *in* 'Proceedings of the 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013)', Lisbon, Portugal, pp. 677–679.

Ajzen, I. (1991), 'The theory of planned behavior', *Organizational Behavior and Human Decision Processes* **50**(2), 179–211.

Bandura, A. (1986), *Social foundations of thought and action: A social cognitive theory*, Prentice Hall.

Bernsen, N. O. (1993), Modality Theory in support of multimodal interface design, *in* 'ERCIM Workshop on Multimodal Human-Computer Interaction', Nancy, France, pp. 37–44.

Bernsen, N. O. (1994), 'Foundations of multimodal representations: a taxonomy of representational modalities', *Interacting with Computers* **6**(4), 347–371.

Bernsen, N. O. (1997), 'Defining a taxonomy of output modalities from an HCI perspective', *Computer Standards & Interfaces* **18**(6-7), 537–553.

Berry, T. R. & Latimer-Cheung, A. E. (2013), 'Overcoming Challenges to Build Strong Physical Activity Promotion Messages', *American Journal of Lifestyle Medicine* **7**(6), 371–378.

Biddle, S. J. H. & Mutrie, N. (2008), *Psychology of Physical Activity - Determinants, well-being and interventions*, 2nd edn, Routledge.

Blissmer, B. & McAuley, E. (2002), 'Testing the Requirements of Stages of Physical Activity Among Adults: The Comparative Effectiveness of Stage-Matched, Mismatched, Standard Care, and Control Interventions', *Annals of Behavioral Medicine* **24**(3), 181–189.

Boudreau, F., Godin, G. & Poirier, P. (2011), 'Effectiveness of a computer-tailored print-based physical activity intervention among French Canadians with type 2 diabetes in a real-life setting', *Health Education Research* **26**(4), 573–585.

Bull, F. C., Kreuter, M. W. & Scharff, D. P. (1999), 'Effects of tailored, personalized and general health messages on physical activity', *Patient Education and Counseling* **36**(2), 181–92.

Campbell, M. K., DeVellis, B. M., Strecher, V. J., Ammerman, A. S., DeVellis, R. F. & Sandler, R. S. (1994), 'Improving Dietary Behavior: The Effectiveness of Tailored Messages in Primary Care Settings', *American Journal of Public Health* **84**(5), 783–787.

Cardinal, B. J. & Sachs, M. L. (1996), 'Effects of mail-mediated, stage-matched exercise behavior change strategies on female adults' leisure-time exercise behavior', *The Journal of Sports Medicine and Physical Fitness* **36**(2), 100–107.

Consolvo, S., Mcdonald, D. W., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., Klasnja, P., Lamarca, A., Legrand, L., Libby, R., Smith, I. & Landay, J. A. (2008), Activity Sensing in the Wild: A Field Trial of UbiFit Garden, *in* 'Proceedings of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems', Vol. 08, Firenze, Italy, pp. 1797–1806.

Cortellese, F., Nalin, M., Morandi, A., Sanna, A. & Grasso, F. (2009), Personality Diagnosis for Personalized eHealth Services, *in* 'Proceedings of the 2nd International ICST Conference on Electronic Healthcare for the 21st Century', pp. 173–180.

de Oliveira, R. & Oliver, N. (2008), TripleBeat: Enhancing Exercise Performance with Persuasion, *in* 'Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services', Amsterdam, the Netherlands, pp. 255–264.

Dekker-van Weering, M. G. H., Vollenbroek-Hutten, M. M. R. & Hermens, H. J. (2012), 'Do Personalized Feedback Messages about Activity Patterns Stimulate Patients with Chronic Low Back Pain to Change their Activity Behavior on a Short Term Notice?', *Applied Psychophysiology and Biofeedback* **37**(2), 81–9.

Dey, A. K. & Abowd, G. D. (1999), Towards a Better Understanding of Context and Context-Awareness, *in* 'Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing', Vol. 4, Karlsruhe, Germany.

Dijkstra, A. (2014), 'The Persuasive Effects of Personalization through Name Mentioning in a Smoking Cessation Message', *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change (online)* .

Ekeland, A. G., Bowes, A. & Flottorp, S. (2010), 'Effectiveness of telemedicine: A systematic review of reviews', *International Journal of Medical Informatics* **79**(11), 736–771.

Erriquez, E. & Grasso, F. (2008), Generation of Personalised Advisory Messages: An Ontology Based Approach, *in* 'Proceedings of the 21st IEEE International International Symposium on Computer-Based Medical Systems (CMBS)', Jyvaskyla, Finland, pp. 437–442.

Evering, R. M. H. (2013), Ambulatory Feedback at Daily Physical Activity Patterns: A Treatment for the Chronic Fatigue Syndrome in the Home Environment?, PhD thesis, University of Twente.

Hager, R. L., Hardy, A., Aldana, S. G. & George, J. D. (2002), 'Evaluation of an Internet, Stage-Based Physical Activity Intervention', *American Journal of Health Education* **33**(6), 329–337.

Hawkins, R. P., Kreuter, M. W., Resnicow, K., Fishbein, M. & Dijkstra, A. (2008), 'Understanding tailoring in communicating about health', *Health Education Research* **23**(3), 454–466.

Jaimes, A. & Sebe, N. (2007), 'Multimodal human-computer interaction: A survey', *Computer Vision and Image Understanding* **108**(1-2), 116–134.

Klaassen, R., op den Akker, R. & op den Akker, H. (2013), Feedback Presentation for Mobile Personalised Digital Physical Activity Coaching Platforms, *in* 'Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA 2013)', Island of Rhodes, Greece.

Klasnja, P. & Pratt, W. (2012), 'Healthcare in the pocket: Mapping the space of mobile-phone health interventions', *Journal of Biomedical Informatics* **45**(1), 184–198.

Latimer, A. E., Brawley, L. R. & Bassett, R. L. (2010), 'A systematic review of three approaches for constructing physical activity messages: What messages work and what improvements are needed?', *International Journal of Behavioral Nutrition and Physical Activity* **7**(36), 17.

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N. & Katzmarzyk, P. T. (2012), 'Impact of Physical Inactivity on the World's Major Non-Communicable Diseases', *Lancet* **380**(9838), 219–229.

Locke, E. A. & Latham, G. P. (2002), 'Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey', *American Psychologist* **57**(9), 705–717.

Marcus, B. H., Bock, B. C., Pinto, B. M., Forsyth, L. H., Roberts, M. B. & Traficante, R. M. (1998), 'Efficacy of an individualized, motivationally-tailored physical activity intervention', *Annals of Behavioral Medicine* **20**(3), 174–180.

Marcus, B. H., Emmons, K. M., Simkin-Silverman, L. R., Linnan, L. A., Taylor, E. R., Bock, B. C., Roberts, M. B., Rossi, J. S. & Abrams, D. B. (1998), 'Evaluation of Motivationally Tailored vs. Standard Self-help Physical Activity Interventions at the Workplace', *American Journal of Health Promotion* **12**(4), 246–253.

Marcus, B. H., Lewis, B. A., Williams, D. M., Dunsiger, S., Jakicic, J. M., Whiteley, J. A., Albrecht, A. E., Napolitano, M. A., Bock, B. C., Tate, D. F., Sciamanna, C. N. & Parisi, A. F. (2007), 'A Comparison of Internet and Print-Based Physical Activity Interventions', *Archives of Internal Medicine* **167**(9), 944–949.

Marcus, B. H., Napolitano, M. A., King, A. C., Lewis, B. A., Whiteley, J. A., Albrecht, A. E., Parisi, A., Bock, B., Pinto, B., Sciamanna, C., Jakicic, J. & Papandonatos, G. D. (2007), 'Telephone Versus Print Delivery of an Individualized Motivationally Tailored Physical Activity Intervention: Project STRIDE', *Health Psychology* **26**(4), 401–409.

Marshall, A. L., Bauman, A. E., Owen, N., Booth, M. L., Crawford, D. & Marcus, B. H. (2003), 'Population-Based Randomized Controlled Trial of a Stage-Targeted Physical Activity Intervention', *Annals of Behavioral Medicine* **25**(3), 194–202.

Michie, S., Abraham, C., Whittington, C., McAteer, J. & Gupta, S. (2009), 'Effective techniques in healthy eating and physical activity interventions: A meta-regression', *Health Psychology* **28**(6), 690–701.

Morandi, A. & Serafin, R. (2007), A personalized motivation strategy for physical activity promotion in diabetic subjects, *in* 'Proceedings of the 2nd Workshop on Personalisation for E-Health', Corfu, Greece, pp. 51–55.

Napolitano, M. A., Fotheringham, M., Tate, D. F., Sciamanna, C. N., Leslie, E., Owen, N., Baumann, A. & Marcus, B. H. (2003), 'Evaluation of an Internet-Based Physical Activity Intervention: A Preliminary Investigation', *Annals of Behavioral Medicine* **25**(2), 92–99.

Naylor, P. J., Simmonds, G., Riddoch, C., Velleman, G. & Turton, P. (1999), 'Comparison of stage-matched and unmatched interventions to promote exercise behaviour in the primary care setting', *Health Education Research* **14**(5), 653–666.

Noar, S. M., Benac, C. N. & Harris, M. S. (2007), 'Does tailoring matter? Meta-Analytic Review of Tailored Print Health Behavior Change Interventions', *Psychological Bulletin* **133**(4), 673–693.

Obrenovic, Z. & Starcevic, D. (2004), 'Modeling Multimodal Human-Computer Interaction', *Computer* **37**(9), 65–72.

Oliver, N. & Flores-Mangas, F. (2006), MPTrain: A Mobile, Music and Physiology-Based Personal Trainer, *in* 'Proceedings of the 8th Conference on Human Computer Interaction with Mobile Devices and Services', Espoo, Finland, pp. 21–28.

op den Akker, H., Jones, V. M. & Hermens, H. J. (2010), Predicting Feedback Compliance in a Teletreatment Application, *in* 'Proceedings of ISABEL 2010: the 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies', Rome, Italy.

op den Akker, H., Jones, V. M. & Hermens, H. J. (2014), 'Tailoring real-time physical activity coaching systems: a literature survey and model', *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change (online)* .

op den Akker, H., Moualed, L. S., Jones, V. M. & Hermens, H. J. (2011), A Self-Learning Personalized Feedback Agent for Motivating Physical Activity, *in* 'Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)', Barcelona, Spain.

op den Akker, H., Tabak, M., Jones, V. M. & Hermens, H. J. (2014), 'Reaching Kairos: Adaptive Prediction of the Opportune Moment for Stimulating Physical Activity', *... (submitted)* .

Petty, R. E. & Cacioppo, J. T. (1986), The Elaboration Likelihood Model of Persuasion, *in* 'Communication and Persuasion - Central and Peripheral Routes to Attitude Change', Springer New York, pp. 1–24.

Petty, R. E. & Wegener, D. T. (1999), The Elaboration Likelihood Model: Current Status and Controversies, *in* S. Chaiken & Y. Trope, eds, 'Dual-process theories in social psychology', The Guilford Press, chapter Chapter 3, pp. 37–72.

Prochaska, J. O. & Velicer, W. (1997), 'The Transtheoretical Model of Health Behavior Change', *American Journal of Health Promotion* **12**(1), 38–48.

Qian, H., Kuber, R. & Sears, A. (2010), Maintaining Levels of Activity using a Haptic Personal Training Application, *in* 'CHI '10 Extended Abstracts on Human Factors in Computing Systems', Atlanta, Georgia, USA, pp. 3217–3222.

Qian, H., Kuber, R., Sears, A. & Murphy, E. (2011), 'Maintaining and modifying pace through tactile and multimodal feedback', *Interacting with Computers* **23**(3), 214–225.

Rodgers, W. M., Wilson, P. M., Hall, C. R., Fraser, S. N. & Murray, T. C. (2008), 'Evidence for a Multidimensional Self-Efficacy for Exercise Scale', *Research Quarterly for Exercise and Sport* **79**(2), 222–234.

Rothman, A. J. & Salovey, P. (1997), 'Shaping Perceptions to Motivate Healthy Behavior: The Role of Message Framing', *Psychological Bulletin* **121**(1), 3–19.

Skinner, C. S., Campbell, M. K., Rimer, B. K., Curry, S. & Prochaska, J. O. (1999), 'How effective is tailored print communication?', *Annals of Behavioral Medicine* **21**(4), 290–298.

Spittaels, H., De Bourdeaudhuij, I. & Vandelanotte, C. (2007), 'Evaluation of a website-delivered computer-tailored intervention for increasing physical activity in the general population', *Preventive Medicine* **44**(3), 209–217.

Tabak, M., op den Akker, H. & Hermens, H. J. (2014), 'Motivational Cues as Real-Time Feedback for Changing Daily Activity Behavior of Patients with COPD', *Patient Education and Counseling* **94**(3), 372–378.

Wieringa, W., op den Akker, H., Jones, V. M., op den Akker, R. & Hermens, H. J. (2011), Ontology-Based Generation of Dynamic Feedback on Physical Activity, *in* 'Proceedings of the 13th Conference on Artificial Intelligence in Medicine (AIME)', Bled, Slovenia, pp. 55–59.

World Health Organization (2013), 'The World Health Report 2013 - Research for Universal Health Coverage'.

Chapter 5

Exploring Representations in Physical Activity Coaching Applications

Randy Klaassen, Jordi Hendrix, Dennis Reidsma, Rieks op den Akker, Betsy van Dijk and Harm op den Akker, in *International Journal on Advances in Telecommunications*, **6** (1&2), 45-56.

Randy Klaassen, Rieks op den Akker and Harm op den Akker, in *Proc PETRA2013*.

Harm op den Akker, Monique Tabak, Mihai Marin-Perianu, Rianne Huis in't Veld, Valerie M. Jones, Dennis Hofs, Thijs M. Tönis, Boris W. van Schooten, Miriam M.R. Vollenbroek-Hutten and Hermie J.Hermens, in *Proc EHST2012*.

Chapter 5

Abstract

This chapter describes an exploration of two different representations for a physical activity coaching application. We describe the development and evaluation of two prototype systems: a multi-device system prototype and an embodied agent prototype. Both systems are developed as extensions of an existing telemedicine system: the Continuous Care & Coaching Platform. The focus of the coaching representation has in the past been on simplicity, due to the common target populations of elderly users. The two prototypes developed here are targeted at inactive office workers, and as such the requirements regarding low-technological skills are relaxed, allowing more complex representations. The multi-device system prototype supports three different devices: a personal desktop PC, a smartphone and a shared screen at the office coffee corner. The system can provide feedback and engage with the user through questionnaires at the most appropriate device given the user's location. The embodied agent prototype employs a virtual embodied agent, running on the smartphone, to provide motivational messages to the user using speech and facial animations and aims to provide a representation that matches the *virtual coach* concept. Both systems have been evaluated in small-scale pilot trials. Results indicate that the multi-device concept is understood by the users, and that the system works technically. The embodied agent prototype is in a further stage of development and underwent a more thorough usability evaluation, showing that users preferred receiving coaching in plain-text messages over the embodied agent representation.

5.1 Introduction

The promotion of daily physical activity in individuals through technology-mediated services is a large focus point of current eHealth/telemedicine research. At Roessingh Research and Development, a large amount of effort has been put into developing interventions for various chronic disease patient populations. These efforts are documented in a number of publications and are summarized in three PhD dissertations on Chronic Low Back Pain (CLBP) (van Weering 2011), Chronic Fatigue Syndrome (CFS) (Evering 2013), and Chronic Obstructive Pulmonary Disease (COPD) (Tabak 2014). But, achieving and maintaining a healthy level of physical activity is not only important for patients suffering from a chronic disease, but in fact for everyone. The American College of Sports Medicine recommends in their 2011 position stand that the majority of adults should perform moderate-intensity cardio respiratory exercise training for at least thirty minutes each day (Garber et al. 2011).

Office workers form a large group of people at risk of not adhering to these recommendations, due to the sedentary nature of their day-to-day work. As such, office workers are a popular target group for physical activity interventions in research (Chan et al. 2004, Levine & Miller 2007, Blangsted et al. 2008) and, one can argue, most likely a prominent target audience for commercial physical activity tracking tools such as FitBit[1] or Withings[2]. This chapter will focus on office workers and their wish to combat a sedentary lifestyle. In accordance with the user-centered design philosophy (Van Velsen et al. 2013) we will clarify the context of the work in the form of the following scenario.

5.1.1 Scenario of a sedentary office worker

Bob is 35 years old, slightly overweight, and has been working as a programmer for the past ten years. His company values a healthy workforce, and it is decided to purchase a system that helps their office workers in achieving and maintaining a healthy lifestyle, inside the office, but also in their private lives. The system consists of a sensor for tracking daily activity, and three devices that are used to

[1] http://www.fitbit.com/
[2] http://www.withings.com/

provide feedback to the user: the user's smartphone, the desktop computer and a public display at the coffee corner.

Bob arrives at work, sits down, and spends the first two hours of the day writing a document. At 11:15 AM, a message with a graph pops up on his desktop screen. The graph shows Bob's activity so far, with a projection on his daily goal. The message displayed is an advice for Bob: "*I see that you have meetings in the afternoon. Now would be a good time to go for a lunch walk!*". This seems like a good idea to Bob and he decides to go for a walk. After his lunch, Bob gets a cup of coffee at the coffee corner. He chats a bit with his colleagues and looks at the screen suspended above the coffee machine. The screen shows, for every person currently standing at the coffee corner, the number of steps taken that day, with one of Bob's colleagues at the top of the chart.

A full workday has passed and Bob drives back home in his car. When he is at home, he prepares his meal and sits on the couch to have dinner and watch some television. Suddenly his phone vibrates. The activity coaching system presents Bob with a message urging him to do a ten minute walk around the block in order to reach his daily quota. Bob thinks he can squeeze that in before the start of his favorite television program and takes the walk. After the walk and the television program Bob goes to bed, knowing he has done enough exercise for that day. The next day Bob looks at his activity history online via his web browser. He is proud of himself, knowing that he is managing to reach his daily goal more often.

5.1.2 Representation in physical activity coaching

As posed by op den Akker, Jones & Hermens (2014), communication in a physical activity coaching application can be seen as having four distinct properties: timing, intention, content and representation. In op den Akker, Cabrita, op den Akker, Jones & Hermens (2014) these properties are explained in further detail, identifying also which specific methods can be used to tailor each of these properties. A model of motivational messages is presented, and a practical framework for generating such messages is given. The focus of this framework is on showing how decision support rules can be implemented to tailor the message generation process to the individual user, with a focus on the *intention* and *content* steps (while *timing* is discussed in detail in op den Akker, Tabak, Jones & Hermens (2014)). In this work we focus on the last remaining aspect of communication in a physical activity coach: *representation*.

In previous physical activity intervention studies (e.g. (van Weering 2011, Evering 2013, Tabak 2014)), the target populations consisted mostly of elderly users with limited technological skills. Therefore, simplicity was a prominent user requirement in the development of the activity coaching platform; and the decision was made to develop a smartphone application with a minimalistic user interface. The activity coach used in the interventions consists of a wearable inertial sensor for activity tracking and a smartphone with a simple interface to provide the coaching, shown in Figure 5.1.

Figure 5.1: Screenshots of the activity coaching application. On the left: cumulative physical activity plotted over a reference activity line indicating the user's goal between 08:00 and 22:00. On the right: example of an encouraging motivational cue.

5

Given the target population of office workers — personified by Bob in the scenario above — we can now relax the assumptions regarding low-technological skills, in order to present a more complex example of an activity coaching application. With that in mind we have developed and tested two prototype health promotion applications. The first application is a multi-device solution, leveraging the fact that our office worker interacts with several different interaction devices throughout his day. The second application is a demonstration of using an embodied agent to represent a virtual physical activity coach.

5.1.3 The multi-device prototype

The main motivation for extending the smartphone-based coaching application into a multi-device (or cross media) application relates to the richness of feedback modalities available on different devices. In (op den Akker, Jones & Hermens 2014) we ranked the different types of feedback devices (sensor, smartphone, web portal and PC application) in terms of timeliness and richness (repeated here in Figure 5.2). This shows that the smartphone is a good choice in being able to provide *real-time* feedback combined with reasonable capabilities of interface presentation. However, the limited screen size of a mobile phone limits the potential of providing detailed insight into the user's physical activity behavior. Web portals and applications that run on a PC can be used to provide much richer information, at the cost of timeliness. One way of leveraging the richness of web portals and PC applications while still maintaining the ability to provide real-time coaching is to support these devices simultaneously in a multi-device setup. Such a multi-device application poses a number of new technical, and interaction related challenges. In Section 5.3 we will discuss these challenges, the design, implementation, and pilot evaluations of the multi-device prototype.

5.1.4 The embodied agent prototype

The second application focuses on the use of an Embodied Agent, or avatar, to represent the activity coaching application as a virtual coach. The Human Media Interaction group at the University of Twente have spent considerable effort in developing an embodied agent framework that runs on the Android platform for use in mobile coaching applications (Klaassen et al. 2012, Klaassen, op den Akker & op den Akker 2013, Klaassen, Hendrix, Reidsma, op den Akker, van Dijk &

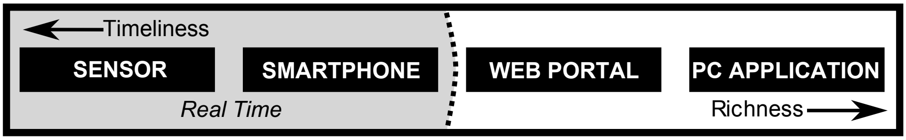


Figure 5.2: Timeliness versus richness of feedback modalities. Modalities to the left are increasingly readily available, while modalities to the right are increasingly rich in their capabilities of providing feedback. Sensor and smartphone are real time feedback modalities (op den Akker, Jones & Hermens 2014).

op den Akker 2013). People are used to receiving motivational support from a human coach. Because the application is designed to *coach* users in changing their physical activity behavior, it makes intuitive sense to also present the application as a *virtual coach*, and thus adopt a human-like representation. However, as with many applications of embodied (conversational) agents in the field, it is unclear whether this form of representation will actually have a positive effect on — in this case — health- or usability outcomes. In Section 5.4 we will present the case of using an embodied agent as a virtual physical activity coach for office workers, shortly summarizing the implementation efforts and focusing on an evaluation comparing text-based feedback with feedback presented by the virtual agent.

5.1.5 Outline

The rest of this chapter is outlined as follows. Both application prototypes are built as extensions of the telemedicine platform developed at Roessingh Research and Development. In Section 5.2 we provide relevant background information on this platform, called the Continuous Care & Coaching Platform. In Section 5.3 we describe the multi-device system prototype, and in Section 5.4 we discuss the embodied virtual coach, describing the developed prototype, while focusing on the performed evaluations. Section 5.5 contains a discussion, challenges for future work and conclusions.

5.2 Background: The Continuous Care & Coaching Platform

Over the past years, Roessingh Research and Development has been working on telemedicine services that aim to support elderly and patients with chronic diseases to achieve and maintain an active, healthy lifestyle; either independently, or supervised remotely by a healthcare professional. The **C**ontinuous **C**are & **C**oaching **P**latf**o**rm, or *C3PO*, focuses on continuous monitoring and providing feedback to the user. The platform consists of sensors, a smartphone application, a server and a web portal. Development was guided by our experience with previous telemedicine applications, and driven by a need for several improvements.

Earlier versions of the platform were mainly designed for research purposes, but the experimental studies performed, indicated that the system had the potential of growing towards a commercial behavioral intervention tool. However, a redesign was necessary to better fit the user requirements regarding usability, as well as a better integration into daily work practice of the healthcare professionals. The early involvement of patients and professionals in the requirements analysis and design process is crucial for this. Therefore, we applied an iterative, user-centered design approach, as outlined in Figure 5.3.

In this approach the designers take the users as the starting point for design and involve them in the evaluation of design choices. How to effectively involve users in the development process is not obvious because of the knowledge gap between the users and developers and their differences in language use, hampering effective communication (Huis In't Veld et al. 2010). Scenario's, as our office worker scenario described earlier (Section 5.1.1), are considered an effective technique to bridge this gap (Carroll 2000). For the requirements of the C3PO platform, scenarios were developed following the user-centered design perspective of the People-Activity-Context-Technology (PACT) framework (Benyon et al. 2004). The scenarios were transformed into functional requirements using the FICS framework (Function and events, Interactions and usability issues, Content and structure, Style and aesthetics (Benyon & Macaulay 2002)), to provide a more descriptive specification and allowing technicians to derive technical requirements (op den Akker et al. 2012).

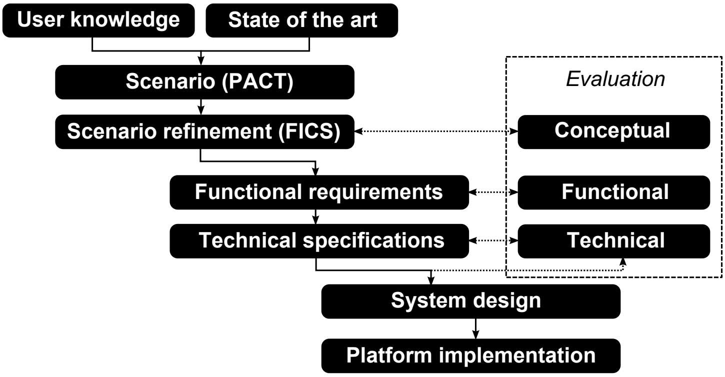


Figure 5.3: The iterative design approach of the Continuous Care & Coaching Platform (op den Akker et al. 2012).

From applying this requirement elicitation approach, it became clear that the activity monitoring system must be highly flexible and configurable to accommodate the various requirements coming from different research projects and research goals. Each research project has specific experiment protocols and procedures, including the level of detail of data shown to the users, and the involvement of healthcare professionals. The input required is also dependent on the research with some projects requiring multiple types of physiological data (e.g. heart rate, SpO_2) and others requiring input from various questionnaires. To fulfill the various needs we aimed to offer support for easy integration of different sensor types and the possibility to use multiple sensors simultaneously. Figure 5.4 gives a high level view of the architecture of the C3PO platform.

The most relevant components for the current work are the *sensor* and *smartphone*, as both prototype systems described here are built using these components as a starting point. The sensor used is a ProMove-3D wireless activity tracker, developed by Inertia Technology[3] and was designed to provide a trade-off between performance, computational and storage resources, wireless capabilities, low-power

[3] http://www.inertia-technology.com/

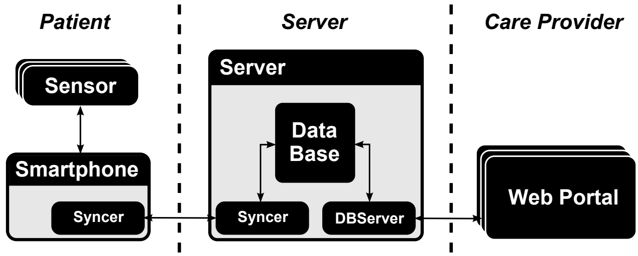


Figure 5.4: High level component overview of the Continuous Care & Coaching Platform (op den Akker et al. 2012).

operation and wearable form factor (Bosch et al. 2009). The sensor node contains a 3D-accelerometer for daily physical activity tracking (as well as a 3D gyroscope and 3D magnetic compass for more advanced body tracking), and communicates wirelessly over Bluetooth to the smartphone. An algorithm to convert the raw accelerometer values to IMA units (Bouten et al. 1997) — a unit that is demonstrated to correspond with daily energy expenditure — is implemented on the sensor's micro controller.

The smartphone application receives the activity data from the sensor, shows the activity data in a graph to the patient, and can trigger motivational cue messages as well as questionnaires for the user to answer. Based on the requirements phase described above, the software architecture developed was designed to form a flexible framework in which various types of applications can be defined by linking together various *modules* that support e.g. different sensors or different user interface functionalities. Each module performs a typically small, clearly defined task and delivers its output to a central communication module, or *Hub*. Other modules can subscribe to this information and are notified when new output becomes available. For example: a *BluetoothModule* is charged with the task of opening, maintaining and closing Bluetooth connections within the device (but knows nothing about specific communication protocols). As output it delivers a stream of data coming from the sensor, as well as updates regarding the status of the connection. The Graphical User Interface, implemented in a *GUIModule* can

5

subscribe to the status messages and provide the user with a warning if a Bluetooth connection is lost. Software developers are free to develop their modules without specific limitations and can communicate with other modules by passing messages through the Hub, to which other modules can subscribe.

All measured data (e.g. questionnaire results, activity data, message logs) can be synchronized to a central server that can store this data in a dedicated database. The server provides an API through which other applications, such as web portals can request data and provide more complicated views for patients as well as healthcare professionals. More details regarding the server, web portals, platform evaluations as well as discussions can be found in (op den Akker et al. 2012).

5.3 The multi-device system prototype

Multi-device systems, also known as *cross media systems*, are systems that extend across a range of different devices. Such systems are gaining in popularity due to their ability to support human activities in a range of different contexts (Segerståhl 2008). However, new challenges arise regarding the interaction paradigm of switching between the different devices. Segerståhl summarizes the main challenges as follows (Segerståhl 2008):

Heterogeneity: Users may need higher skills regarding information technology to use the various devices and applications in a multi-device system. Users may also have different expectations for different devices. Finally, the user interaction complexity grows because the user's primary task is determined by different variables depending on the device and the task context.

Interoperability: Besides interoperability in terms of connectivity, the conceptual definition of the roles of devices within the system, and the definition of how functionality is distributed, is equally important.

Consistency: Especially in the case of systems that are developed by multiple parties, there is a risk to lose interaction semantics and logic in terms of both industrial design and user interface aspects.

The challenges related to a multi-device setting relate to different levels of the system, The Open Innovation Platform (OIP) developed in the context of the

SOFIA project, provides an architecture that distinguishes three layers: the *device world* (devices, networks), the *service world* (applications, services) and the *smart world* (information-level) (Katasonov & Palviainen 2010). For the development of the multi-device system prototype (MDS Prototype) we will focus on the service world, and the smart world, i.e. on the communication between applications, and the consistency of information semantics on the different devices. The most important aspect that arises on these levels is inter-usability, or the ability to transition seamlessly between devices (Denis & Karsenty 2004). Inter-usability deals with the level of ease at which users can reuse their knowledge and skills for an action when using another device. Inter-usability encompasses two types of continuity which should be handled by a multi-device system: knowledge continuity and task continuity. The concept of *knowledge continuity* is about how easily users can understand how to perform certain functions supported by the system with the different devices, while *task continuity* is about storing intermittent task-operations to enable handing over incomplete tasks between devices.

The MDS Prototype is built to demonstrate the possibilities of leveraging the strengths of various interaction devices in a single health coaching application. The smartphone as coaching device — being the choice of device in previous interventions — can be considered a very *personal* device that is assumed to be ubiquitously available. These properties offer distinct advantages that motivate the choice of having the smartphone as a primary coaching device. But our target user, the sedentary office worker, interacts with many devices throughout his day, each offering unique advantages and disadvantages. In Table 5.1 we summarize some unique properties in terms of *privacy*, *availability*, *input*, and *output* for a list of possible interaction devices. For example, the *desktop PC* at work is a device that can be considered personal, as the user will be the only one able to sign on. The desktop PC is only available for coaching when the user is at his desk (situational). Regarding the inputs and outputs, sound and voice can be considered optional, depending on whether the user shares a workspace with one or more colleagues (in which case it would be inappropriate). The *activity sensor* is an example of an interaction device that is always with the user (personal, ubiquitous), but has only very limited input and output possibilities (e.g. a small display showing current steps, calories burned, and the clock, with a single button to cycle through display modes).

Device	**Privacy**	**Availability**	**Input**	**Output**
Smartphone	Personal	Ubiquitous	Touchscreen, (Voice)	Small screen, (Sound)
Desktop PC	Personal	Situational (work/desk)	Keyboard, Mouse (Voice)	Large screen, (Sound)
Kiosk PC	Shared	Situational (coffee-corner)	None	Large screen
Television	Personal, Shared	Situational (home)	Remote control	Large screen, Sound
Activity Sensor	Personal	Ubiquitous	Minimal	Minimal

Table 5.1: A selection of different interaction devices with different properties in terms of privacy, availability, input and output.

The MDS Prototype described here will target three of the devices listed in Table 5.1: the smartphone, desktop PC, and kiosk PC. The goal of the MDS Prototype is to leverage the advantages of the various involved devices, while overcoming some of the challenges, described above, that arise in multi-device systems, such as heterogeneity, interoperability, and consistency (Segerståhl 2008). First, we describe the functional design of the MDS Prototype (Section 5.3.1). Then we describe the technology aspects as well as the architecture underlying the system (Section 5.3.2), and finally describe the design and results of a short usability experiment using the prototype (Section 5.3.3).

5.3.1 Design

The design of the multi-device system prototype is described according to the four main requirements of *multiple devices*, *handover*, *knowledge continuity*, and *task continuity*.

Multiple devices

The obvious primary requirement for the multi-device system prototype is that it should support multiple devices. The devices with which our office worker persona interacts most frequently are his *smartphone*, and his *desktop PC* in the office. A third device that we chose to add is the *kiosk PC*, a standard computer with a large display hanging above the coffee machine in the central space of the office, where employees regularly meet for short (informal) conversations. The kiosk device adds a potentially interesting aspect, as the communication between this device and the user is not considered to be private (see Table 5.1). Together these three devices cover a broad spectrum of the different privacy, availability, input and output properties.

Smartphone: The smartphone device contains the multi-device coaching system's main application. For the current prototype, several functionalities, additional to the existing activity coaching application, are implemented. First, in order to use the smartphone's unique input- and output modalities, the application supports different input- and output-*realizers*. A realizer is a specific software implementation of an input- or output modality, and in the current prototype the user can switch between *text* and *speech* for both input and output. Additionally, the system allows the user to switch to an *avatar*, or embodied agent, as output modality, which is described in detail in Section 5.4. For speech output, the motivational messages, as well as the questionnaire question and possible answers are spoken by the system using the Android Text-to-Speech engine (enhanced by facial animations for the avatar option). When speech input is enabled, the system allows the user to use spoken language to answer the questionnaires by uttering either the specific answer, or the number of the answer (e.g. "the second option"). The Google Speech-to-Text service is used to implement this functionality. In the prototype, users can switch between input- and output modality using a menu option.

Desktop PC: The desktop PC is implemented as the first additional device for the multi-device system, as our office worker will interact with this device during most of his working day. The desktop application is run as a 'background' program with a taskbar icon and can provide motivational messages as well as questionnaires to the user. Input to the questionnaires can be given using the mouse to select the answers to a question from a drop-down menu (see Figure 5.5, left). The desktop supports a simple *text* output realizer, as well as a more advanced *avatar* (embodied agent) realizer that uses text-to-speech and facial animations to present the questionnaires. More details regarding this avatar is given in Section 5.4.

Kiosk PC: The Kiosk PC application runs on a standard desktop PC, but is attached to a large screen suspended above the coffee machine in the office workspace. This device can communicate with the user whenever he is present at the coffee corner. The coffee corner is considered to be a 'social hub', where employees meet for ad-hoc meetings or casual conversation. Providing feedback at this location means that the communication is directed towards everyone who is present, instead of just our primary user. This can be used to enhance interaction between various users of the system (an example of Inter-Human Interaction, see (op den Akker, Jones & Hermens 2014)), by e.g. showing the current activity progress of the user compared to all other users currently present at the coffee corner. The large public display also calls for quick- and easy to understand graphical user interfaces, through the use of *glanceable* feedback (Boerema et al. 2012). As the initial MDS Prototype evaluation is performed with a single user scenario, these features are not yet implemented. Instead, as a minimal working example of the device, the kiosk for now displays reminders about questionnaires waiting for user input.

Handover

The MDS Prototype is designed for a user to interact with only one device at a time. In order to provide a smooth interoperable experience for the user, handovers to the most relevant device occur automatically. In order to facilitate this, the system includes a server component, described in more detail in the architecture section (5.3.2), that handles requests from the individual devices to become the *active device*. By default, when the system is started, the smartphone device becomes the active device. Given no contextual information, we assume that the smartphone is the most likely to be in the vicinity of the user. Each device can then request to become active device based on knowledge of interacting with the user. The desktop PC will request to become active device whenever the user is using the mouse or keyboard, indicating that the user is working on his PC. The kiosk

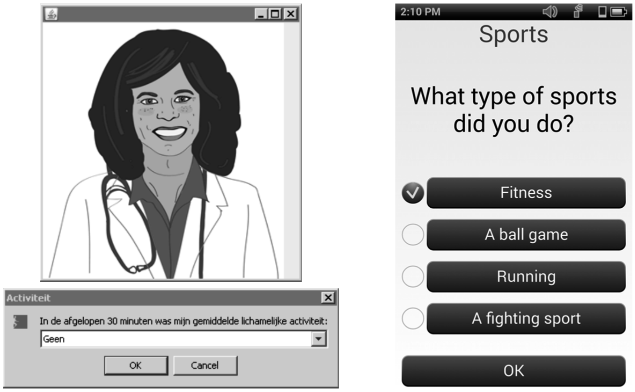


Figure 5.5: The desktop realizer (avatar and drop-down menu) for questionnaires on the left, and the basic smartphone realizer on the right.

PC, having no input modalities, has to rely on a proximity detection, for example requesting to become active device whenever a Bluetooth connection with the user's smartphone, or other portable device can be established. This functionality is currently not implemented, but is instead simulated in the evaluations (Section 5.3.3). Similar to the desktop application, the smartphone will request to become active device when interactions with the device are registered. The desktop and kiosk devices will actively relinquish the active device status after a period of user inactivity, in which case the smartphone device will again become active by default. Every handover action is handled and synchronized through the server component, ensuring that no two devices can assume to be the active device at the same time.

Knowledge continuity

In order to ensure knowledge continuity between the devices, the two main personal feedback devices (smartphone and desktop) share the same functionalities in terms of motivational message and questionnaire support. The third device — the kiosk — is different in terms of privacy- and input properties, making it unrealistic to extend knowledge continuity to this device as well. The simple text output realizers for motivational messages on smartphone and desktop PC are represented in very similar ways, as the message is simply printed on the screen, while allowing the user only to dismiss the message by touching the *OK* button on the smartphone, or clicking *OK* on the desktop. Larger differences are present with the representation of questionnaires. On the smartphone, the possible answers are presented as a list of buttons (see Figure 5.5, right), while on the desktop PC a drop-down menu is presented as the most intuitive way of interacting with this device. In order to increase continuity between these device, both support the use of an embodied agent for presenting the question to the user. The idea is that having the same "virtual coach agent" as representation on both devices will support the notion that the different devices are merely different output platform for *the same* virtual coach. This notion of having the same virtual coach presentation in multiple interfaces to accommodate persistence is one of the main features of the *Mobile Fitness Companion* described in (Ståhl et al. 2008, Ståhl et al. 2009, Turunen et al. 2011).

Task continuity

The notion of task continuity in multi-device systems relates to the ability of the user to recover the state of data and the context of an activity (Florins et al. 2004). The example from the MDS Prototype in which task continuity is of key importance is presented by the ability of users to engage in short health-related dialogues with the system, which can be interrupted at any time and continued on a different device. In practical terms, each of the devices are programmed to carry out short multi-stage questionnaires, to inquire the user about their current objective physical activity behavior. As an example, consider the following interaction between the system and the user:

System: Do you feel well rested at this moment?
User: Not very well.
System: What causes this in your opinion?
User: I was very busy today.
System: Did you perform any sports yesterday?
User: Yes.
System: What type of sports did you do?
User: I performed fitness exercises.

In order to accommodate task continuity, both the smartphone and desktop devices have the same questionnaires defined. During the execution of the questionnaire on either device, each step — questions asked, and answers given — are synchronized through the server with the other device. In this way, both devices can progress the internal state of the questionnaire. In case the user aborts the 'dialogue', he can then pick up where he left when switching to the other active device.

5.3.2 Architecture & technology

There are three common architectural frameworks for multi-device systems: a central server-based architecture, a decentralized architecture, or a hybrid approach. In a central setup, all data is stored and all interaction decisions are made on the server. In this setup, the clients are charged only with the actual user interactions. This approach is simple from a technology point of view, being a classic example of a multiple client-server architecture. The downside however is, that in case of a loss of connectivity, the entire system will fail. An opposite approach is to use a decentralized architecture in which each device acts as a "smart" autonomous agent. The agents would negotiate amongst themselves which is the most appropriate for engaging in interactions with the user. Such an approach is much more robust towards network failure, as each individual device will continue to operate by itself. It does however impose a considerable increase in complexity. The hybrid approach introduces a server that primarily handles the assigning of the *active device*, but also facilitates data synchronization. Each individual device stores a copy of its own data and makes its own decisions regarding initiating interactions with the user (granted that they are the *active device*).

For the MDS Prototype we have opted for a hybrid approach, introducing a central server for data synchronization and active device selection. Figure 5.6 shows the high level architecture of the system. The server's three primary responsibilities are handled by the *Connection Manager*, the *Data Storage* module and the *Active Device Selector*. Each *Device* is implemented using the Hub-architecture, as used by the smartphone application in the C3PO platform, explained in Section

5.2. Each device has its own *Local Storage*, a *Connection Module* dealing with the client-server communication, and a number of *Functional Modules* that contain the device's main functionalities.

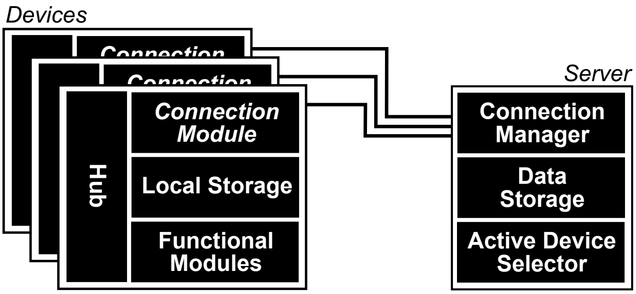


Figure 5.6: High level architecture of the Multi-Device System Prototype showing the Server and Device's main building blocks.

Each of the client device applications are implemented using the same Hub architecture, and are very similar from an architectural point of view. Figure 5.7 shows the modular building blocks of the most 'complicated' of the three devices: the smartphone application. The modules and their functionalities can be grouped into three broad categories: data collection, user interaction and multi-device related modules.

The *data collection* modules are unique to the smartphone application and deliver the user's activity data as well as contextual data. Activity data is collected through the *Bluetooth*, *ProMoveReader*, and *IMAData* chain of modules that handle the Bluetooth connection, the sensor data protocol, and the storage of activity data respectively. The *Location* module provides high level location information so that the *Weather* module can provide data about the current weather at the user's current location. This information can be used to provide more relevant motivational messages to the user, but has no primary function for the current system.

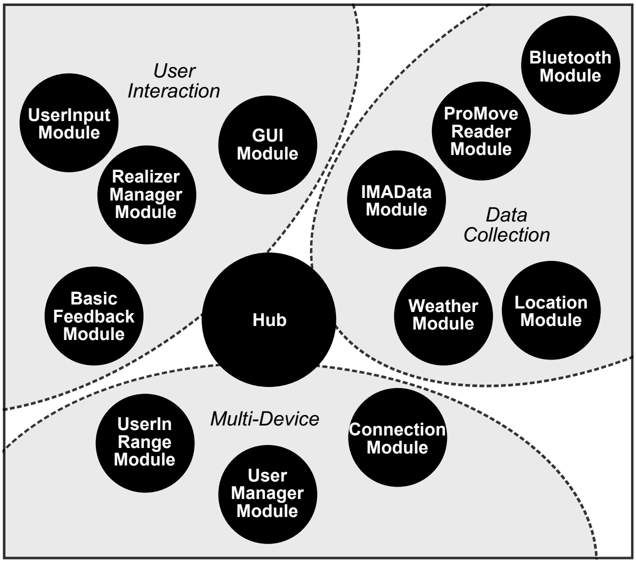


Figure 5.7: Modular Hub-based architecture of the smartphone device application. Every module communicates directly to the Hub, which takes care of passing messages to relevant receivers.

The *interaction* related modules are present in each of the three different device applications and together handle the communication with the user. The *UserInput* module contains a predefined schedule that is used to trigger certain interactions. An interaction in our prototype can either be a motivational message (handled by the *BasicFeedback* module), or a dialogue in the form of a questionnaire (handled by the *GUI* module). Besides triggering interactions, the *UserInput* module also takes care of storing intermediate questionnaire results, implementing jointly with the *IMAData* module the 'Local Storage' functionality described in the high level architecture overview (Figure 5.6). The *RealizerManager* module handles for each different device application the selection of different input- and output realizers, allowing the user to e.g. switch between spoken input and touch input on the smartphone, or to switch between text-based messages and the embodied agent visualization of motivational messages in the desktop application.

The *multi-device* related modules handle the functionalities that enable user-handover and active device selection from the client application side. The *Connection* module handles the network connection to the server. The *UserManager* module provides the decision support functionality for requesting user-handovers, i.e. requesting to become the active device. The *UserInRange* module facilitates this process by providing information about whether the user is currently able to interact with the current device.

The most valuable aspect of the architecture described here is the re-usability of modules that provide the same functionality regardless of the specific device implementation. This feature simplifies the development of the client applications and makes it relatively straightforward to add new types of devices to the prototype, especially when such a device supports a Java application (such as e.g. an Android tablet). The main difference between each application is the implementation of the *GUI* module, that deals with the actual *representation* of the communication (the topic of this chapter). Another notable difference in the module implementation relates to the *UserInRange* module that uses different ways of determining whether the user is currently able to interact with the device. For example, the smartphone application could use the phone's accelerometer to determine whether the user is carrying the phone, the desktop application could use keyboard/mouse activity (as implemented in the current prototype), while the kiosk application would need to implement a more sophisticated localization service.

5.3.3 Evaluation

The MDS Prototype was evaluated in a small scale technical- and usability study with six participants (four female, two male), aged between 24 and 54. Five of the six participants work in an office environment, while the other one was a student. All participants ranked themselves as being proficient with technology, and three of the six participants owned a smartphone, which they carried with them almost all the time.

The evaluations were conducted as a scripted scenario, in which participants were asked to perform an office task at the desktop PC, get a coffee at the coffee corner, and go for a walk around the office. Participants were informed about the basic functionalities of the system, being aware that they would receive feedback on physical activity, as well as activity related questionnaires. The test moderator recorded the user while interacting with the system, and guided the participants in performing a "thinking aloud" test procedure. The *thinking aloud* procedure is a method where the user verbalizes his thoughts during the use of a new technology, and it is deemed particularly effective in early exploratory research, such as prototype testing (Rubin & Chisnell 2008). The test scenario consisted of a total of six tasks that were guided partly by the scripted feedback- and questionnaire triggers in the desktop and smartphone device, and partly by the test moderator. Figure 5.8 shows the activity flow of the test scenario for participant 1. Activity flows for the other test participants are similar, except for minor variations.

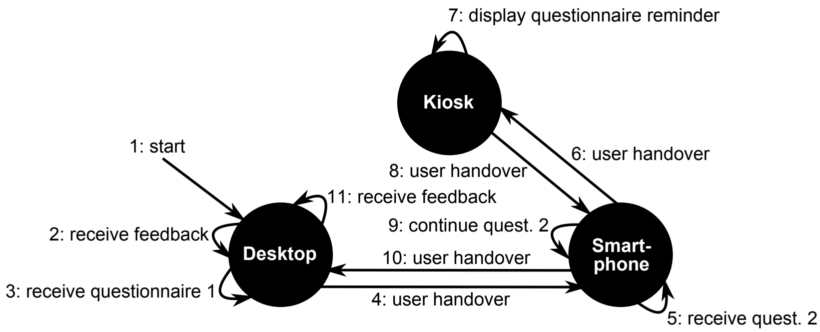


Figure 5.8: Activity flow for the evaluation scenario of one of the six test participants.

Overall, there was variety in the user's preferences towards either the desktop or smartphone device. Two users preferred to answer the questionnaires on the smartphone, one preferred the desktop, while the others had no preference. All participants understood how to interact with the devices after the initial short explanation of the prototype. From the small scale evaluations the main lessons learned are as follows. The detection of the user being at the desktop PC (through mouse and keyboard usage) was not always working correctly, causing in some cases excessive handovers between smartphone and desktop while the user was e.g. reading from the desktop screen. This may be improved by simply tuning some of the parameters in the desktop-usage and smartphone-usage detection algorithms. However, combined with the fact that we had to mimic the detection of the user at the kiosk PC, the overall issue of reliably detecting which device the user is currently engaged with, seems an important topic that needs to be further addressed. Furthermore, three of the six participants did not — by themselves — bring their smartphone when going to the coffee corner or a walk through the office, making a proximity detection approach between smartphone and other devices not an ideal solution. Our idea is that polling the *activity sensor*'s Bluetooth address from the desktop and kiosk PC, while assuming the smartphone to be the active device when not in range, might be the most promising approach.

Unfortunately, switching between different realizers in the application was not used by any of the participants. For future evaluations we believe it would be most important to develop input- and output realizers that maximize consistency, and the feeling of interacting with the same system across the various devices. Perhaps the most suitable option for this would be the use of the same embodied agent on each of the devices. Context awareness can be used to e.g. switch between spoken output in a private setting, and text output when the user is in a public space. The kiosk device's functionality implemented for this prototype was to remind users when an open questionnaire is available to answer. This reminder at the coffee corner did actually prove useful in one of the evaluation runs.

The proof of concept for the multi-device system prototype was successful. Except for minor, internet-access related issues, the prototype worked as intended and no major usability issues arose from the small scale evaluations. From a design and development perspective, the C3PO platform offered a useful architecture that allowed for a quick implementation and we hope to extend the prototype in the context of future research projects.

5.4 The embodied agent prototype

An embodied agent is a virtual, computer generated, graphical representation of a character that can interact with the user through natural language. The use of an embodied agent is one of the possible, promising ways of implementing the representation aspect of a coaching system for physical activity promotion. The embodied agent platform, called *Elckerlyc*, developed at the Human Media Interaction group of the University of Twente (Klaassen, Hendrix, Reidsma, op den Akker, van Dijk & op den Akker 2013) was also used in the MDS Prototype described above (Section 5.3). In this section we discuss the embodied agent prototype as an extension of the smartphone component of the C3PO platform, as used in a small two-week trial with 14 office workers.

The term 'embodied agent' usually includes the 'conversational' component: *embodied conversational agent* or ECA. The motivation for using ECAs as a way of delivering the coaching in a health behavior change intervention is perhaps best phrased by Bickmore in 2009 as "*. . . the most effective technologies are those which come closest to the "gold standard" of one-on-one, face-to-face counseling with an expert health provider*" (Bickmore et al. 2009). But research into ECAs dates back to the late 1990's where the first embodied agent systems were developed to help users find information on specific, limited domains (Beskow & McGlashan 1997, Thórisson 1997, Cassell et al. 2000). Even in the beginning of the 2000's, the first mobile versions of embodied agents were developed for PDA's (Kadous & Sammut 2002, Chittaro et al. 2006, Tomlinson et al. 2006, Bickmore 2002), although the limitations of these devices often required help from a back-end server to facilitate heavy processing.

Embodied Agents are also used in the field of physical activity promotion. In a review on tailoring for real-time physical activity promotion (op den Akker, Jones & Hermens 2014), four of the 12 described applications use embodied agents as the chosen form of representation of the coaching system. The MOPET system (Buttussi et al. 2006, Buttussi & Chittaro 2008) uses a full body agent to present feedback to the user, but also to provide exercise guidance by showing 3D animations of how to perform certain exercises along a fitness trail. The Handheld Exercise Agent (Bickmore et al. 2009) by Bickmore uses an embodied agent to stimulate the user to perform more daily physical activity (walking). In both these systems, the embodied agents represents some health coach, while the NEAT-o-Games (Fujiki et al. 2007, Fujiki et al. 2008) and Move2Play (Bielik et al. 2012) applications use the embodied agent more as a representation of the user himself.

The physical activity coaching application of the C3PO platform is designed to provide real-time coaching on daily physical activity behavior. Therefore, the embodied agent prototype described here is focused on reinforcing the idea of the application as a *coach*. Below we will shortly summarize the design of the embodied agent prototype (Section 5.4.1) as an integration of the *Elckerlyc* platform (Klaassen, Hendrix, Reidsma, op den Akker, van Dijk & op den Akker 2013) and the C3PO platform (op den Akker et al. 2012). In Section 5.4.2 we will focus on a small-scale evaluation of the system with 14 office workers.

5.4.1 Design

The embodied agent prototype is an integration of the *Elckerlyc* embodied conversational agent platform, developed at the Human Media Interaction department of the University of Twente and the C3PO telemedicine platform, developed at Roessingh Research and Development. The Elckerlyc platform is a Behavior Markup Language (Vilhjalmsson et al. 2007) realizer for real-time generation of behaviors of virtual humans. Originally Elckerlyc was designed as a rich platform for generating multi-modal verbal and nonverbal behavior with a focus on continuous interaction and precise temporal and spatial control between output modalities (van Welbergen et al. 2010, Reidsma et al. 2011, Reidsma & van Welbergen 2011), and was built to run on a standard desktop computer. Recent efforts by Klaassen et al. have been made to develop a lightweight version of the agent platform to run on Android smartphones. In this version, many of the architectural components remain unchanged, but the output realizer is replaced by the lightweight *PictureEngine*. Figure 5.9 depicts the realization of the Elckerlyc platform on the desktop (left), and the light weight version developed for Android phones on the right. These development efforts are documented in detail in (Klaassen et al. 2012, Klaassen, op den Akker & op den Akker 2013, Klaassen, Hendrix, Reidsma, op den Akker, van Dijk & op den Akker 2013), and examples of use are given for diabetic type II patients, as well as our office worker scenario.

Figure 5.9: The desktop realization of the Elckerlyc platform (left), and the lightweight PictureEngine version for Android phones (right) (Klaassen, Hendrix, Reidsma, op den Akker, van Dijk & op den Akker 2013).

The integration of the Elckerlyc lightweight PictureEngine into the Android application that is part of the C3PO platform, was done by utilizing the Android application's modular architecture. In this case, the integration meant extending the *GUI Module* (see Figure 5.7) to include the classes of the Android version of the Elckerlyc platform. Using a configuration option, the application could be run to either display it's motivational messages in the standard plain text representation, or by using the embodied agent, including a text-to-speech implementation using the Google TTS engine. As the goal of this study was to evaluate the potential added value of an embodied agent on a mobile health coaching system, this different representation of the messages was the only change implemented. The configuration of the *ECA system* and the *Basic system* were otherwise identical. The application is set to provide a motivational message every hour between 09:00 AM and 22:00 PM, selected from a list of predefined messages. Some messages could only be generated during office hours (09:00 - 17:00), and some messages could only be provided after office hours.

5.4.2 Evaluation

The evaluation of the Embodied Agent Prototype targets healthy office workers, or adults with a sedentary profession. The main goal of the evaluations was to answers the following questions:

- Does the use of an embodied agent lead to an increase in physical activity?
- Does the use of an embodied agent improve the user experience?
- Does the use of an embodied agent increase the perceived credibility of the system?

Methodology

The experiment performed has two conditions: *plain feedback*, and *embodied agent feedback*, and follows a within-subjects design, allowing participants to make an explicit comparison between the two experiment conditions. A counterbalanced measures design is used to limit the influence of the order in which the participants are exposed to the two forms of feedback. Half of the participants start with the plain feedback, while the other half starts by receiving the embodied agent feedback. The total experiment lasted for two weeks, switching between conditions after the first week.

Several types of data were collected during the experiment. First, the smartphone application logs the physical activity performed by the participants in IMA-counts per minute. Because of the relatively short duration of the experiment however, a significant physical activity behavior change is not expected. More importantly, the system logs the start-up- and shutdown times of the application, as well as the times at which feedback messages are presented, and the time at which they are observed. More subjective measures form the most important source of information. Each participant is required to fill in a first survey before starting the experiment, measuring basic user characteristics, the user's stage of change, and their technological skills. A second survey is completed after the first condition is finished (after one week), as well as after the second condition was finished. This surveys consist of four questionnaires, asking the participants to rate the previously observed condition on aspects of user experience (AttrakDiff2), credibility (SCS), acceptance (UTAUT), and quality of coaching (modified CBS-S).

The AttrakDiff2 user experience questionnaire is presented in the form of a semantic differential scale consisting of 28 bipolar pairs of adjectives, describing extreme ends on the spectrum of different aspects of system quality and attractiveness (Hassenzahl et al. 2003). The SCS or Source Credibility Scale (12 items) is used to measure perceived credibility of the system (McCroskey 1966), an important aspect and motivation for the use of embodied agents in general. In order to measure system acceptance, a subset of the Unified Theory of Acceptance and Use of Technology scale (UTAUT) questionnaire is used. The questions included relate to effort expectancy, attitude towards using technology, anxiety, and behavioral intention to use the system. Finally, a modified version of the Coaching Behavioral Scale for Sport (CBS-S) (Côté et al. 1999) questionnaire to measure the quality of coaching from the system. The modified version targets digital coaching in particular, and is developed by Philips as part of the DirectLife[4] project. It consists of 21 statements related to certain behaviors of the coach.

After the second condition has finished, each participant is presented with an additional questionnaire that aims to evaluate a direct comparison between the two conditions of plain feedback and embodied agent feedback. Finally, open interviews were held with each of the participants in order to discover any issues that were not covered by the questionnaires, and could not be observed by the logged data.

Results

A total of 14 participants were recruited from the University of Twente and Roessingh Research and Development offices. Of the 14 participants, 8 were male, and ages varied between 22 and 61. Unfortunately, none of the participants were in the *action* stage of change; with 2 participants in *precontemplation*, 4 in *contemplation*, 3 in *preparation*, and 5 in the *maintenance* stage of change.

The four individual questionnaires on user experience, credibility, acceptance and quality of coaching as well as the additional comparative questionnaire formed the most important source of results for the evaluation. Unfortunately, for many of the statements in the various questionnaires, no significant differences between

[4]http://www.directlife.philips.com/

the two conditions were observed. The AttrakDiff2 user experience questionnaire resulted in only one of the 28 aspects being rated with a significant difference between the plain feedback and embodied agent feedback. The {*complicated – simple*} scale showed a significant difference (p = 0.035) in favor of the plain feedback condition. For the Source Credibility Scale (SCS), again most differences were minor. When looking at the total scores over all constructs, the embodied agent scored 4.51, versus 4.83 for the plain feedback version (higher is better). This indicated a significant difference (p = 0.007) in favor of the plain feedback version. Looking at the individual pairs, the only significant differences found were that the plain feedback scored higher on the {*unpleasant – pleasant*} scale, and the {*awful – nice*} scale. Both the acceptance questionnaire (UTAUT) and the quality of coaching (modified CBS-S) questionnaires showed no significant differences in any of the constructs.

The final, comparative questionnaire showed the most significant results. Participants were asked to make a direct comparison between the two feedback conditions on 14 statements. All positively phrased statements (e.g. "Please indicate which version you though was more pleasant to use") were judged in favor of the plain feedback condition, while all negatively phrased statement (e.g. "Please indicate which version you though was more irritating") were judged "in favor" of the embodied agent condition. Of these 14 statements, six are deemed statistically significant with p-values below 0.05. Participants would prefer to use the plain feedback version for longer periods of time, thought it was more pleasant to use, and thought it gave better advice. For the embodied agent version, participants would ignore messages more often, thought it was more irritating and thought it was more cumbersome to use.

The exact reasons for participants to prefer the plain feedback version of the embodied agent prototype became more clear after the interviews. The main reason given by all participants can be summarized as *glanceability*. With the embodied agent version, the user had to wait for the agent to utter the motivational message, while with the plain feedback version, the message could be seen at a glance. This issue becomes bigger over time, as the messages become predictable due to a lack of variety. Given the very limited 'conversational' nature of the communication, the embodied agent is perceived as "*a bit of a gimmick*". Other comments that were given regularly relate to the quality of the speech output and the visual appearance of the agent. On the positive side, three participants did mention that the embodied agent prototype felt more personal, by adding personality to the system.

5.5 Discussion & Conclusion

The two prototypes described here are two first attempts at developing and testing new representations for the physical activity coach. The first prototype demonstrates how the addition of a multi-device server and a number of functional modules in the interaction device architecture can be used to extend the C3PO platform to communicate seamlessly across devices. The second prototype shows the integration of an embodied agent into the activity coaching smartphone application (Klaassen, Hendrix, Reidsma, op den Akker, van Dijk & op den Akker 2013), and focuses on the assessment of the effects of the embodied agent on user experience. Both prototypes are in an early stage of development but their evaluations point out a number of future research directions.

For the multi-device prototype, the development process and evaluations showed the advantage of the modular and platform-independent JAVA based application architecture. The most important aspect in terms of developing applications for multiple devices in a multi-device setup is leveraging the output (representation) and input modalities of the various devices. The modular architecture allows to reuse application components related to client-server communication, data storage, as well as *smart* components dealing with the generation of coaching communication. For example, a module that implements self-learning feedback timing such as the Kairos system discussed in (op den Akker, Tabak, Jones & Hermens 2014) or a module for the generation of motivational message content as discussed in (op den Akker, Cabrita, op den Akker, Jones & Hermens 2014) can simply be embedded into each device application. In the evaluations we showed that the users understood the concept of device handover, and positively showed a variety in preferences towards using the desktop or smartphone device for their interactions.

Unfortunately, the lessons learned from the MDS prototype evaluations were mostly related to the technical functioning of the system. Further development efforts are needed to fully explore and leverage the unique advantages of coaching on different devices. The main motivation for using different devices is that e.g. a desktop device allows providing detailed information in a richer environment. In the current implementation of the prototype this was not sufficiently addressed, as the functionalities of the desktop device were largely the same as those on the smartphone. The same applies to the kiosk PC, where the potential benefits of providing group-based feedback was not addressed in the current prototype. Future developments should focus on the creation of user interfaces specifically tailored to the device. The desktop PC could provide more detailed overviews of the user's activity behavior as well as richer motivational coaching focusing on health benefits and meaningful dialogues about physical activity behavior between the system and user. These detailed overviews should use the large screen display capabilities of the device, while the processing power of the desktop device can be used to support a realistic embodied agent supporting natural interaction.

The second prototype described in this work discussed the use of an embodied agent as virtual coach in the office worker scenario. The user evaluations showed a preference for receiving motivational messages in a plain text representation. We believe that this is due to the fact that the embodied agent representation does not match the content it is meant to convey. In the current prototype the motivational messages are short and offer little variation. This means that after some time, message content will start repeating and the purpose of the messages is reduced to a simple reminder (or *cue*) to draw attention to physical activity behavior. Most likely, an embodied agent representation has a higher probability of showing a positive effect on usability or health outcomes if the intention of the communication is more than a simple reminder. Research should focus on identifying the need for a digital coaching system to have meaningful health dialogues with the user. Only when there is clear evidence that such interactions could positively influence user's physical behavior change, further research into the development of virtual embodied agents would be warranted.

Acknowledgments

The authors would like to thank Wilko Wieringa and Jordi Hendrix for their efforts in developing and evaluating the two prototypes.

Bibliography

Benyon, D. & Macaulay, C. (2002), 'Scenarios and the HCI-SE design problem', *Interacting with Computers* **14**(4), 397–405.

Benyon, D., Turner, P. & Turner, S. (2004), *Designing Interactive Systems - People, Activities, Contexts, Technologies*, Pearson Education.

Beskow, J. & McGlashan, S. (1997), Olga - a conversational agent with gestures, *in* 'Proceedings of the IJCAI'97 workshop on Animated Interface Agents: Making them Intelligent', Nagoya, Japan.

Bickmore, T. W. (2002), Towards the Design of Multimodal Interfaces for Handheld Conversational Characters, *in* 'CHI '02 Extended Abstracts on Human Factors in Computing Systems', Minneapolis, Minnesota, USA, pp. 788–789.

Bickmore, T. W., Mauer, D. & Brown, T. (2009), 'Context Awareness in a Handheld Exercise Agent', *Pervasive and Mobile Computing* **5**(3), 226–235.

Bielik, P., Tomlein, M., Krátky, P., Mitrík, v., Barla, M. & Bieliková, M. (2012), Move2Play: an innovative approach to encouraging people to be more physically active, *in* 'Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium', Miami, Florida, USA, pp. 61–70.

Blangsted, A. K., Søgaard, K., Hansen, E. A., Hannerz, H. & Sjøgaard, G. (2008), 'One-year randomized controlled trial with different physical-activity programs to reduce musculoskeletal symptoms in the neck and shoulders among office workers', *Scandinavian Journal of Work, Environment & Health* **34**(1), 55–65.

Boerema, S. T., Klaassen, R., op den Akker, R. & Hermens, H. J. (2012), Glanceability Evaluation of a Physical Activity Feedback System, *in* 'Proceedings of the Sixth International Symposium on e-Health Services and Technologies (EHST 2012)', Vol. 2, Geneva, Switzerland, pp. 52–57.

Bosch, S., Marin-Perianu, M., Marin-Perianu, R., Havinga, P. J. & Hermens, H. J. (2009), Keep on Moving! Activity Monitoring and Stimulation Using Wireless Sensor Networks, *in* 'Proceedings of the 4th European Conference on Smart Sensing and Context (EuroSCC 2009)', Vol. 5741 of *Lecture Notes in Computer Science*, Guildford, United Kingdom, pp. 11–23.

Bouten, C. V. C., Koekkoek, K. T. M., Verduin, M., Kodde, R. & Janssen, J. D. (1997), 'A Triaxial Accelerometer and Portable Data Processing Unit for the Assessment of Daily Physical Activity', *IEEE Transactions on Biomedical Engineering* **44**(3), 136–147.

Buttussi, F. & Chittaro, L. (2008), 'MOPET: A Context-Aware and User-Adaptive Wearable System for Fitness Training', *Artificial Intelligence in Medicine* **42**(2), 153–163.

Buttussi, F., Chittaro, L. & Nadalutti, D. (2006), Bringing mobile guides and fitness activities together: a solution based on an embodied virtual trainer, *in* 'Proceedings of the 8th conference on Human-computer interaction with mobile devices and services', Espoo, Finland, pp. 29–36.

Carroll, J. M. (2000), 'Five Reasons for Scenario-Based Design', *Interacting with Computers* **13**(1), 43–60.

Cassell, J., Bickmore, T. W., Vilhjálmsson, H. & Yan, H. (2000), More Than Just a Pretty Face: Affordances of Embodiment, *in* 'Proceedings of the 5th International Conference on Intelligent User Interfaces', New Orleans, LA, USA, pp. 52–59.

Chan, C. B., Ryan, D. A. & Tudor-Locke, C. (2004), 'Health benefits of a pedometer-based physical activity intervention in sedentary workers', *Preventive Medicine* **39**(6), 1215–1222.

Chittaro, L., Buttussi, F. & Nadalutti, D. (2006), MAge-AniM: a system for visual modeling of embodied agent animations and their replay on mobile devices, *in* 'Proceedings of the Working Conference on Advanced Visual Interfaces', Venice, Italy, pp. 344–351.

Côté, J., Yardley, J., Hay, J., Sedgwick, W. & Baker, J. (1999), 'An exploratory examination of the coaching behavior scale for sport', *Avante* **5**(3), 82–92.

Denis, C. & Karsenty, L. (2004), Inter-Usability of Multi-Device Systems - A Conceptual Framework, *in* A. Seffah & H. Javahery, eds, 'Multiple User Interfaces: Cross-Platform Applications and Context-Aware Interfaces', John Wiley & Sons, chapter 17, pp. 373–385.

Evering, R. M. H. (2013), Ambulatory Feedback at Daily Physical Activity Patterns: A Treatment for the Chronic Fatigue Syndrome in the Home Environment?, PhD thesis, University of Twente.

Florins, M., Trevisan, D. G. & Vanderdonckt, J. (2004), The Continuity Property in Mixed Reality and Multiplatform Systems: A Comparative Study, *in* 'Proceedings of the Fifth International Conference on Computer-Aided Design of User Interfaces (CADUI'2004)', Funchal, Isle of Madeira, pp. 323–334.

Fujiki, Y., Kazakos, K., Puri, C., Buddharaju, P. & Pavlidis, I. (2008), 'NEAT-o-Games: Blending Physical Activity and Fun in the Daily Routine', *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment* **6**(2), 1–22.

Fujiki, Y., Kazakos, K., Puri, C., Pavlidis, I., Starren, J. & Levine, J. (2007), NEAT-o-Games: Ubiquitous Activity-Based Gaming, *in* 'CHI '07 Extended Abstracts on Human Factors in Computing Systems', San Jose, California, USA, pp. 2369–2374.

Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I.-M., Nieman, D. C. & Swain, D. P. (2011), 'American College of Sports Medicine position stand. Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise', *Medicine & Science in Sports & Exercise* **43**(7), 1334–1359.

Hassenzahl, M., Burmester, M. & Koller, F. (2003), AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität, *in* G. Szwillus & J. Ziegler, eds, 'Mensch & Computer: Interaktion in Bewegung', Vieweg+Teubner Verlag, pp. 187–196.

Huis In't Veld, R. M., Widya, I. A., Bults, R. G., Sandsjö, L., Hermens, H. J. & Vollenbroek-Hutten, M. M. (2010), 'A scenario guideline for designing new teletreatments: a multidisciplinary approach', *Journal of Telemedicine and Telecare* **16**(6), 302–307.

Kadous, M. W. & Sammut, C. (2002), Mobile Conversational Characters, *in* 'Proceedings of HF2002: Virtual Conversational Characters: Applications, Methods, and Research Challenge', Melbourne, Australia, pp. 1–4.

Katasonov, A. & Palviainen, M. (2010), Towards Ontology-Driven Development of Applications for Smart Environments, *in* '8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)', Mannheim, Germany, pp. 696–701.

Klaassen, R., Hendrix, J., Reidsma, D. & op den Akker, R. (2012), Elckerlyc goes Mobile: Enabling Technology for ECAs in Mobile Applications, *in* 'The Sixth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies', Barcelona, Spain, pp. 41–47.

Klaassen, R., Hendrix, J., Reidsma, D., op den Akker, R., van Dijk, B. & op den Akker, H. (2013), 'Elckerlyc Goes Mobile - Enabling Natural Interaction in Mobile User Interfaces', *International Journal on Advances in Telecommunications* **6**(1 & 2), 45–56.

Klaassen, R., op den Akker, R. & op den Akker, H. (2013), Feedback Presentation for Mobile Personalised Digital Physical Activity Coaching Platforms, *in* 'Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA 2013)', Island of Rhodes, Greece.

Levine, J. A. & Miller, J. M. (2007), 'The energy expenditure of using a "walk-and-work" desk for office workers with obesity', *British Journal of Sports Medicine* **41**(9), 558–561.

McCroskey, J. C. (1966), 'Scales for the measurement of ethos', *Speech Monographs* **33**(1), 65–72.

op den Akker, H., Cabrita, M., op den Akker, R., Jones, V. M. & Hermens, H. J. (2014), 'Tailored Motivational Message Generation: A Model and Practical Framework for Real-Time Physical Activity Coaching', *Journal of Biomedical Informatics (submitted)* .

op den Akker, H., Jones, V. M. & Hermens, H. J. (2014), 'Tailoring real-time physical activity coaching systems: a literature survey and model', *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change (online)* .

op den Akker, H., Tabak, M., Jones, V. M. & Hermens, H. J. (2014), 'Reaching Kairos: Adaptive Prediction of the Opportune Moment for Stimulating Physical Activity', *... (submitted)* .

op den Akker, H., Tabak, M., Marin-Perianu, M., Huis In't Veld, R. M., Jones, V. M., Hofs, D., Tönis, T. M., van Schooten, B. W., Vollenbroek-Hutten, M. M. & Hermens, H. J. (2012), Development and Evaluation of a Sensor-Based System for Remote Monitoring and Treatment of Chronic Diseases - the Continuous Care & Coaching Platform, *in* 'Proceedings of the Sixth International Symposium on e-Health Services and Technologies (EHST 2012)', Geneva, Switzerland, pp. 19–27.

Reidsma, D., Kok, I., Neiberg, D., Pammi, S., Straalen, B., Truong, K. & Welbergen, H. (2011), 'Continuous interaction with a virtual human', *Journal on Multimodal User Interfaces* **4**(2), 97–118.

Reidsma, D. & van Welbergen, H. (2011), Elckerlyc in Practice - On the Integration of a BML Realizer in Real Applications, *in* 'Intelligent Technologies for Interactive Entertainment - Proceedings of the 4th International ICST Conference', Genova, Italy, pp. 83–92.

Rubin, J. & Chisnell, D. (2008), *Handbook of Usability Testing: How to plan, design and conduct effective tests*, Wiley Publishing.

Segerståhl, K. (2008), Utilization of Pervasive IT Compromised? Understanding the Adoption and Use of a Cross Media System, *in* 'Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia', Umeå, Sweden, pp. 168–175.

Ståhl, O., Gambäck, B., Hansen, P., Turunen, M. & Hakulinen, J. (2008), A Mobile Fitness Companion, *in* 'Proceedings of the Fourth International Workshop on Human-Computer Conversation', Bellagio, Italy.

Ståhl, O., Gambäck, B., Turunen, M. & Hakulinen, J. (2009), A mobile health and fitness companion demonstrator, *in* 'Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics', Association for Computational Linguistics, pp. 65–68.

Tabak, M. (2014), Telemedicine for Patients with COPD - New Treatment Approaches to Improve Daily Activity Behaviour, PhD thesis, University of Twente.

Thórisson, K. R. (1997), Gandalf: An Embodied Humanoid Capable of Real-Time Multimodal Dialogue with People, *in* 'Proceedings of the First ACM International Conference on Autonomous Agents', Marina del Rey, California, USA, pp. 536–537.

Tomlinson, B., Yau, M. L. & Baumer, E. (2006), Embodied Mobile Agents, *in* 'Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems', Hakodate, Japan, pp. 969–976.

Turunen, M., Hakulinen, J., Stå hl, O., Gambäck, B., Hansen, P., Rodríguez Gancedo, M. C., de la Cámara, R. S., Smith, C., Charlton, D. & Cavazza, M. (2011), 'Multimodal and mobile conversational Health and Fitness Companions', *Computer Speech and Language* **25**(2), 192–209.

Van Velsen, L., Wentzel, J. & Van Gemert-Pijnen, J. E. (2013), 'Designing eHealth that Matters via a Multidisciplinary Requirements Development Approach.', *JMIR research protocols* **2**(1), e21.

van Weering, M. G. (2011), Towards a New Treatment for Chronic Low Back Pain Patients - Using Activity Monitoring and Personalized Feedback, PhD thesis, University of Twente.

van Welbergen, H., Reidsma, D., Ruttkay, Z. M. & Zwiers, J. (2010), 'Elckerlyc - A BML Realizer for continuous, multimodal interaction with a Virtual Human', *Journal on Multimodal User Interfaces* **3**(4), 271–284.

Vilhjalmsson, H., Cantelmo, N., Cassell, J., Chafai, N. E., Kipp, M., Kopp, S., Mancini, M., Marsella, S., Marshall, A. N., Pelachaud, C., Ruttkay, Z., Thórisson, K. R., Van Welbergen, H. & Van Der Werf, R. J. (2007), 'The Behavior Markup Language: Recent Developments and Challenges', *Lecture Notes in Artificial Intelligence* **4722**, 99–111.

Chapter 6

General discussion

Chapter 6

Contributions and future work

Physical inactivity is a clear societal issue with a large impact. The envisioned solution of using technology to stimulate lasting behavior change is a promising one, but has not yet reached its maximum potential. Tailoring, or matching the coaching to individual user's needs, is regarded as a promising paradigm to increase the potential of physical activity coaching systems. But tailoring is a broad and complex concept. In this Thesis we aimed to shed light on the concept of tailoring in the field of physical activity coaching systems. The complexity of tailoring is highlighted in Chapter 2 where our literature survey showed the variety of tailoring methods employed in the various systems under review. As a first step in providing clarity and structure to the topic we developed a model of tailoring that describes seven tailoring concepts and the way in which they interact with each other. Perhaps more importantly, it is shown how each tailoring concept can be used in different ways to tailor to four properties of a communication instance: *timing*, *intention*, *content*, and *representation*. In addition to the development of the tailoring model itself, we use these four communication properties to discuss the further contributions of this Thesis.

Timing

The first communication property is *timing*. A system that aims to change the user's behavior needs to utilize the opportune moments to coach the user. Earlier studies conducted with the activity coaching smartphone application used a fixed-timing model, sending a motivational message to the user at regular fixed times throughout the day. In studies targeting chronic low back pain patients, these messages were configured to trigger once every hour (van Weering 2011). In later studies targeting chronic fatigue syndrome patients, this was configured to occur once every two hours (Evering 2013). But it is hard to say which configuration is better, and even more difficult to argue why these messages should sent exactly on the passing of the hour. The reason to focus on finding the opportune moment for providing feedback is best explained through examples of when things can go wrong. A user who receives a message encouraging him to go for a walk is unable to take action when he just got into his car to drive to work. If you are prompted to go for a walk in the park at the moment a heavy rainstorm started, you are also unlikely to do so.

In general the assumption is that everyone has a daily schedule in which there are moments that are more suitable, and moments that are unsuitable for receiving a reminder to change activity behavior. Since there are so many factors that can influence this, the idea of the Kairos module was to develop a system that takes as much as possible information regarding the user's context, learns a model of the user's responses in different contexts, and triggers the motivational messages at those times that are considered likely to yield a response. Our evaluation of the Kairos system in Chapter 3 shows that the system works technically, but only performs better — in terms of motivational message compliance — in a limited way. Still, we believe that the system can be considered successful. The previous, fixed-timing approach, has no structured way of improving (other than changing the time interval, or perhaps including some randomness). The Kairos system on the other hand, has many possible paths to improvement.

Overall, improvements can be made to the input, feature processing and selection, and training steps. The input to the system can be improved by providing a more accurate assessment of the user's activity behavior, including semantic knowledge of the user's activities such as walking, running, or cycling — but also classifications of driving, or use of public transport. Activity classification is an active field of research (Altun et al. 2010) that already lead to results in various commercially available sensors, such as the FitBit Flex[1] that tracks climbing stairs. The more accurately the user's daily routine can be assessed, the better the classifier's chances of finding the opportune moments to intervene. Additional input processing can help reduce the complexity of the input data and make life easier for the machine learning algorithms in the training phase. For example, weather information is provided to the system as a number of features describing e.g. cloud cover, minimum- and maximum temperature or humidity. However, from the point of view of predicting motivational message compliance, the humidity outside is not so much of interest, but rather the question of whether it is "nice weather to go outside". A pre-processing step on the weather feature vector could summarize all the weather related information into a single continuous variable that essentially describes a *weather rating*, reducing the input complexity to

[1]http://www.fitbit.com/flex/

the motivational message predictor. Finally, the evaluation of different machine learners in the Kairos system should be considered exploratory, and optimization and tuning could further improve results.

Intention and Content

On the subject of motivational coaching, it is perhaps obvious that it is important *what* to say to the user. Overall this can be seen as the topic of the 4th Chapter of this Thesis. Taking a step back, the first contribution of the chapter is a structured model of a motivational message. In the literature survey of Chapter 2, we identified the four communication properties, and gave examples of how various combinations of tailoring methods can be used to tailor to these specific properties. In Chapter 4 we address the problem of generating real-time tailored motivational messages according to a more detailed model of a message and its four properties of timing, intention, content and representation. The implementation of the framework focused on *intention* and *content*. The reasons for this are that *timing* was already addressed in detail in Chapter 3, while for the *representation* we chose to focus on a single modality — natural language text — to reduce the complexity of the examples and implementations. An analysis of existing literature on motivational messaging identified the relevant concepts that can influence the motivational quality of a message. Based on earlier work on modeling tailoring in Chapter 2 and specifically on motivational message modeling in Erriquez & Grasso (2008) and Wieringa et al. (2011) has guided the development of the proposed model in which the identified relevant concepts find its place. In the practical framework that was built from the model, the separation between intention and content allows for a structured, sequential way of generating motivational messages. This structural, sequential decomposition is useful because in every step of the process different tailoring methods can be applied. For example. the decision of whether or not to provide e.g. an *argument* (e.g. informing about the benefits of physical activity) is influenced by different tailoring methods (e.g. based on adaptation through stage of change) than the content of such an argument (e.g. user targeting the argument to the user's medical condition).

In contrast to the *data-driven* approach to tailoring in Chapter 3, the message generation framework in Chapter 4 allows for relatively simple rule-based decision making regarding each separate step in the message generation process. As such we believe that the framework can be used to quickly evaluate different tailoring methods in the various steps of the message generation process. In the Discussion (Section 2.7) of the literature survey on tailoring we identified that most of the tailoring methods applied in existing applications are relatively simple and intuitive, but their effectiveness remains often unproven. This issue of evaluation is one that remains problematic for two reasons. First, evaluating a behavior change application requires longitudinal studies that are time consuming and costly, and second, the tailoring method itself is just a cog in the wheel of an often complex intervention tool, making it complicated to attribute effectiveness of the overall intervention to any single part of the system. Although the model of motivational messages does not necessarily solve these problems, we do believe that it can contribute to a solution. Describing a behavior change intervention at the level of detail of the message model can enable a better comparison between different approaches to behavior change. For example, a system described as providing "*personalized feedback messages*" can be more specifically classified as implementing e.g. "*Name Mentioning* (Dijkstra 2014), a form of *User Targeting* applied to the *Argument* component of the *Content* of the *Motivational Message*". Furthermore, specific, isolated tailoring algorithms could be evaluated for technical and conceptual soundness using simulated user responses. For example, by simulating responses of persona's that are modeled to exhibit a range of exaggerated preferences towards e.g. name mentioning, or certain types of physical activities, or persona's that show a very fixed daily routine, researchers can test the validity of their methods before going through expensive user-based trials. Such simulation-based trials become increasingly necessary, in particular in case of complex solutions such as the Kairos system discussed earlier, but can also be used to test the validity of a simpler rule-based system.

The message generation framework is implemented in Java to run as a module in the Android based activity coaching application. Examples of rule-based decision components have been implemented, but many gaps remain to be filled. Every *end-node* in the message model, such as the Content — Follow-up — Suggestion — **Lifestyle Suggestion** node, presents a clearly defined research area in its own right. Future work in the development of the message generation framework will thus focus on carrying out the research in each of these end-nodes. Taking the

*Lifestyle Suggestion*s as an example, there is a need to define a structured set of lifestyle suggestions and parametrize them in such a way that they can be mapped to the user's preferences and personal lifestyle. Similar efforts need to be carried out to define the *Activity Suggestion* node of the model where suggestions should be filtered according to the capabilities, preferences and context of the user. Initial work on creating an ontology of activity suggestions and selecting relevant activities using probabilistic functions is presented in (Wieringa et al. 2011) and could form the basis of these efforts. The idea is that using an ontological representation of *activities* can help the system to quickly learn about user's preferences. For example, when the user reacts positively to messages that suggest to go to a supermarket, he could be more likely to also respond positively to going to the local market, as these activities would be represented close to each other in the activity ontology. Indeed, the system could infer that the user prefers outside activities over indoors activities in general and could consequently increase the likelihood of selecting those message in future communications.

The envisioned future work for the *timing* aspects of motivational coaching already mentions the importance of more accurate input data to improve performance. From the structure of the message generation framework, this need becomes even more apparent. Tailoring communication to the user is defined as deciding on the timing, intention, content or representation based on information from the user model, context model, interaction model, or domain model. We believe that the effectiveness of such tailoring is dependent on the information models just as much as on the implemented decision rules. Thus, in order to foster efficient tailoring, there is a need to develop complete, and easy accessible information models. Context models, containing information about the user's location, weather, and nearby points of interest and events could largely be filled automatically as demonstrated by services such as Google Now. Interaction models that contain information on the user's previous interactions with the system can also be populated automatically by efficient logging mechanisms. However, information that makes up the user model (containing e.g. user characteristics and preferences) and the domain model (containing information about the user that is relevant to the current domain, e.g. *physical activity*) are often more difficult to extract automatically. When information is difficult to obtain in an unobtrusive way (e.g. whether the user has a dog) the user needs to provide this information explicitly. We believe that instead of providing a complicated *settings* menu where this information can be provided, it is worthwhile exploring the possibility of obtaining the information through short, natural dialogues with the system. As the activity coaching application is meant to be a kind of *virtual coach*, such interactions may also be used as a tool to form a relationship between the user and the system, possibly increasing the overall effectiveness of the coach (Bickmore & Picard 2005).

Representation

Looking at the *representation* aspect of the communication, the concept of an embodied *virtual coach* is explored in the 5th Chapter of the Thesis. In the message model discussed above, the representation aspect is exemplified only with *natural language text*, as this was the modality used in the physical activity coaching application thus far. The representation, or user interface, to the coaching application was designed to be very simplistic, consisting for the primary part of a daily activity graph and short textual motivational messages. The reason for having such a minimalistic interface is that the target population for the intervention often consists of elderly users with limited technological skills. In Chapter 5 we focus on a target group that is more familiar with technology — office workers — allowing ourselves to experiment with more complex forms of representations.

We demonstrated how the modular architecture of the smartphone application can be used to quickly experiment with different, more complex forms of representation. Two prototype applications were developed, focusing on a multi-device setting, and an embodied agent representation respectively. The multi-device system prototype evaluation focused on technical evaluations, showing that the service worked, that users understood the concept of device handover, and showing promise in the variety of user preferences towards using certain devices. However, the system developed can be considered an early stage prototype, requiring a second round of development to implement additional functionalities that effectively leverages the strengths of the various devices involved. Development should focus on creating a "*social*" interface for the kiosk device, fostering competition by showing comparative activity progress of colleagues gathered at the coffee corner. Furthermore, the desktop application should be extended to provide in-depth overviews of activity progress over time. With such developments completed, we believe the multi-device system is a promising tool to stimulate physical activity among office workers by providing feedback and coaching at the desk, in social situations, and on the go.

The embodied agent prototype developed in collaboration with the Human Media Interaction group of the University of Twente, is another example of the ease of integration allowed by the smartphone architecture. The prototype was developed up to a more advanced stage, and evaluated with a focus on usability. The results of the evaluation show a preference for the 'plain text' condition in favor of the embodied agent representation. We believe that this is caused by a mismatch between representation and the content of the coaching provided. The only 'interaction' provided by the coaching application was in the form of a limited set of short, motivational messages that would quickly become predictable and repetitive. We believe that the embodied agent representation was too complex given the low-complexity of information presented. Compare this to looking up the time on a phone, which we expect to be available at a glance, while a more complex task such as starting a route guidance is acceptable to be achieved through a speech dialogue with the phone ("***U:*** *Okay Google. Navigate to work.* ***S:*** *Navigating to Roessinghsbleekweg 33b, Enschede*"). Developing natural interaction through e.g. embodied agents is then last on the list of future work efforts and should only be further developed when more complex interactions with the system are supported. An example of such an interaction scheme in the context of physical activity promotion could be a collaborative goal-setting approach, where the coach could initiate a dialogue about setting a new daily or weekly goal for the user.

Summary of future work

We summarize the future research directions based on the analysis, modeling, development and experiments performed on the topic of tailoring real-time activity coaching systems in this Thesis. Smart tailoring algorithms are only as good as the quality of the input data available. Thus (**1**) research should focus on developing and integrating technology to obtain better, semantically rich, relevant input data. In order to allow easy integration with e.g. a tailored message generation framework as defined in Chapter 4, (**2**) there is a need to build structured, easily accessible models containing *user-*, *context-*, *interaction-*, and *domain* information. Although an intelligent coach will be able to acquire implicit knowledge about the user and his context to fill these models, some information needs to be provided explicitly by the user. Thus, (**3**) the system should support meaningful dialogue in order to facilitate this knowledge acquisition while possibly providing a more engaging bond between system and user. To further increase engagement and the bond between system and user (**4**) such dialogues should be represented in a natural way, possibly providing a more suitable use case for embodied virtual agent technology. Finally, (**5**) novel evaluation methodologies should be explored that are fast and effective in highlighting the strengths and weaknesses of tailoring and interaction techniques.

Outstanding issues

Finally, there are two topics related to the work in this Thesis that have not yet received any attention. The first topic is gamification. Gamification, or the use of game-design elements in non-gaming contexts (Deterding et al. 2011) — such as health promotion systems — is a research field that is on a rapid rise since 2011 (Hamari et al. 2014). The review on tailored real-time physical activity coaching systems in Chapter 2 included applications that apply gaming elements to stimulate physical activity. The NEAT-o-Games system (Fujiki et al. 2008) is a good example, in which the user gathers "activity points" throughout the day which can be used to progress in a virtual race, or acquire hints in embedded cognitive games. The Move2Play system (Bielik et al. 2012) specifically mentions gamification as a design principle, describing the use of *achievements* (virtual rewards for long-term user efforts), *badges* (short-term recognition of user efforts) and *unlock-able features* (such as visual modifications of the user's avatar that are only available after reaching some activity goal). These techniques were not implemented or evaluated in the Move2Play system, and to the best of our knowledge there is no consensus on the effectiveness of gamification in physical activity promotion systems. However, the professional gaming industry has used concepts such as achievements, badges, and unlock-able content successfully for years to keep players engaged. Overall, video games have been extremely successful in keeping players engaged (Schoenau-Fog 2011) and it seems worthwhile to explore the methods and technologies used to achieve this goal. Even commercial activity tracking products such as FitBit already include gamification elements such as badges on their user web portal. It is difficult to say what the relationship is between gamification and tailoring without an extensive analysis of this cross-section, but it seems prudent to investigate how gamification techniques can be tailored to individual users.

The second point of discussion is privacy. In the context of eHealth and telemedicine applications, privacy issues are a common concern due to the processing and handling of patient data. Most of the work described in this Thesis focuses on the development of stand-alone, intelligent coaching services. *Stand-alone* in this case, means that the data collection and processing is designed to occur locally on the user's smartphone without the need for a data connection to a remote service. Within this paradigm, possible concerns regarding privacy could be limited by adopting standard security measures such as encrypting stored data on the smartphone. However, whenever a telemedicine component — connecting the activity coach to healthcare professionals — is included in an intervention, data storage, transfer and access policies become a more serious issue. Such issues should not be ignored, but are out of the scope of the current work to address properly. An exception to the local processing paradigm introduced in this work is presented in the Kairos system. Kairos uses, during the cold start phase, data from other users to provide the prediction of motivational message timing. In a practical application, the cold start classifier should be trained, and regularly updated by some third party service provider, meaning that possibly privacy sensitive data needs to be send from the user to this third party. A possible solution to this issue is presented by (de Hoogh et al. 2014) who have used a portion of the data used for the Kairos system to develop and evaluate algorithms for privacy-preserving data mining, effectively allowing predictions to be made based on encrypted data.

Conclusion

In this Thesis we performed an in-depth analysis of tailoring in the context of physical activity promotion systems. We developed a high-level model of tailoring, and a detailed model of motivational messages, their communication properties and the relationship to tailoring concepts. We designed, developed and evaluated an intelligent, tailored application that predicts the optimal timing of delivering motivational messages to the user and adjusts itself over time to the individual. A practical message generation framework was designed and implemented to tailor the content of such messages to the individual using *adaptation*, *context awareness*, *goal-setting*, *user targeting* and *self learning*. We demonstrated two advanced forms of representations for a coaching application in a multi-device setting and using an embodied agent. Lastly we discussed the limitations and prepared ourselves for future work. With this, we have contributed to a new way of coaching that is relevant to anyone dealing with a chronic condition or anyone aiming to reach or maintain a healthy physically active lifestyle.

Chapter 6

Bibliography

Altun, K., Barshan, B. & Tunçel, O. (2010), 'Comparative study on classifying human activities with miniature inertial and magnetic sensors', *Pattern Recognition* **43**(10), 3605–3620.

Bickmore, T. W. & Picard, R. W. (2005), 'Establishing and Maintaining Long-Term Human-Computer Relationships', *ACM Transactions on Computer-Human Interaction* **12**(2), 293–327.

Bielik, P., Tomlein, M., Krátky, P., Mitrík, v., Barla, M. & Bieliková, M. (2012), Move2Play: an innovative approach to encouraging people to be more physically active, *in* 'Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium', Miami, Florida, USA, pp. 61–70.

de Hoogh, S., Schoenmakers, B., Chen, P. & op den Akker, H. (2014), Practical Secure Decision Tree Learning in a Teletreatment Application, *in* 'Proceedings of the 18th International Conference on Financial Cryptography and Data Security', Bridgetown, Barbados.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K. & Dixon, D. (2011), Gamification. using game-design elements in non-gaming contexts, *in* 'CHI '11 Extended Abstracts on Human Factors in Computing Systems', Vancouver, Canada, pp. 2425–2428.

Dijkstra, A. (2014), 'The Persuasive Effects of Personalization through Name Mentioning in a Smoking Cessation Message', *User Modeling and User-Adapted Interaction, Special Issue on Personalization and Behavior Change (online)* .

Erriquez, E. & Grasso, F. (2008), Generation of Personalised Advisory Messages: An Ontology Based Approach, *in* 'Proceedings of the 21st IEEE International International Symposium on Computer-Based Medical Systems (CMBS)', Jyvaskyla, Finland, pp. 437–442.

Evering, R. M. H. (2013), Ambulatory Feedback at Daily Physical Activity Patterns: A Treatment for the Chronic Fatigue Syndrome in the Home Environment?, PhD thesis, University of Twente.

Fujiki, Y., Kazakos, K., Puri, C., Buddharaju, P. & Pavlidis, I. (2008), 'NEAT-o-Games: Blending Physical Activity and Fun in the Daily Routine', *Computers in Entertainment (CIE) - Theoretical and Practical Computer Applications in Entertainment* **6**(2), 1–22.

Hamari, J., Koivisto, J. & Pakkanen, T. (2014), Do Persuasive Technologies Persuade? - A Review of Empirical Studies, *in* 'Persuasive Technology, Proceedings of the 9th International Conference PERSUASIVE', Padua, Italy, pp. 118–136.

Schoenau-Fog, H. (2011), The Player Engagement Process - An Exploration of Continuation Desire in Digital Games, *in* 'DiGRA: Think Design Play', p. 18.

van Weering, M. G. (2011), Towards a New Treatment for Chronic Low Back Pain Patients - Using Activity Monitoring and Personalized Feedback, PhD thesis, University of Twente.

Wieringa, W., op den Akker, H., Jones, V. M., op den Akker, R. & Hermens, H. J. (2011), Ontology-Based Generation of Dynamic Feedback on Physical Activity, *in* 'Proceedings of the 13th Conference on Artificial Intelligence in Medicine (AIME)', Springer, Bled, Slovenia, pp. 55–59.

Appendices

Summary

The lack of physical activity in the general population is recognized internationally as an important societal issue. For apparently healthy adults, inactivity leads to overweight, and increased risk of numerous chronic and acute diseases including coronary heart disease, type 2 diabetes, as well as breast and colon cancers. One of many recommendations relating to increasing physical activity is that of the American College of Sports Medicine who advise that the majority of adults should perform moderate-intensity cardio respiratory exercise for at least thirty minutes each day. Over the last century, human lifespan has increased yearly with a rate of 0.24 years in the United States, Europe and Japan. This increase in longevity means that more and more people spend a large part of their lives at an older age, coping with chronic diseases such chronic obstructive pulmonary disease (COPD) or chronic back pain. The treatment for such chronic diseases often includes of a programme of daily physical activity.

This Thesis deals with the promotion of daily life physical activity — walking, cycling, gardening, or housework — all the activities regularly performed in everyday life that require a person to move his body. In recent years, researchers and policy makers have focused on technology as a tool to deliver coaching on physical activity. Technology in the form of web sites or smartphone applications can be used to reach many people at low costs in order to enable self-care for patients (eHealth) or to provide technology-mediated professional care in the home environment (Telemedicine). An important factor in the popularity of such technology mediated coaching is the rise of cheap, miniaturized wireless sensors that enable the tracking of a person's physical activity throughout the day. A second

important factor is the ubiquity of smartphones that can receive this activity data, process it, and provide user friendly feedback and coaching at any place and any time. Through the combination of these technologies, effective activity coaching tools have been developed and are now available on the market for reasonable prices.

Although the popularity of physical activity tracking and motivation is apparent in both research and commercial areas, the effectiveness of the coaching mechanism employed are far from perfect. It is widely believed that coaching strategies can be improved by catering to individual users' needs and preferences, a field of research known as *tailoring* or *personalization*. What tailoring is, and how it can be applied in the field of real-time coaching on everyday physical activity, are some of the issues addressed in this Thesis.

In **Chapter 1** a general rationale for the work in this Thesis is given. The issues of healthcare for the aging population, and the rise of the field of physical activity monitoring are explained in more detail. The focus of this Thesis on high-tech interventions — real-time coaching methods — is explained, as well as the need to improve their effectiveness through tailoring. The general introduction concludes by posing the four major research questions that are tackled in the four major Chapters of the Thesis.

Chapter 2 provides the theoretical framework of the Thesis through an extensive literature survey on the topic of tailored real-time activity coaching systems. The chapter investigates twelve different applications that each approach tailoring in various different ways. From the analysis of these applications, a model of tailoring is developed, describing a total of seven different tailoring concepts: *feedback*, *inter-human interaction*, *adaptation*, *goal-setting*, *user targeting*, *context awareness*, and *self learning*. Furthermore, a definition of a *communication instance* is given, encompassing aspects of *timing*, *intention*, *content*, and *representation*. For each communication to the user, it is shown how each of these aspects can be tailored in various different ways.

Each of the following chapters provide a focus based on the defined communication aspects. In **Chapter 3** the focus is on tailoring the timing aspect of communication in order to increase compliance to individual motivational messages — or *cues*. In this chapter the rationale, design, development and evaluations are described of a system that can decide when the optimal moment for providing a motivational cue to the user is. Based on the user and his current context, the system uses machine learning to classify instances of time as being opportune for initiating a message, and learns from its successes and failures. In a study with 10 COPD patients it is shown that this way of tailoring the timing to individual users helps increase the compliance to individual messages.

In **Chapter 4**, the next step towards a fully automated coach is taken, by focusing on the question of *what* to say to the user. Based on an analysis of literature and a corpus of example messages used in earlier studies, a model of motivational messages is developed. This model is implemented in JAVA, providing a practical framework for generating motivational messages that are tailored on each of the communication aspects. Examples of easy to understand tailoring rules are given, that can be used to determine a suitable message intention and content given a user in some known context. The framework allows for a clear, sequential way of generating messages by implementing clear, simple, isolated tailoring rules, that when combined tackle the complex issue of providing a suitable message to the user.

The final communication aspect is treated in the last major chapter of the Thesis. **Chapter 5** deals with the *representation* of coaching, focusing on novel ways of representing the coaching to the users. The chapter explores the concept of representation through examples of the development and evaluation of two prototype coaching applications. A first prototype describes the use coaching mediated by multiple devices, where a virtual coach *migrates* with the user across devices and interfaces. The second prototype focuses on the use of an embodied agent that uses speech and facial animation to represent a real *virtual coach* concept. Evaluations show the potential merits of a multi-device system, but the early stage of the prototype's development does not allow us to draw any conclusions on its effectiveness at coaching. For the embodied agent prototype, evaluations show that users prefer a simple, text-based representation over the embodied agent, pointing out in this case a mismatch in complexity between representation and content.

Overall, the Thesis provides structure to the field of tailoring in the context of physical activity coaching. It is commonly believed that tailoring is a promising way of increasing the effectiveness of healthcare interventions, and many examples of applications that aim to apply tailoring in some shape or form exist. But so far, there is little consensus in the field of tailoring, either in regard to terminology or methodology. This Thesis aims to provide this structure and provides examples on how to navigate within this structure to implement effective tailoring mechanics in the context of an existing activity coaching application.

Samenvatting

Het gebrek aan fysieke activiteit in de algemene bevolking wordt erkend als een belanglangrijk maatschappelijk probleem. Voor ogenschijnlijk gezonde volwassenen, leidt inactiviteit tot overgewicht, en een verhoogd risico van talrijke chronische en acute ziekten zoals hartziekten, Type 2 diabetes, borst- en darmkanker. Een van de vele aanbebevelingen omtrendt meer bewegen is dat van de *American College of Sports Medicine*, die adviseert dat de meerderheid van de volwassenen ten minste dertig minuten per dag matig intensieve cardio-respiratoire oefeningen uitvoerd. In de afgelopen eeuw, is de menselijke levensduur jaarlijks gestegen met gemiddeld 0,24 jaar in de Verenigde Staten, Europa en Japan. Deze stijging van de levensverwachting betekent dat meer en meer mensen een groot deel van hun leven doorbrengen op een oudere leeftijd, en moeten omgaan met chronische ziekten zoals chronische obstructieve longziekte (COPD) of chronische rugpijn. De behandeling van deze chronische ziekten omvat vaak een programma van dagelijkse fysieke activiteit.

Dit proefschrift gaat over de bevordering van dagelijkse lichamelijke activiteit — wandelen, fietsen, tuinieren of huishoudelijk werk — alle activiteiten die regelmatig worden uitgevoerd in het dagelijks leven waarbij het lichaam moet worden bewogen. In de afgelopen jaren hebben onderzoekers en beleidsmakers zich gericht op technologie als een instrument om coaching op de lichamelijke activiteit te leveren. Technologie in de vorm van websites of smartphone toepassingen kan worden gebruikt om veel mensen te bereiken, tegen lage kosten, met het oog op zelfzorg voor patiënten (eHealth) of technologie gemedieerde professionele zorg in de thuisomgeving (Telemedicine). Een belangrijke factor in de populariteit van deze technologische coaching is de opkomst van goedkope, geminiaturiseerde draadloze sensoren die het mogelijk maken de lichamelijke activiteiten van een persoon gedurende de dag te meten. Een tweede belangrijke factor is de alomtegenwoordigheid van smartphones die deze activiteiten data kan verwerken en gebruiksvriendelijke feedback en coaching op elke plaats en elk moment kan geven. De combinatie van deze technologieën hebben er voor gezorgd dat effectieve activiteiten coaching instrumenten zijn ontwikkeld en nu beschikbaar zijn op de markt voor een redelijke prijs.

Hoewel de populariteit van applicaties die dagelijkse activiteiten meten en daar op coachen duidelijk is uit zowel de onderzoekswereld als commerciële gebieden, is de doeltreffendheid van de toegepaste coachings strategiën verre van perfect. Het wordt algemeen aangenomen dat coaching strategieën kunnen worden verbeterd door te cateren aan de behoeften en voorkeuren van individuele gebruikers, een onderzoeksgebied bekend als *tailoring* of *personalisatie*. Wat *tailoring* is, en hoe het op het gebied van real-time coaching op dagelijkse fysieke activiteit kan worden toegepast, zijn enkele van de thema's in dit proefschrift.

In **Hoofdstuk 1** wordt een algemene motivatie voor het werk in dit proefschrift gegeven. De problemen van de gezondheidszorg voor de vergrijzende bevolking, en de opkomst van het gebied van de fysieke activiteiten monitoring worden nader toegelicht. De focus van dit proefschrift op hoog technologische oplossingen — real-time coaching methoden — wordt uitgelegd, evenals de noodzaak om de doeltreffendheid ervan te verbeteren door middel van tailoring. De algemene inleiding eindigt met het stellen van de vier grote onderzoeksvragen die worden behandeld in de vier grote hoofdstukken van dit proefschrift.

Hoofdstuk 2 geeft het theoretische kader van het proefschrift door middel van een uitgebreide literatuurstudie over het onderwerp van op maat gemaakte real-time activiteiten coaching systemen. Het hoofdstuk onderzoekt twaalf verschillende toepassingen die elk op verschillende manieren hun coaching aanpak afstellen. Uit de analyse van deze toepassingen is een model van *tailoring* ontwikkeld, die een totaal van zeven verschillende tailoring concepten beschrijft: *feedback*, *inter-human interaction*, *adaptation*, *goal-setting*, *user targeting*, *context awareness* en *self-learning*. Verder is een definitie van een *communicatie in-*

stantie gegeven, bevattende aspecten van *timing intentie*, *inhoud* en *representatie*. Voor elke communicatie instantie met de gebruiker wordt getoond hoe elk van deze aspecten op verschillende manieren kunnen worden aangepast (*ge-tailored*).

Elk van de volgende hoofdstukken geven een focus op basis van de gedefinieerde communicatie-aspecten. In **Hoofdstuk 3** ligt de focus op het afstemmen van de timing van de communicatie met het oog op de naleving van de individuele motiverende berichten. In dit hoofdstuk wordt de motivatie, het ontwerp, de ontwikkeling en evaluatie beschreven van een systeem dat kan bepalen wanneer het optimale moment is voor het verstrekken van een motiverende *hint* voor de gebruiker. Op basis van de gebruiker en zijn huidige context, maakt het systeem gebruik van machine learning om momenten te classificeren als opportuun voor het initiëren van een bericht, en leert van de successen en mislukkingen. In een studie met 10 COPD patiënten is aangetoond dat deze manier van afstemmen van de timing aan individuele gebruikers helpt bij het verhogen de naleving van individuele berichten.

In **Hoofdstuk 4**, wordt de volgende stap naar een volledig geautomatiseerde coach genomen, door te focussen op de vraag *wat* te zeggen tegen de gebruiker. Gebaseerd op een analyse van de literatuur en een corpus van voorbeeld berichten gebruikt in eerdere studies, is een model van motiverende berichten ontwikkeld. Dit model is geïmplementeerd in JAVA, en bied een praktisch kader voor het genereren van motiverende berichten die zijn afgestemd op elk van de communicatie aspecten. Er worden voorbeelden gegeven van gemakkelijk te begrijpen tailoring regels, die kunnen worden gebruikt om een geschikte boodschap te bepalen in termen van intentie en inhoud voor een individuele gebruiker in een bepaalde context. Het kader zorgt voor een heldere, sequentiële manier van genereren van berichten door het implementeren van duidelijke, eenvoudige, geïsoleerde tailoring regels, die samen het hoofd bieden aan de complexe kwestie van het verstrekken van een passende boodschap aan de gebruiker.

Het laatste communicatie aspect wordt behandeld in het laatste grote hoofdstuk van het proefschrift. **Hoofdstuk 5** behandeld de *representatie* van coaching, en richt zich op nieuwe manieren om die coaching naar de gebruikers te brengen. Het hoofdstuk verkent het concept van representatie door middel van twee voorbeeld applicaties waarbij de ontwikkeling en evaluatie wordt besproken. Het eerste prototype behandeld het gebruik van coaching verspreid over meerdere apparaten, waarbij een virtuele coach meereist met de gebruiker over verschillende apparaten en interfaces. Het tweede prototype richt zich op de ontwikkeling van een *embodied agent* waarbij natuurlijke spraak en mimiek wordt gebruikt om een ware *virtuele coach* te vormen. Uit evaluaties blijkt dat er potentiële voordelen zijn van een systeem dat verspreid over meerdere apparaten werkt, maar het vroege stadium van de ontwikkeling van het prototype laat ons niet toe om conclusies te trekken over de specifieke effectiviteit van de coaching. Voor het embodied agent prototype blijkt uit evaluaties dat de gebruikers de voorkeur geven aan een eenvoudige, op tekst gebaseerde representatie, wat er in dit geval op wijst dat de complexiteit van inhoud en representatie niet op elkaar aansluiten.

Tot slotte brengt dit proefschrift structuur aan het gebied van tailoring in de context van het coachen op fysieke activiteit. Het wordt algemeen aangenomen dat tailoring een veelbelovende manier van het verhogen van de effectiviteit van interventies in de gezondheidszorg is, en er bestaan vele voorbeeld toepassingen die er op gericht zijn om tailoring op een of andere manier toe te passen. Maar tot nu toe is er weinig consensus in het vakgebied van tailoring, noch betreft de terminologie, noch de methodologie. Dit proefschrift heeft als doel om deze structuur te bieden en geeft voorbeelden van hoe binnen deze structuur te navigeren en te komen tot effectieve tailoring mechanismen in de context van een bestaande activiteiten coaching applicatie.

Acknowledgements

The work presented in this Thesis could never have been completed without, obviously a lot of help and support from a number of people.

My loving parents, Marion and Rieks, have since the beginning played a very instrumental role in the process leading up to the work here before you. I will refrain from writing a sentimental essay that may leave the reader under the impression that my last days on earth are drawing near, but it should be clear that I am very grateful for their support and care. Since I have attended the defense of my father Rieks he has been my major source of inspiration to becoming a scientist!

Clearly I am thankful to Hermie Hermens for this scientific opportunity. As a promotor and guide you have always given me the freedom to work on a very broad number of topics, and finally steering me towards finishing this final report on tailoring, healthcare, physical activity, and more! Val Jones, for all the interesting discussions, detailed commentary on all the writing, and all the many theoretical and philosophical discussions, a big thank you! I hope you are both as pleased as I am that the PhD chapter has come to an end!

Certain University policies and/or scientific traditions dictate that the work written in this book should be printed to paper. So, in order to limit my impact on deforestation, I shall limit the spatial requirements of the remainder of this acknowledgements section by switching to a nineties rock album format.

Many thanks to: Miriam Cabrita for discussions, care and support. All my colleagues at RRD for an enjoyable work environment, open for discussions and critique. All my co-authors for pleasant collaboration on various topics. All the people at StackExchange who take the time to write comprehensive answers, and working examples for all kinds of LaTeXand JAVA related issues. The developers and contributors of various awesome, free software packages: NotePad++, Eclipse, MiKTeX, TeXnicCenter, Mendeley, Inkscape, GIMP, Sumatra PDF, and DavMail. My friends, for providing the necessary distractions throughout all these years. Het Bolwerk, for providing the necessary distractions throughout all these years. Project partners in the various scientific national and European projects in which I was involved (with a special mention for Inertia Technology and the rest of the ISActive crew). The committee members for taking the time to read and evaluate this Thesis. My coffee cup for the steady supply of caffeine. Google Translate for helping out with the Dutch summary. Shampoo. Rieks op den Akker for the cover illustration, and Bob Dickhoff for the cover design. And everyone else I forgot...

Curriculum vitae

Harm op den Akker was born in Enschede, the Netherlands on the 20^{th} of February, 1985. In 2002 he started his Bachelor of Science education in Computer Science at the University of Twente in the Netherlands. In 2006 he received his Bachelor of Science diploma — with a final Thesis on '*Question Answering as Meeting Browser Interface*'. Immediately after, he continued his education at the University of Twente in the Master's program of Human Media Interaction, following courses in a.o. Machine Learning, Speech and Language Processing, and Conversational Agents. During the Master's program, he enrolled in a 3-Month Traineeship program, financed by the European FP6 project AMIDA, at the Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) in Saarbrücken, Germany. There he learned the LaTeX special character encoding for ü, but mainly worked on Automatic Dialogue Act Segmentation using Machine Classifiers, published in 2008. On the 20^{th} of March 2009, Harm finished his Master of Science degree with a Master Thesis titled '*On Addressee Prediction for Remote Hybrid Meeting Settings or how to use Multiple Modalities in Predicting whether or not you are being addressed in a live Hybrid Meeting Environment*', which was awarded with a score of 9/10.

In April 2009, he joined Roessingh Research and Development, in Enschede, the Netherlands, as a PhD researcher in the field of Telemedicine. He worked on a number of European and National projects, including IS-ACTIVE (AAL), AlWEN (PointOne), CareBOX (SBIR), Senior (Pidon), COPDdotCOM (ZonMW), and THeCS (COMMIT). Harm was focused on the development of intelligent coaching mechanisms, the results of which are presented in this Thesis.

Author publications

2014

Harm op den Akker, Valerie M. Jones, and Hermie J. Hermens, "Tailoring real-time physical activity coaching systems: a literature survey and model", in User Modeling and User-Adapted Interaction (available online), 2014.

Sebastiaan de Hoogh, Berry Schoenmakers, Ping Chen and **Harm op den Akker**, "Practical Secure Decision Tree Learning in a Teletreatment Application", in Proceedings of the 18th International Conference on Financial Cryptography and Data Security, Bridgetown, Barbados, March 2014.

Miriam Cabrita, **Harm op den Akker**, Reinoud Achterkamp, Hermie Hermens and Miriam Vollenbroek-Hutten, "Automated Personalized Goal-Setting in an Activity Coaching Application", in Proceedings of the 3rd International Conference on Sensor Networks (SensorNets2014), Lisbon, Portugal, pages 389-396, January 2014.

2013

Monique Tabak, **Harm op den Akker** and Hermie J. Hermens, "Motivational cues as real-time feedback for changing daily activity behavior of patients with COPD", in Patient Education and Counseling (available online), 2013.

Linda Vos-Maneschijn, Monique Tabak, **Harm op den Akker** and Hermie J. Hermens "Activity coach motivates an active lifestyle [Article in Dutch:

ActiviteitenCoach motiveert tot actieve levensstijl]", in FysioPraxis, December 2013.

Linda Vos-Maneschijn, Monique Tabak, **Harm op den Akker** and Hermie J. Hermens "Using an ambulant activity coach to motivate patients with COPD in having an active lifestyle [Article in Dutch: Gebruik van een ambulante activiteitencoach om mensen met COPD te motiveren tot een actieve levensstijl]", in Fysiotherapie en Ouderenzorg, Vol. 27(4), 2013.

Reinoud Achterkamp, Miriam Cabrita, **Harm op den Akker**, Hermie J. Hermens and Miriam M.R. Vollenbroek-Hutten, "Promoting a Healthy Lifestyle: Towards an Improved Personalized Feedback Approach", in Proceedings of the 15th IEEE International Conference on e-Health Networking, Applications and Services (Healthcom2013), Lisbon, Portugal, pages 677-679, October 2013.

Eleftheria Makri, Maarten H. Everts, Sebastiaan de Hoogh, Andreas Peter, **Harm op den Akker**, Pieter H. Hartel and Willem Jonker, "Privacy-Preserving Verification of Clinical Research", in Workshop Program of the Western European Workshop on Research in Cryptology (WEWoRC2013), Karlsruhe, Germany, pages 43-46, July 2013.

Harm op den Akker, Randy Klaassen, Rieks op den Akker, Valerie M. Jones and Hermie J. Hermens, "Opportunities for Smart & Tailored Activity Coaching", in Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems (CBMS2013), Porto, Portugal, pages 546-547, June 2013.

Randy Klaassen, Jordi Hendrix, Dennis Reidsma, Rieks op den Akker, Betsy van Dijk and **Harm op den Akker**, "Elckerlyc Goes Mobile - Enabling Natural Interaction in Mobile User Interfaces" in International Journal on Advances in Telecommunications, Vol. 6 (1 & 2), pages 45-56, 2013.

Randy Klaassen, Rieks op den Akker and **Harm op den Akker**, "Feedback Presentation for Mobile Personalised Digital Physical Activity Coaching Platforms", in Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA2013), Rhodes Island, Greece, May 2013.

Monique Tabak, **Harm op den Akker** and Hermie J. Hermens, "Time-Related Feedback Messages for Changing Activity Behaviour of Patients with COPD", in Proceedings of the 4th Dutch Bio-Medical Engineering Conference (BME2013), Egmond aan Zee, the Netherlands, pages 161-161, January 2013.

2012

Harm op den Akker, Valerie M. Jones, Laura S. Moualed and Hermie J. Hermens, "A Context-Aware Adaptive Feedback Agent for Activity Monitoring and Coaching", in Proceedings of the 8th International Conference on Methods and Techniques in Behavioral Research (Measuring Behavior 2008), Utrecht, the Netherlands, pages 192-132, August 2012.

Harm op den Akker, Monique Tabak, Mihai Marin-Perianu, Rianne Huis in't Veld, Valerie M. Jones, Dennis Hofs, Thijs M. Tönis, Boris W. van Schooten, Miriam M.R. Vollenbroek-Hutten and Hermie J.Hermens, "Development and Evaluation of a Sensor-Based System for Remote Monitoring and Treatment of Chronic Diseases - the Continuous Care & Coaching Platform", in Proceedings of the 6th International Symposium on eHealth Services and Technologies (EHST2012), Geneva, Switzerland, July 2012.

2011

Harm op den Akker, Laura S. Moualed, Valerie M. Jones and Hermie J. Hermens, "A Self-Learning Personalized Feedback Agent for Motivating Physical Activity", in Proceedings of ISABEL 2011: the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies, Barcelona, Spain, October 2011 (invited).

Thijs M. Tönis, **Harm op den Akker**, Simone T. Boerema, Freek van Polen and Hermie J. Hermens, "Exploring Patterns of Activities of Daily Living in the Home Environment", (abstract) for the 3rd Workshop on the pervasive application of wireless technologies (W3), Enschede, The Netherlands, September 2011.

Wilko Wieringa, **Harm op den Akker**, Valerie M. Jones, Rieks op den Akker and Hermie J. Hermens, "Ontology-Based Generation of Dynamic Feedback on Physical Activity", in Proceedings of Artificial Intelligence in Medicine, 13th International Conference on Artificial Intelligence in Medicine (AIME 11), Bled, Slovenia, July 2011.

Val Jones, Ricardo Batista, Richard Bults, **Harm op Den Akker**, Ing Widya, Hermie Hermens, Rianna Huis in't Veld, Thijs Tönis, Miriam Vollenbroek-Hutten. "Interpreting streaming biosignals: in search of best approaches to augmenting mobile health monitoring with machine learning for adaptive clinical decision support". In Learning from Medical Data Streams Workshop, July 6th, 2011.

Harm op den Akker, Hermie Hermens and Val Jones, "A Context-Aware Adaptive Feedback System for Activity Monitoring", in Book of Abstracts of ICAMPAM 2011: the 2nd International Conference on Ambulatory Monitoring of Physical Activity and Movement, Glasgow, United Kingdom, May 2011 (nominated for best student oral presentation).

2010

Harm op den Akker, Val Jones and Hermie Hermens, "Predicting Feedback Compliance in a Teletreatment Application", in Proceedings of ISABEL 2010: the 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies, Rome, Italy, November 2010.

2009

Harm op den Akker and Rieks op den Akker, "Are You Being Addressed? - real-time addressee detection to support remote participants in hybrid meetings", in Proceedings of SIGDIAL 2009: the 10th Annual Meeting of the Special Interest Group in Discourse and Dialogue, pages 21-28, Queen Mary University of London, September 2009. ISBN: 978-1-932432-64-0.

Harm op den Akker, Val Jones and Hermie Hermens, "Personalized Feedback based on Automatic Activity Recognition from Mixed-Source Raw Sensor

Data", in Proceedings of the Intelligent Data Analysis in bioMedicine and Pharmacology (IDAMAP) Workshop, AIME2009, pages 65-66, July 19, 2009.

2008

Rieks op den Akker, Dennis Hofs, Hendri Hondorp, **Harm op den Akker**, Job Zwiers and Anton Nijholt, "Supporting Engagement and Floor Control in Hybrid Meetings", in Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions, COST Action 2102 International Conference, Prague, Czech Republic, pages 276-290, October 15-18, 2008. ISBN: 978-3-642-03320-9.

Harm op den Akker and Christian Schulz, "Exploring Features and Classifiers for Dialogue Act Segmentation", in Lecture Notes in Computer Science - Proceedings of the 5th International Workshop on Machine Learning for Multimodal Interaction, pages 196-207, September 8-10, 2008. ISBN: 978-3-540-85852-2.

Progress range

The following publications have been published in the Progress range by Roessingh Research and Development, Enschede, the Netherlands. Copies can be ordered, when available, via info@rrd.nl.

1. J.W.G.A. Pot, H. Boer, W.H. van Harten, H.J. Hermens, E.R. Seydel, Comprehensive Need-Assessment. Ontwikkeling van een meetinstrument voor zorgbehoeften en kwaliteitsbeoordeling door patiënten, Roessingh Research & Development, the Netherlands, September 1994, ISBN 90-25452-01-2.
2. N.G. A. van Leerdam, H.J. Hermens, Revalidatietechnologie in Euregio, Roessingh Research & Development, the Netherlands, July 1995, ISBN 90-75452-02-0.
3. L. Duda. L.O. van Noort, S. Röseler, B.O.L. Greitemann, W.H. van Harten, N.S. Klazinga, Rehabilitation in Germany and the Netherlands, A comparison of two rehabilitation-systems, Roessingh Research & Development, the Netherlands, August 1995, ISBN 90-75452-03-9.
4. Hermie J. Hermens, Anand V. Nenen, Gerald Zilvold, Electrophysiological Kinesiology, Proceedings of the 11 th congress of the International Society of Electrophysiologie and Kinesiology in Enschede, TheNetherlands 1996, Roessingh Research & Development, the Netherlands, October 1996, ISBN 90-75452-04-7.
5. Wim H. van Harten, Bouwen aan een kwaliteitssysteem in de revalidatiezorg. Een poging tot constructieve technology assessment van een kwaliteitssysteem in een gezondheidszorginstelling, Roessingh Research & Development, the Netherlands, december 1997, ISBN 90-75452-07-1.
6. Gert Baardman, Maarten J. IJzerman, Design and evaluation of a hybrid orthosis for people with paraplegia, Roessingh Research & Development, the Netherlands, November 1997, ISBN 90-75452-08-X.
7. Miriam M.R. Hutten, Lumbar Dynamometry: A useful method for assessment of patients with chronic low back pain?, Roessingh Research & Development, the Netherlands, November 1999, ISBN 90-75452-13-6.
8. Drs. A. van der Salm, Prof. Dr. W.H. van Harten, Dr. C.G. B. Maathuis, Ketenkwaliteit

Cerebrale Parese Zorg. Een beschrijving van de cerebrale parese zorg en mogelijke verbeteringen hierin, Roessingh Research & Development, the Netherlands, april 2001, ISBN 90-75452-19-5.

9. Marc J. Nederhand, Muscle activation patterns in post traumatic neck pain, Roessingh Research & Development, the Netherlands, March 2003, ISBN 90-75452-27-6.
10. Michiel J.A. Jannink, Usabillity of custom-made orthopaedic shoes in patients with degenerative disorders of the foot, Roessingh Research & Development, the Netherlands, September 2004, ISBN 90-75452-28-4.
11. Mariëtte Blokhorst, State-dependent factors and attention in whiplash associated disorder, Roessingh Research & Development, the Netherlands, January 2005, ISBN 90-365-2111-4.
12. Jaap H. Buurke, Walking after stroke co-ordination patterns & functional recovery, Roessingh Research & Development, the Netherlands, February 2005, ISBN 90-365-2140-8.
13. Arjan van der Salm, Spasticity reduction using electrical stimulation in the lower limb of spinal cord injury patients, Roessingh Research & Development, the Netherlands, October 2005, ISBN 90-365-2253-6.
14. Govert J. Snoek, Patient preferences for reconstructive interventions of the upper limb in tetraplegia, Roessingh Research & Development, the Netherlands, December 2005, ISBN 90-365-2255-2.
15. Joke de Kroon, Therapeutic electrical stimulation of the upper extremity in stroke, Roessingh Research & Development, the Netherlands, December 2005, ISBN 90-365-2269-2.
16. Van Dijk H. Motor skill learning, age and augmented feedback. Roessingh Research and Development, the Netherlands, March 2006, ISBN 90-365-2302-9.
17. Mes CAJ. Improving non-optimal results in chronic pain treatment. Roessingh Research and Development, the Netherlands, January 2007, ISBN 90-365-2435-0.
18. Voerman GE. Musculoskeletal neck-shoulder pain: a new ambulant myofeedback intervention approach. Roessingh Research and Development, The Netherlands, March 2007, ISBN 90-365-2460-1
19. Kallenberg LAC. Multi-channel array EMG in chronic neck-shoulder pain. Roessingh Research and Development, the Netherlands, March 2007, ISBN 90-365-2459-8
20. Huis in 't Veld MHA. Work-related neck-shoulder pain: The role of cognitive-behavioural factors and remotely supervised treatment. Roessingh Research and Development, The Netherlands, December 2007, ISBN 978-90-365-2584-8
21. Judith F.M. Fleuren. Assessment of Spasticity: From EMG to patients' perception. Roessingh Research and Development, The Netherlands, October 2009, ISBN 978-90-365-2869-6
22. Jasper Reenalda. Dynamic sitting to prevent pressure ulcers in spinal cord injured. Roessingh Research and Development, The Netherlands, October 2009, ISBN 978-90-365-2884-9
23. G.B. (Gerdienke) Prange. Rehabilitation Robotics: Stimulating restoration of arm function after stroke. Roessingh Research and Development, The Netherlands, October 2009, ISBN 978-90-365-2901-3
24. Vos-van der Hulst M. Prognostic factors and underlying mechanisms in chronic low back pain. Roessingh Research and Development, The Netherlands, November 2009, ISBN 978-90-365-2881-8

25. Anke I.R. Kottink-Hutten. Assessment of a two-channel implantable peroneal nerve stimulator post-stroke. Roessingh Research and Development, The Netherlands, February 2010, ISBN: 978-90-365-2959-4
26. Marit G.H. van Weering. Towards a new treatment for chronic low back pain patients. Roessingh Research and Development, The Netherlands, May 2011, ISBN: 978-90-365-3180-1
27. Jitske Gulmans. Crossing Boundaries: Improving Communication in cerebral palsy care. Roessingh Research and Development, The Netherlands, February 2012, ISBN: 978-90-365-3305-8
28. Birgit Molier. Influence of augmented feedback on learning upper extremity tasks after stroke. Roessingh Research and Development, The Netherlands, March 2012, ISBN: 978-90-365-3296-9
29. Rosemary Dubbeldam. Towards a better understanding of foot and ankle kinematics in rheumatoid arthritis. Roessingh Research and Development, The Netherlands, October 2012, ISBN: 978-90-365-3407-9
30. Richard Evering. Ambulatory feeback at daily physical activity patterns. Roessingh Research and Development, The Netherlands, April 2013, ISBN: 978-90-365-3512-0
31. Malhotra S. Does spasticity interfere with functional recovery after stroke? Roessingh Research and Development, The Netherlands, November 2013, ISBN: 978-90-365-3567-0
32. Monique Tabak. Telemedicine for patients with COPD. New treatment approaches to improve daily activity behaviour. Roessingh Research and Development, The Netherlands, February 2014, ISBN: 978-94-6108-590-0
33. Hester Trompetter. ACT with pain. Measurement, efficacy and mechanisms of Acceptance & Commitment Therapy. Roessingh Research and Development, The Netherlands, September 2014, ISBN: 978-90-365-3708-7

Roessingh
Research and Development
34

Progress in rehabilitation science

ISBN 978-90-365-3762-9