THE FUTURE OF SELF-CONTROL:

DISTRIBUTED MOTIVATION AND COMPUTER-MEDIATED EXTROSPECTION

Lars Hall, David de Léon & Petter Johansson
Lund University Cognitive Science
Kungshuset, Lundagård
222 22 Lund
E-mail: {lars.hall, david.de_leon, petter.johansson}@lucs.lu.se

Abstract: In this article we discuss the social implications of context-aware computing technology for the perennial human problem of *self-control*. We present a basic model of the domain of self-control, and provide a range of suggestions for how modern sensor and computing technology might be of use in scaffolding and augmenting our self-control abilities. The model consists of two core concepts. The first we call *Computer-Mediated Extrospection*, which builds upon the familiar idea of self-observation or self-monitoring, and concerns itself with the crucial need for accumulation and explication of self-knowledge in any rational person-centered decision process. The second concept is *Distributed Motivation*, which we see as a natural extension of the idea of precommitment and self-binding that is often discussed in the self-control literature. The article ends with a discussion of issues of flexibility, and ethical concerns about privacy and persuasion in possible context-aware applications for self-control.

Keywords: Self-control, precommitment, distributed motivation, computer-mediated extrospection, persuasive computing, affective computing, contextual awareness, ubiquitous computing, pervasive computing.

1. INTRODUCTION

The ubiquitous vision is one in which computers take an increasing part in our everyday activities, in ways that mesh naturally with how people think, act and communicate. We are excited by this vision, but feel that the full possibilities offered have yet to be explored. Work within ubiquitous computing and context awareness has made us increasingly familiar with computers that mediate our interactions with the world, *but what about computers that mediate our interactions with ourselves?* We believe that computers can be made more powerful by letting them gain information about the user, but in a similar manner we also believe users can be made smarter and more powerful by letting them gain additional knowledge about themselves. To this end we will here propose some ways in which sensor and computing technology might be used for purposes of *self-control*. This is an avenue that, surprisingly, has remained largely unexplored.

Recently, HCI researchers have shown a growing interest in the motivational role that computers might serve in human conduct – i.e. what has become known as *persuasive computing* (Fogg, 2003). However, this field is still in its infancy, and very few explicit connections to theoretical or empirical research on self-control have been made (e.g. to important work like Rachlin, 2000; Elster, 2000; or Ainslie, 2001).

The lack of qualified research on the possible use of information technology to alleviate problems of self-control is, we feel, a very unfortunate state of affairs. Issues of self-control are extremely pervasive in modern societies. Take for example the use of tobacco. According to the latest *World Health Report* (2001) our planet harbors more than 1.2 billion smokers worldwide and tobacco accounts for well over three million annual deaths. Costs are more difficult to calculate, but a recent World Bank report on the economics of tobacco control estimates that in high-income countries smoking-related health care account for 6–15% of all annual health care cost (World Bank, 1999). Similar figures can be found in relation to regulation of dietary intake. As measured by the standardized Body Mass Index we now have roughly 1.1 billion overweight people in the world (Gardner & Halweil, 2000). In the US alone an estimated 300 000 people die each year of causes related to obesity (Mokdad et al., 2001). However, problems of self-control and self-regulation are not only operative in such salient and life-threatening

Lund University Cognitive Studies - LUCS 95 ISSN 1101 - 8453 © 2002 by the authors

domains as craving and addiction, but also in the minute workings of everyday plans, choices and actions. Ameliorative action is as pertinent to the dreadful experience of withdrawal from heroine, as it is to innocuously hitting the snooze-button on the alarm clock, and missing the first morning bus to school (Rachlin, 2000; Ainslie, 2001).

The purpose of this article is to present a succinct model of the domain of self-control that maps both the degree of severity and the great variety of self-control problems, as well as possible remedial actions using modern sensor and computing technology. The model consists of two core concepts, or tools, that we believe may serve an important role in elucidating the problem of self-control from a ubiquitous computing perspective. First, we introduce the concept of *Computer-Mediated Extrospection*, which builds on and expands the familiar idea of *self-observation* or *self-monitoring*, and concerns itself with the crucial need for accumulation and explication of self-knowledge in any rational person-centered decision process. Secondly, we present the idea of *Distributed Motivation,* as a natural extension of previous discussions of *precommitment* and *self-binding* in the self-control literature.

Issues of context awareness occupy center stage in the field of ubiquitous computing and human–computer interaction (Dey, Abowd & Salber, 2001). The most relevant aspects of context are also generally agreed upon. For instance, according to Dey, Abowd and Salber (2001): "Context is typically the location, identity, and state of people, groups and computational and physical objects," where state refers to "physical, social, emotional, or informational state". In practice however, modeling context has been mostly confined to information about identity and location, and far less attention has been paid to the psychological states of people. For purposes of research and development this has been a pragmatically sound strategy, and more recently it has also been amply demonstrated how pertinent contextual constructs can be leveraged from the combination of simple environmental measures (see e.g. Gellersen, Schmidt & Beigl, 2002). However, if the computer is to become a tool for augmenting interactions with ourselves these kinds of measures must be combined with more intimate, psychophysiological measures. In the kinds of applications we discuss in the present paper the affective and cognitive states of users, and the emotional context of interaction, are essential and inescapable aspects of context.

In presenting our model, we draw upon existing research in ubiquitous computing and context awareness (and from conceptual neighbors like wearable computing, telemedicine, affective computing, and the aforementioned field of persuasive computing), but to make our points clearly we also include references to future scenarios and hypothetical cases. It is our hope that the model and our discussion will provide a principled and useful way for designers of human–computer interfaces and context-aware systems to approach the domain of self-control, as well as to provoke further debate on the possible role of computing technology in matters of human motivation.

The outline of the article is as follows. First we present an overview of the problem of self-control, and then, in consecutive sections, we introduce and discuss our two conceptual tools. The article ends with a discussion of different self-control scenarios drawn from our model, and some suggestion for how modern sensor and computing technology might be used to alleviate problems of self-control.

2. SELF-CONTROL

In its simplest form, the problem of self-control consists of the fact that we tend to choose smaller, sooner rewards rather than larger later rewards *despite* knowing that this is against our best interest. At a descriptive level such situations show a characteristic profile. At T1, a safe distance from the reward, we decide that we prefer the greater reward to be delivered at T3, to the lesser reward delivered at T2. However, at an intermediate time right before T2 we succumb to the imminent lesser reward, which is then followed by regret and lament at T3. Obviously, not all self-control problems are so straightforward, but clearly delineated conflicts between smaller sooner and larger later rewards (what Rachlin, 2000, calls *simple ambivalence*) form the core of the issue of self-control. Importantly, the problem of self-control lies not simply in the act of impulsively choosing an immediate and "lesser" reward, but in doing so against ones own recognized best interest. Examples of this would include failure to follow through on decisions to start exercising, or quit smoking, or a constant tendency to put off writing important assignments at school.

But the problem of self-control is not just a problem manifested in the behavior of certain "weak-willed" individuals: it is a basic, universal and hardwired feature of reward-anticipation in the human brain. A great deal of research has been conducted into choice-behavior in relation to different temporal distributions of reward, both in animals and humans (see for example Mazur, 2001; Frederick, Loewenstein & O'Donoghue, 2003). The main result that has emerged from this research is a general mathematical function that concisely expresses the diminution of the motivational force of a reward in relation to the length of delay. This is called the *temporal discount function* (Monterosso & Ainslie, 1999). All organisms prefer having immediate access to reward, rather than having to wait for it. The most important empirically derived property of the discount function is its hyperbolic shape (Ainslie, 2001; Monterosso & Ainslie, 1999; see also Read, 2003). What this means is that our subjective evaluation of the reward grows much faster when we are closer to the reward than when we are far from it, and that the mere passage of time can lead to sudden preference reversals.

However, for animals such preference reversals only represent a kind of *manufactured* irrationality. Choice behavior that produces preference reversals in the laboratory is consistent with models of adaptive foraging in the wild (roughly captured by the adage "a bird in the hand is worth more than ten in the woods"). The motivational systems of rats, pigeons, chimpanzees and other animals are naturally attuned to the reward contingencies of ecologically valid environments, and not to cleverly designed laboratory settings (Rolls, 1999). For humans, on the other hand, temporally induced preference reversals present a serious problem. Unlike most animals, we *constantly* live in a manufactured environment, and the distribution and supply of rewards we face bears little resemblance to the environment in which our species evolved (Duchaine, Cosmides, & Tooby, 2001; but see Kacelnik, 2003). Given our ability to plan ahead and form long-term goals, a natural conflict arises when ancestral systems of reward evaluation entice us with short-term immediate gains. What is revealed by the hyperbolic discounting curve is that failure to follow through on long term goals takes place *just because* we do not have an evolved capacity to wisely, disinterestedly and steadfastly select between short and long-term rewards.

2.1. Computer-Mediated Extrospection

What is it that people do, when they acquire, analyze and act upon self-relevant knowledge? According to folk wisdom, to arrive at such knowledge, people engage in a process of *introspection*, of looking inwards and inspecting the contents of their own minds. Even if this process often is believed to be both fallible and arduous, it is also believed to be more or less transparent to the person involved in the activity. If it is anything in the world that people know with certainty, it is what they themselves think, feel, believe and desire (Goldman, 1993). From this perspective, it would seem that a scheme of capturing and representing aspects of user-context, *for the supposed benefit of the users themselves*, would be of limited value. Such information, it seems, would at best be redundant and superfluous, and at worst a gross mischaracterization of the user's true state of mind.

On the other hand, common-sense psychology has always acknowledged an imperfect access and command over *some* aspects of our mental lives. Memory, for example, is a case at hand: it is common knowledge that processes of encoding and remembering often are fragile and sometimes inscrutable. The fact that memory is a fundamentally reconstructive process, often at risk of seriously distorting the past, also seems to be generally agreed upon (Buckner & Wheeler, 2001; Paller & Wagner, 2002). Here, then, it is obvious that context aware, memory-enhancing technology could provide a valuable service to users (Mann, 2001; Beigl, 2000). However, computing technology can do more than just emulate the old diary function, and does not have to rest content with capturing information that would have been available to the user if she only had been more attentive or vigilant. Technology can provide information about the state of the user that is *uniquely accessible* by such means. For example, functional magnetic resonance imaging (fMRI) measures of brain activation taken at the time of encoding of a certain material, can (in contrast to the people doing the actual encoding), accurately *predict* levels of recall for a period of several weeks afterwards (Schacter & Dodson, 2001). Similarly, fMRI activity-measures obtained concomitantly with an event can be used to separate out *true* from *false* memories about that event – to a degree not nearly approximated by the remembering agents themselves (Schacter & Dodson, 2001).

Once one starts questioning the scope of our introspective access, it soon becomes clear that it is even more circumscribed than what first appeared to be the case. Memory, for example, is much more than a simple process of routine encoding and retrieval: it is an inseparable component of reasoning and reflection, and deeply involved in our concurrent efforts to gain self-knowledge and regulate our behavior (LeDoux, 2002). As an illustration, take the case of phobia. For many phobias the subject is unambiguously and acutely aware of the fact that the specific fears they harbor (spiders, open spaces, heights, etc.) are irrational and unreasonable, while at the same time completely failing to act upon this belief (Öhman & Mineka, 2001). The reason for this being that the phobias have been laid down as memories by dedicated subcortical fear-learning mechanisms that are all but cognitively impenetrable (Medina et al., 2002). We simply cannot "look inwards" and divine or correct the workings of these brain systems without extensive training, external prompting, or deliberate relearning (as is practiced in cognitive-behavioral therapy and other similar techniques). But phobias are only an extreme case of the constant, day-to-day multilevel learning and responding that takes place in our lives. Using methods of implicit measurement (including everything from hulking basement-dwellers like fMRI scanners, to simple ambulatory sensing of galvanic skin responses) striking dissociations between subjective experience and cognitive/emotional activity have been established in a wide variety of domains and behaviors (e.g. see Gazzaniga, 2000, for a wealth of examples). Evidently, the process of introspection is powerless to survey and regulate a great and important part of our mental economy.

In fact, in our view, these examples demonstrate a general principle about the human cognitive architecture: implicit processing of one or other variety is the norm, not the exception (Rolls, 1999; Dehaene & Naccache, 2000). Even most forms of learning have strong elements of implicitness. The competencies we acquire tend to be anchored in the specific tissues that are modified by training. They are *embedded* competencies, in the sense that they are incapable of being transported readily to be brought to bear on other problems faced by the individual, or shared with other individuals. *It is knowledge in the system, but not yet knowledge to the system* (Clark & Thornton, 1997).

Importantly, this does not mean that we are powerless to acquire and act upon knowledge about ourselves. Implicit knowledge can be observed in its external, somatic and behavioral manifestations, and it can be subjected to educated, situational "probes" (or sometimes just plain trial and error), in order to generate significant patterns of reaction. We call this process *extrospection*, as we believe it makes for a salient and informative contrast to the traditional concept of *intro*-spection. In its basic form, extrospection involves the observation and extraction of regularities that represent the outward expression of implicit information processing in the brain. From such regularities (or sometimes even single instances) the extrospecting agent must infer likely causes and reasons for their occurrence. But, as we mentioned above, extrospection can also involve subtle "provocations" of specialist brain subsystems, in order to evoke noteworthy reactions (i.e. a sort of quasi-experimental approach to self-reactivity), indicating how external feedback-loops can be used to probe and direct our own brain-internal processes[1].

A currently much discussed example of this latter process is the use of peripheral emotional reactions to guide and constrain decision-making. Studies have shown how loss of peripheral emotional reactions (such as galvanic skin responses) can result in critically impaired decision-making on a variety of tasks (Bechara et al., 2002). The theory behind this being that if we loose extrospective access to the embodied wisdom of our bodily reactions (our so-called somatic markers), then we also loose some of our ability to make fast and appropriate responses to everyday choice-situations.

Still, as a process of inference, extrospection is subject to the same limitations and problems as any other form of reasoning. In the following sections we elaborate upon possible ways in which the process of extrospection can be augmented by the use of sensor and computing technology – what we call computer-mediated extrospection (CME). CME has many potential uses in the wider process of self-regulation, but here we focus on its particular application to problems of self-control.

2.1.1 Computer-Mediated Extrospection and Self-Control

The starting point for many discussions of self-control is the observation that people are often aware of their self-control problems, but seldom *optimally aware* of the way these problems are expressed in their behavior, or under what contingencies or in which situations they are most prone to lapses in control (what is called *partial naiveté* in behavioral economics). Most likely, this is due to a mix of biased self-perception, cognitive limitations, and lack of inferential activity (Frederick, Loewenstein & O'Donoghue, 2003). Here CME could serve an important role in correcting faulty self-perceptions. The types of systems we envisage show considerable overlap with initiatives in personal imaging and remembrance agents within wearable and ubiquitous computing (e.g. Mann, 2001; Rhodes, 2000; Singletary & Starner, 2001). However, CME would differentially emphasize the elucidation of information with specific relevance for self-knowledge and self-regulation (not just any task in which augmented memory could be employed). Within this domain, we see four rough categories of CME-tools.

1. Enhanced Perception. As a first measure of a CME-tool geared towards improving self-perception, the focus would be on capturing and representing valuable information in our immediate surrounding that we normally fail to register and/or encode, but which we generally believe ourselves to have at least some inkling of. While it may seem like the category of things we falsely believe ourselves to have seen, heard, felt, etc., ought to be very small, evidence suggests otherwise. As the phenomenon of change blindness (i.e. of not noticing potentially gross and remarkable changes to scenes or pictures under conditions of degraded low-level motion information) makes clear, the essence of vision is not as a *form of representation*, but rather a *mode of exploration* (O´Regan & Noe, 2001), and that few things in our near-self environment are registered and retained in any enduring detail (O´Regan & Noe, 2001). On the other hand, our avowed self-knowledge about such matters tends to assume the existence of a much more detailed and reliable impression, and this can lead to a wide variety of self-related misconceptions (Levin et al., 2000).

2. Macro Prediction. As a possible means of mitigating problems of self-perception CME could also be used for purposes of *macro prediction*, by finding subtle regularities in behavior over time and situations. Even if humans are obsessive, incessant and adept pattern-recognizers (whether we know it or not) we are ill suited to process data that is scattered over many different contexts and time-scales. The role of CME could be one of *personal data-mining* (Clarkson, 2002), to discover quirks of acting and responding that are well nigh invisible from the subjective perspective.

3. Self-Monitoring. It is also of great importance to apply CME to capture and represent information that we normally successfully access and monitor, but which we sometimes *momentarily* fail to survey. Studies have shown that while humans are quite capable at self-monitoring when given clear directives and timely external prompts, performance quickly deteriorates under natural conditions (Rachlin, 2000). (Compare not trying to scratch an itch under stern scrutiny in the doctor's office, and not scratching it later while watching TV.) The degree of self-monitoring, in turn, greatly influences the nature of our self-control behavior. There is a big difference between smoking a cigarette

---

[1] We specifically want to stress that there can be no clear dividing line between seeing these strategies as unconsciously applying themselves when the situation calls for it, and us deliberately and consciously employing them in the service of a particular goal. As far as cognitive operations go, there is no systematic relation between the complexity of a process and whether it executed in a conscious manner or not (see Dehaene & Naccache, 2000).

that happens to be the 24th of the day, and being aware that one is about to light up the 24th cigarette for the day. The simple fact of providing accurate monitoring of self-control related context has been shown to markedly reduce the incidence of self-control lapses (Rachlin, 2000). The problem is of course that it is almost as difficult to stay constantly vigilant and attentive to such context as it is to control the behavior in the first place. This is an area where the use of context aware technology and CME would be of great use.

4. Micro Prediction. CME can also be applied in a more direct and intimate manner to measure and influence cognitive and emotional brain activity. A more complex (and potentially more powerful) form of CME would be to apply ubiquitous sensing and computing technology to explicate the relations between different levels of explanation of behavior. A standard approach in context-aware applications is to build up relevant context from a variety of simple features (or primitives). Combinations of contextual elements like time, location, position, etc. can be used to derive a specific *action*, or *activity* (Gellersen, Schmidt & Beigl, 2002). In explanations of human behavior the gold-standard of context-abstraction is the intentional level: the level at which we can determine what *purpose* an action has, what it *means*, and what the agents involved *intend*, *desire* and *believe*. There is ample evidence that the human ability to identify intentional states is built up from many semi-autonomous, interdependent processes (detection of self-propelled motion, eye-gaze, joint-attention, etc. see Malle, Moses and Baldwin, 2001). Used in concert, and applied similarly to both oneself and to others, these mechanisms secure the capture of relevant high-level *patterns* in human behavior, and give our folk-psychology great powers of explanation and prediction (Dennett, 1991). However, such patterns are still abstractions, and leave out much information that could be (and often is) critical to explanations of human behavior. Here, CME finds several different uses. Most importantly, CME could enable a user to perform various forms of *micro prediction* of her own behavior. For example, in the fMRI studies of memory encoding described earlier, the subjects involved did not *intend* to forget the material, or held some odd *beliefs* that made forgetting understandable, it was simply the case that the functional-level, brain-based explanation of the process, was *more powerful* than their own self-explanations. This type of prediction does not necessarily have to be based on “in-skull” measurement. As is commonly the case in human-factors studies of error-performance, attention-lapses, and similar micro behaviors, it could just as easily be based on surface psychophysiology, or even reaction-time performance (Kramer & Weber, 2000; Parson & Hartig, 2000). In a similar way we envision that CME can be used in the context of self-control to set up series of micro predictions of lapse-critical behavior in the presence of temptation cues, or specific contexts previously associated with relapse.

However, as important as the process of acquiring and processing self-relevant knowledge by CME, is the further use this knowledge is put to in processes of regulation and control. The possibilities of CME as a new *interface* for ourselves go far beyond simple feedback-control. Only detailed experimentation can determine what function – modulating, communicative, explanatory, metacognitive, rewarding, facilitating, distractive, evidential, etc. – that CME might play in any given ubiquitous and context-aware system. In the next section we discuss how the *output* from CME can play a crucial role in instigating and shaping wider processes of motivation and self-control. We also introduce *distributed motivation*, the second of the two general conceptual tools we believe to be important in engaging the problem of self-control.

2.2 Distributed Motivation

As has become evident from our discussion of the nature of the self-control dilemma, and the various means of attaining self-knowledge (whether by our natural senses, or by CME), there is no simple and patented solution to the problem of self-regulation and control. The interesting question is rather what we ordinary folks do when we decide to set out to pursue some lofty goal – to start to exercise on a regular basis, to finally write that film script, to become a less impulsive and irritable person – if we cannot just look inside our minds, exercise our “will,” and simply be done with it.

The answer, we believe, is that people cope as best they can with a heterogeneous collection of culturally evolved and personally discovered strategies, skills, tools, tricks and props. We write authoritative lists and schedules, we rely on push and pull from social companions and family members, we rehearse and mull and exhort ourselves with linguistic mantras or potent images of success, and we even set up ceremonial pseudo-contracts (trying in vain to be our own effective enforcing agencies). Often we put salient markers and tracks in the environment to remind us of, and hopefully guide us onto the chosen path, or create elaborate scenes with manifest ambiance designed to evoke the right mood or attitude (like listening to soundtracks of old Rocky movies before jogging around the block). We also frequently latch onto role models, seek out formal support groups, try to lock ourselves into wider institutional arrangements (join a very expensive tennis-club with all its affiliated activities), or even hire personal pep coaches. In short, we prod, nudge and twiddle with our fickle minds, and in general try to *distribute* our motivation out into stable social and artifactual structures in the world. Like Odysseus facing the Sirens we often know that we will find ourselves in conditions where we are likely to do something detrimental to our long-term goals, and like Odysseus tying himself to the mast we would often like to be able to *self-bind* or *precommit*, and avoid or resist such temptations.

While various extrospective processes provide the core input and overall shape of our proposed self-regulatory efforts (identifying needs, judging the effectiveness of potential measures, testing solutions to get crucial feedback, etc.) the general strategy of using stable features of the environment to scaffold the process of goal-attainment deserves a special mention. This is what we call *distributed motivation*. As such, distributed motivation is a subclass of the well-established theory of *distributed cognition* (Hutchins, 1995a). Distributed cognition deals with computational processes distributed among agents, artifacts and environments. It is a set of tools and methodologies that allow the researcher to look beyond simple "cognizant" agents, and shift the unit of analysis to wider computational structures (among which the human brain of course is an important part). Distributed motivation aims to achieve the same shift of emphasis in the realm of motivational problems as distributed cognition has done for problems of reasoning (Zhang, 1997), memory (Hutchins, 1995b), and collaboration (Hutchins, 1995a; Rogers & Ellis, 1994). We do not believe there is any principal difference between the "cold" cognitive phenomena normally studied, and the "hot" motivational and emotional processes that are our focus here.

The adoption of an explicit framework of distributed motivation will not only provide a platform in the search for potential remedial applications, but also, we believe, capture often overlooked aspects of how people actually go about trying to overcome problems of self-control (see Elster, 2000, for a similar sentiment). Primary among these aspects, and one of the most central features of our notion of distributed motivation, is the concept of *precommitment* or self-binding.

The tale of Odysseus and the Sirens is a standard illustration of this principle (Elster, 2000; for an in depth treatment, see Sally, 2000a,b). Odysseus, knowing the inevitable consequences of the siren song, orders himself to be tied to the mast (and plugs the ears of the oarsmen crew) thus arranging the environment in such a fashion as to allow him to sail by unharmed. Going back to our outline of the self-control problem, Odysseus suspects at time T1 that he will most likely experience a preference reversal at T2, and so he guarantees, by precommitment, that his original preference will not be violated, and receives the larger long-term reward at T3. What we would like to argue here is that the image of the clever Odysseus foiling the Sirens, might serve as a promising template for the design of modern remedies based on ubiquitous and context-aware technology. While people generally strive to approximate the Odyssean ideal in their daily self-regulation behavior they seldom manage to create conditions of precommitment stable enough to sustain them through complex and difficult problems. As sure as the fact that the majority of folk-strategies of self-control have been tried and tested in harsh conditions of cultural evolution, or over the full life span of incessantly extrospecting individuals, and that they embody considerable pragmatic wisdom, is also the fact that they fail miserably when looked at on a societal scale (e.g. the extreme pervasiveness of failures to self-regulate that we elaborated upon in the introduction).

2.2.1 Distributed Motivation and Ubiquitous Precommitment Technology.

The problem with most folk-strategies is of course that *they do not have enough binding power* (sadly the injunctions are often no stronger than the glue on the back of the post-it notes they are written on). For example, an often-told anecdote in the context of research on self-control is that of the young Afro-American man that made a "powerful" commitment to pay US$ 20 to the Ku Klux Klan every time he smoked a cigarette. In contrast to many other cases it is easy to understand the force this commitment *might* have on his behavior, but the fact still remains that once he has succumbed to the temptation, nothing really compels him to transfer money to the KKK. But if no such crucial deterrent for future behavior can be established, then why on earth should he adjust his behavior in relation to the commitment to begin with? Without going into philosophical niceties, it is easy to see that there is something deeply paradoxical about this kind of self-punishment. Indeed, if one really could exert the type of mental control that effectively *binds* oneself to pay the smoking fee to the KKK, then why not just simply bind oneself not to smoke in the first place?

The main weakness of the strategy employed is the lack of enforcement. The key to improving on the strategy is clearly to increase the binding power of the initial precommitment; in this case, ensuring that the "fine" for smoking is reliably incurred and that lapses are reliably detected. There are in fact several possible solutions, both to monitoring and enforcement. Of the more extreme variety are the agencies that offer round the clock surveillance of dieters and smokers. Although effective, there is a very understandable general resistance to these kinds of schemes. In addition, they are usually cumbersome, inflexible and costly (Ainslie, 1999, 2001; Rachlin, 2000). What is needed are solutions which do not compromise individual integrity, and were the cost of setting up and maintaining the scheme is in parity with the expected benefits. We believe that a ubiquitous infrastructure will be able to meet all of these demands. In the next section, we introduce our schematic model, with illustrations and examples of actual and potential context aware applications to scaffold our self-control behavior.

3. SELF-CONTROL SCENARIOS

The issue of self-control is a very complicated phenomenon. Despite the fact that all humans share the same basic cognitive machinery for evaluation of short and long-term rewards (as revealed by the extremely wide applicability of the hyperbolic discount curve) each case has to be evaluated on an individual basis.

Obviously, some problems are more severe than others. Common forms of laziness should not be equated with full-blown addiction just because both conditions find their root in similar mechanisms of reward evaluation in the brain. In the following sections we present a model and a discussion of how the conceptual tools we have proposed and discussed in the paper (*computer-mediated extrospection and distributed motivation*) can be applied and tailored to the demands of particular self-control problems. We start with comparatively less difficult problems, and move on to harder ones (this progression and our theoretical tools are summarized in figure 1).

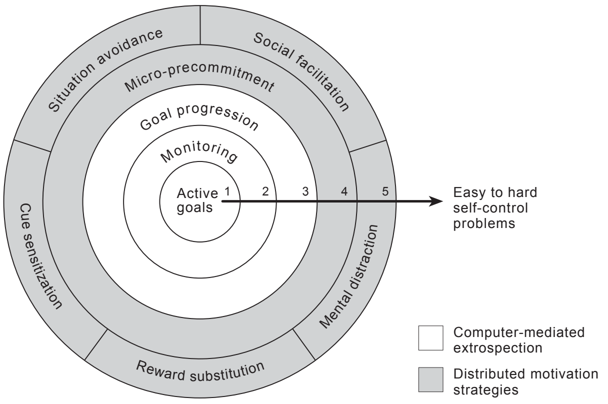

Figure 1. Basic model and applications of the concept of CME and distributed motivation to the problem of self-control. The five circles represent a progression from (comparatively) easy to harder problems. The outer circle contains some representative examples of cultural strategies of distributed motivation that can be plugged into any scheme of precommitment. The fact that the model does not cover other more traditional approaches to self-control (i.e. purely mentalistic approaches, or pharmaceutical inter-ventions) should not be taken as evidence of an opposition to such endeavors; it is only meant to represent approaches that are amenable to manipulation by sensor and computing technology.

3.1 Active Goal Representation

In our discussion of the concept of distributed motivation we catalogued some of the many cultural strategies of self-control that people employ in their daily lives, and noticed how they often fail because of the lack of crucial binding power. However, degree of binding is not the only variable that determines success or failure of any particular attempt at self-control. Sometimes the solution is actually easier than we might first think.

At the most basic level of analysis an often overlooked factor is the nature of the representation of the goals we are striving for. An example from the clinical literature provides a good illustration of this. Patients who have suffered damage to the prefrontal cortex (PFC) often face dramatic impairments in their ability to engage in behaviors that depend on knowledge of a goal and the means to achieve it. They distract too easily, and are said to be "stimulus bound" (Miller, 2000; see also Manuck et al., 2003). Despite this, rehabilitation studies have shown that performance on difficult tasks can be fully restored to the level of control subjects, by the simple use of a wireless, auditory pager system that alerts the patients at random intervals to think about their goals and what they are currently doing (Manly et al., 2002). In this example the pager does not function as a specific memory prosthesis, like a day-planner, or a PDA; it is not telling the patients *what* to do. It is a cheap, global signal that tells them to think about what it was they *really wanted to do*. Similarly, for normal people, there is reason to believe that many of our common failures to follow through on goals and plans, simply stem from an inability to continuously keep our goals active in the face of a bewildering array of distracting (and of course, often tempting) stimuli. Maintenance of behavioral goals is a full time job even for people with perfectly intact prefrontal structures (Miller & Cohen, 2001). "Preferences are not effortlessly stable, the truth is that we manage them, construct them, treat them strategically, we confound them, avoid them, expect change in them and suppress them" (Sally, 2000a, p.690). As is revealed by the wireless pager example, the representational and coordinative power of the human PFC can easily be eclipsed by the intelligence inherent in well-designed cultural artifacts and environments.

Thus, the first tier in any CME-based program for alleviating problems of self-control focuses on maintaining important goals in an active state. Specific types of enhancements to prospective memory exist in countless forms: from post-it notes, to computerized calendars, to ubiquitous context-aware systems like MemoClip (Beigl, 2000) that allow users to associate items or actions to be remembered with specific geographical locations. More general systems, like the wireless pager system described above, have been far less extensively explored. This is unfortunate, because such systems could occupy an important niche that traditional remembrance agents cannot fill. What CME-systems like the wireless pager promise to do, is to act like *a pacemaker for the mind,* a steady signal or beacon to orient our own thinking efforts. It would not require us to specify all our actions in advance (and then give reminders to do those things), but instead encourage us to think back, and apply the knowledge of our prior goals to whatever situation we happen to find ourselves in at the time of the alert.

A further reason to explore such applications comes from recent findings in basic learning theory. Nelson and Bouton (Nelson & Bouton, 2002; see also Myers & Davis, 2002) have found that a basic asymmetry exists between initial learning in any domain, and subsequent attempts at unlearning such behavior (for example, eating or drinking habits we would like to change). With few exceptions, initial learning is far less *context-dependent*, while attempts at unlearning generally only work in the specific context where the training took place (for example, in a specific environment, or in a specific state of mind, or even at a specific time, see Nelson and Bouton, 2002[2]). This means that the risk of relapse is always great unless meticulous care is taken to control for contextual variables that could be of importance. However, Nelson and Bouton (2002) have also shown that this problem can be substantially alleviated by conditioning the retraining to a salient object that is accessible in practically any context (i.e. the object in effect works as a portable context). In the light of the previous discussion, a system like the wireless pager described by Manly et al. (2002) could, with proper preparation, work both as a beacon that is used to re-engage attention to our goals and simultaneously as a signal to (more or less automatically) *inhibit* our bad habits. This would be a powerful example of computing technology that supplies and blends influences from both the "cognitive" and the "motivational" domains.

3.2 Self-Monitoring

The second tier of defense against self-control lapses introduces a more powerful form of CME: one that couples the randomized alert with a context-aware system able to recognize user behavior that may signal impeding breakdown, or react when it finds the user in specific "contexts of temptation." In our previous discussion of CME we mentioned the fact that, in relation to real-life self-control problems, people often fail to uphold a sufficiently high level of self-monitoring. The phenomenology of lapse behavior is often completely bereft of any feeling of us having weighed and considered different alternatives, and then finally succumbed to the temptation. Instead we often just *find ourselves*, habitually or absent-mindedly, having performed the act we wanted to avoid.

CME designed to support user self-monitoring could be employed on a scale of both *macro* and *micro*-prediction (see the discussion in section 2.2). Macro prediction would be enabled by sifting through large amounts of context-data relating to lapse-critical behavior (e.g. neural network approaches, or Bayesian user modeling) and identifying "danger-cues" that could serve to augment and increase the self-knowledge of the user. Micro prediction, on the other hand, would be based on more intimate context measures like the psychophysiological state of the user. In this case, the prediction should be situated at the moment of activity, and come (minutes or seconds) *before* the actual action is performed. For some types of self-control problems this will be comparatively easy. For example, any goals having to do with strong emotions (like trying to become a less aggressive person, or trying to stifle unproductive anger in marital disagreements) will be an ideal target for CME micro prediction. As Elster (2000) has pointed out, advice about emotion regulation most often fail simply because they come *after* the unwanted emotion has already been aroused, and taken full effect upon behavior. At an earlier stage such advice might have been perfectly effective (i.e. here the proper assessment of the need for self-control is as important as the control itself). Considerable research already exists on psychophysiological markers that indicate the implicit buildup or expression of emotional states, not only for anger and aggression, but also for more subtle conditions like frustration, stress and anxiety (e.g. Caccioppo et al., 2000; Healy & Picard, 1998). Promising efforts are also underway to identify similarly predictive profiles for less obviously emotional behavior like smoking and gambling (Warren & McDonough, 1999; Blanchard et al., 2000). To increase the chances of finding predictive regularities, CME-technology would add an additional layer to these techniques by allowing the measurements to be individually calibrated over time and multiple contexts. As an example of this, the recently launched *BioMod* project hosted by the MIT Affective Computing Group, aims to develop individually tailored psycho-physiological markers of craving-induced stress in smokers trying to quit, and to use this in a large-scale prevention program (more on this in section 3.5 below).

3.3 Goal Progression

Following up on the theme of self-monitoring, the third tier introduces devices or technologies that enable us to better appreciate our level of goal progression. As we mentioned in the earlier discussion of CME, there is a world of difference between lighting up a cigarette that happens to be the 24th of the day, and knowingly and willingly smoking the 24th cigarette of the day. But while CME technology could provide substantial help with monitoring of goals in relation to clear cut objectives like dieting or smoking (it is a relatively straightforward task to implement context-aware devices that could count the amount of calories or cigarettes consumed) it promises to provide an even greater impact in relation to goals that are more abstract, nebulous or distantly long-term. For example, imagine someone that has decided to become a more amiable and caring person. How would she go about fulfilling this goal, and how would she know when she has fulfilled it? One solution that is realizable by means of context-aware technology is to operationalize the goal in such a way as to be able to get discriminating feedback on the outcome of her behavior. This is a perfect job for context-aware CME-technology. What computers do best is to capture, record, store and analyze data. With the help of ubiquitous or wearable computing devices, conditions of "goal-attainment" could be specified, and used as an objective comparison for the agent involved. Criteria could be set in relation to any behavior, or activity, or reaction of value that can be automatically captured (number of smiles received, time spend in charity organization service, galvanic skin responses

---

[2] Technically, this means that learning to break a bad habit does not involve *unlearning* the old patterns, but rather that *a new form of learning* has been established that (in certain contexts) *inhibits* the old learning. For details, see Nelson and Bouton (2002).

that indicate deception and lying, environmental contexts that suggest pleasurable social interaction, number of scheduled appointments met in time, etc.). But would this really capture all there is to being an amiable person? No, probably not, but that does not detract from the fact that any change in behavior in the direction towards such a goal, would be for the better. In our view, the role of CME in such cases could be seen as a form of *scaffolding* that get people started in the direction towards some abstract or long-term goal. When the behavioral change has gained some momentum, the CME-scaffolding can be dropped in order for more complex (and less measurable) behaviors to flourish.

Another similar, but subtly different role for computational technology in monitoring goal-attainment and goal-criteria is provided by Ainslie (2001). He discusses the difficult problem of trying to establish self-controlled behavior by applying and following *principles*. He argues that in the cultural sphere, and over the lifetime of an individual, a natural evolution of principles takes place, such that (with very few exceptions) principles come to evolve away from what we ideally would like them to do, to instead focus on what is clear and simple and easy to uphold. Thus, an alcoholic that is lucky enough to recover, does not recover as a "social" drinker with a controlled (and presumably) positive intake of alcohol, but as one that abstains from all forms of drinking (Ainslie, 2001; see also discussion in Rachlin, 2000). Total abstinence as a principled approach is much easier to uphold because it leaves no room for subjective interpretation (a beer together with a steak is no real drink, another drink will not hurt me because I have no more cash on me, etc.), and so it does not put the user on a slippery slope. On the other hand, as Ainslie (2001) forcefully argues, what such principles completely ignore, is that this situation might often not be anywhere near what the subject would really want their lives to be like. Again, what CME can bring to this situation is the promise of using computing technology to precisely measure conditions of behavior and criteria for goal-attainment, in order to effectively emulate the function of principles but without having to settle for the few cases that are so clear cut that our ordinary senses can reliably tell them apart (i.e. we could imagine that with finely tuned sensor and computing equipment, the "social" drinker could live by a CME augmented principle that said that she is only allowed to drink once every other month, or only a certain amount each week, or only if she is at a party of a certain size, etc.).

3.4 Micro Precommitment

Returning now to the core question of time-inconsistent reward evaluation, the fourth tier of defense brings us back to the issue of distributed motivation and methods for self-binding. While active goal representation, swift and accurate self-monitoring, and monitoring of goal-progression are important CME-strategies, they are clearly less applicable in cases of genuine reward conflict. In such cases, precommitment is the right strategy to apply. On the other hand, reward-conflicts come in many different flavors, and often it is not the binding power as such that determines the value of any specific scheme of precommitment. As we outlined in our earlier discussion of ubiquitous precommitment technology, what technology has to offer the age-old strategy of precommitment (apart from more binding-power) is a much-lowered cost and a much-increased range of operation. This is good news, because some species of precommitment need to be fast and easy to set up, and should come at a very low cost. For example, we have remote controls for many electrical appliances that enable us to turn them on and off at our convenience. But we have no remotes that allow us to turn appliances off in a way that, within a set limit of time, *we cannot turn them on again* (for TV and web-surfing, we have things like parental or employer control devices, that can block certain channels or domains, but we have no effective equipment for *self*-binding). We can of course always climb under the sofa, pull the plug and the antenna from the TV, and put them in a place we cannot easily reach (to make TV-viewing relatively inaccessible), but such ad-hoc maneuvers are generally too costly and cumbersome to perform in the long run. The trick is to strike a balance between inaccessibility and flexibility. That is, for many behaviors and situations we would like to be able to make quick, easy, but transient precommitments, that allow us to move beyond some momentary temptation, but then expire so as not to further limit our range of alternatives. We call this *micro precommitment* (MPC). MPC finds its primary use when the temptations we are dealing with are not overwhelming, but still noticeable enough to bring us to the fall.

As an example, imagine a cell-phone based location-aware system (using GPS or any other modern positioning technique) where we can instantaneously "tag" different places from which we wish to be kept. The mechanism for tagging could be as simple as having the phone in the same "cell" as the object to be tagged, or having a place-map database in the phone that allows for distance independent blocking. Let us now say we have a minor shoe-shopping compulsion, and walk around town on an important errand. Walking down the street with this system we could, with just a brief moment of forethought, tag an upcoming tempting shoe-store. The tagging could have any number of consequences, like locking our wallet or credit-card, or even tuning the store-alarm to go off if we enter the premises. The point of MPC is *not* to set up consequences that represent maximally strong de-terrents. Quite the opposite: it is a technique suited for temporarily bringing us past small but nagging distractions. Tomorrow, when we have no important errands anymore, we might want to shop for shoes again, and would not want to spend our time unwinding a too forceful and elaborate precommitment scheme. In fact, since MPCs, in our view, should be as easy and cheap as possible to instigate, they *should also not be allowed to have costly or long-term consequences*.

3.5 Precommitment

The final tier in our program starts out where MPC leaves off. While MPCs are swift and cheap and play with low stakes and short-term consequences, regular precommitment holds no such limits. For precommitment the amount of binding power and the cost of engagement are determined in relation to the magnitude of the problem, and may be as strong as any agent desires. In contrast to MPC, regular precommitment should not come easy. To make sure that the binding represents a "true" preference, a certain amount of inertia ought to be built into any precommitment decision procedure (for a sensitive discussion of how to handle this problem, see Elster, 2000). For example, some larger casinos give patrons prone to too much gambling the option of having themselves banned from playing. Since casinos are generally equipped with rigorous security and surveillance systems, the ban can be very effectively enforced. However, one can not just walk up to the entrance cashier and ask to be banned. The decision must be made in dialog and with council from the casino management, because once you are banned *the casino will not be coaxed into letting you in again*. As would be expected from a compulsive gambler, you soon find yourselves back at the gates trying to undo your former decision. It is at this point that the casino enforces the bind by bluntly disregarding your pleas (and if the commitment was made in too light a manner, this would be an unfortunate outcome).

As we explained in our earlier discussion of ubiquitous precommitment technology, the prime strength of such technology is the manifold of new possibilities for manipulating varieties and degrees of binding it introduces. The question is: are these benefits substantial enough to allow us to fashion realistic scenarios for the alleviation of more difficult problems of self-control, such as craving and addiction? We believe so.

Craving and addiction are extremely difficult topics to approach. Behavioral abnormalities associated with addiction are exceptionally long-lived, and currently no reliable remedies exist for the pathological changes in brain-reward systems that are associated with prolonged substance abuse (Nestler, 2001; Everitt, Dickinson & Robbins, 2001; Robinson & Berridge, 2003). With reference to precommitment, it is sometimes said that it is a limited strategy for handling things like addiction, because in the addicted state we supposedly never find a clear *preference platform* from which to initiate the precommitment (i.e. we do not know which of our preferences that are the "true" ones). Rachlin (2000) writes: "Instead of clearly defined points of time where one strong preference gives way to its opposite we generally experience a continuous opposition of forces and apparently random alternation between making and breaking our resolutions" (p. 54). This state of *complex ambivalence* (as Rachlin calls it) also makes it likely that a fierce arms-race will be put in motion by the introduction of any scheme of precommitment, where the addicted subject will waste precious resources and energy trying to slip through the bind of the commitment. The drug Antabuse illustrates these problems. If you take Antabuse and then have a drink, you will experience severe pain. Thus, taking Antabuse is a form of precommitment not to drink alcohol. However, alcoholics have been known to subvert the effects of the drug by sipping the alcohol excruciatingly slowly, and some even drink the alcohol despite the severe pain (Rachlin, 2000). Also, the outcome of Antabuse treatment has been generally less than satisfying because many alcoholics decide against taking the drug in the first place.

In our view, this example should be taken as a cautionary tale for any overly optimistic outlook on the prospects of precommitment technology to handle really tough cases like addiction, but we do not believe it warrants a general doubt about our approach. As is evident by the fantastically prosperous industry for the supply of services and products that purports to alleviate problems of self-control (in practically any domain of life) people are willing to take on substantial commitments, in terms of time, energy, and resources, to change their current ways of life.

Take smoking as an example. What would a ubiquitous precommitment scheme for helping smokers to quit look like? Firstly, as a foundation, some means of detecting the presence or absence of smoking-related context is needed. The context could be built from observation of the actual smoking, from traces of smoking (from smoking-related behavior patterns, or from psychophysiological concomitants of smoking), and many types of sensors could be used to generate the match. For example, one sensor-platform that might be used in the near future to provide robust and efficient measurement, is in-blood substance detection. In relation to diabetes treatment, Tamada, Lesho and Tierney (2002) describe a host of emerging *transdermal* (through the skin) techniques for measuring glucose levels in the blood[3]. While not perfected yet, such sensors can be worn continually and unobtrusively by diabetics to efficiently monitor and manage their blood sugar levels. A similar system could easily be envisaged for nicotine[4]. Yet, as Gellersen, Schmidt, and Beigl (2002) have shown, a combination of many cheap and overlapping environmental sensors (i.e. things like temperature, acceleration, light, movement, etc.) might provide equally robust context-measurement as a specialized subcutaneous device.

---

[3] Nicotine delivery skin patches are an example of transdermal technology working in the other direction, where the molecule of interest is moving into the body rather than out of it.

[4] If we want to limit ourselves to existing technologies, CO is considered to be a very reliable indicator of smoking, and products monitoring the CO level in exhaled air have been used for a number of years (e.g. the Smokerlyzer™). Using saliva samples is currently the fastest and least obtrusive way of detecting nicotine, and products for this purpose have also been around for some time (e.g. NicAlert™, Accutest®).

The great boon of ubiquitous precommitment technology is that once the basic sensing of context is in place (in the previous fictional example, transdermal nicotine blood level detection), a multitude of distributed motivational strategies can be latched onto it, and varieties of binding can be added or subtracted depending on the nature and severity of the case. The versatility of the platform also allows for overlapping and partially redundant incentives to be put in place. To take a dramatic example, for providing strong and relentless binding, a wireless bracelet for nicotine monitoring could be hooked up directly to the bank account of the participating subject, and simply withdraw money in proportion to the amount of smoking the subject does. But to prevent loss of money, an anticipatory CME backup-system that detects "lapse-critical" behavior (as described in section 3.2 above) could be employed alongside the nicotine-bracelet, and make automatic support calls to other participants in the program if the subject is in danger of taking a smoke; a very similar approach to this is taken in the MIT BioMod project we described earlier. The extracted psychophysiological markers of "lapse-critical" stress levels will be used to automatically relay cell-phone calls to a support center where trained professional can answer to the needs of the subject. In all, we foresee that while exceptionally strong single precommitment criteria can be put in place (i.e. you loose all your money if you smoke one single cigarette), it is the possibility of mixing and merging many less forceful strategies in one system that will provide the greatest benefits. Most likely, venerable cultural strategies like situation avoidance (e.g. the shoe-store "tagging" example), social facilitation, reward-substitution, etc., will experience a strong resurgence in the hand of ubiquitous technology for distributed motivation.

4. DISCUSSION

4.1 Flexibility and Rigidity

Some researchers have expressed great pessimism about the ability of context-aware systems to make meaningful inferences about important human social and emotional states, and believe that context aware applications can only supplant human initiative in the most carefully proscribed situations (Bellotti & Edwards, 2001). We are in no position to assess the finer details behind this pessimism, but it must be noted that this problem is far less pressing for the proposed domain of ubiquitous self-control technology discussed in this article. Precommitment technologies offer people the option of temporary, but forceful, binding, aided by computer systems that will not be swayed or cajoled, and *it is through their very inflexibility that these systems have the potential to support individual self realization*. As Dennett (2003) notes, in the domain of self-control effectively constraining our options gives us more *freedom* than we otherwise would have had.

Nevertheless, the rigidity of these technologies may sometimes be a weakness. Although precommitment technology increases the likelihood of attaining individually set goals, there is the attendant risk that people will lock themselves into inappropriate precommitments, and waste time and effort fulfilling needless obligations. It can be relatively easy to precommit, but comparatively hard to foresee potential conflicts with other valued goals and preferences one might have. Preferences are also subject to change and the future can bring unexpected opportunities, as well as emergencies, that we want to be able to respond to. In addition, there is the possibility that, once precommitted, some people will expend wasteful resources on increasingly elaborate countermeasures. These concerns must all be taken seriously, but are not as severe as they might appear at first blush. Clearly, the proposed systems will have to leave room for a host of pre-programmed contingencies, as well as a fixed number of predetermined digressions. If pre-commitment does come with a price (temporally limited freedom) this must be taken in relation to the valuation of the goal one wishes to attain; if a particular precommitment seems arduous this also has to be judged in relation to potential benefits. Ultimately, people will be free to use these systems or not, as they see fit, and to weigh potential benefits against possible costs.

A degree of inflexibility is essential to the successful working of these kinds of systems, but as we discussed earlier, we foresee a range of different kinds of binding, of various degree and type, that can be combined in regimes suitably coupled to particular issues. In a possible scheme, the range of permissible actions is large to start with, but slowly curtailed in response to flagging willpower. In an alternative scheme, permissible actions are limited at the onset, but then expand as the need for support slowly wanes (this kind of regime might be seen as a form of motivational scaffolding or training wheel).

4.2 Ethical Considerations

With the introduction of new technology also come new ethical considerations. With lots of information about the user being picked up and circulated (information about location, behavior, affective and cognitive states etc.) there is the risk that the information could be put to unsavory use. A widely shared worry is that this kind of information is a threat to privacy. The problem of privacy is one that besets the whole field of ubiquitous computing, and there have been some viable and thoughtful suggestions of how this could be handled (Dey, Abowd & Salber, 2001; Bellotti & Edwards, 2001). Nevertheless, with more intimate, psychological, measures afloat the problem of privacy is perhaps even more pressing.

A standard solution to the problem of privacy in context-awareness, is to increase the "transparency" of the applications, making users aware of what kind of information about them is being registered, and what actions are about to be taken in response. On the other hand the obvious problem with trying to increase the transparency of contextually aware applications is that constant requirement of notice could easily overwhelm users, and disrupt their activities (think of such a relatively simple task as management of browser cookies). It is a reasonable question to ask whether users would provide attention and direction if they were constantly bombarded by requests from all kinds of systems (temperature and light settings, image-capture, notes and file-sharing, driver safety customizations, etc.). Studies have shown that people are both poor at handling such updates, and unwilling to receive them (Belloti & Edwards, 2001; Ackerman, Darrell & Weitzner, 2001). In this regard, ubiquitous CME and precommitment technology have a clear advantage over many other context-aware applications. Again, what we would like to stress here is that these concerns are not nearly as pressing for a scheme of representing and augmenting user-perceptions of context – *to the users themselves*. The emphasis on explicitness of interaction (at the loss of some ease and efficiency) is not a problem for the manufacturing of self-control technology; quite the opposite. Given the personal importance potentially at stake in such examples it would be dangerous and irresponsible to allow the process to proceed entirely implicitly. People already spend a great deal of time and effort trying to regulate and manage their cognitions, emotions and behavior. Our project only proposes to usurp resources already devoted with scant success to similar causes.

Another ethical concern for the prospect of ubiquitous CME and precommitment technology, is that with the availability of these kinds of systems there is a risk that people will be put under undue pressure to employ them (by family members or employers, or maybe even government agencies). If precommitment is too easy to set up, and the binding forceful, there is the risk that people get stuck in precommitments they wouldn't have chosen "under a calm moment of reflection." We must therefore ensure that precommitments are not entered into under duress, but at appropriate times and for appropriate reasons. A related concern is that pressure might come from the system itself. Systems like these could be purposefully designed to be persuasive: to lure users into setting up various kinds of precommitments and obligations. This is one point at which we clearly differ from the avowed, but related, goals of *persuasive computing* (e.g. Fogg, 2003). Although we have selected some possible societally beneficial areas for remedial action, our intent is not one of persuading people to participate. All along, the premise of our work have been that self-control problems only apply to situations in which the subject herself considers it to be a problem (choosing a "lesser" reward *against ones own recognized best interest*). This does not mean that a "persuasive" or paternalistic stance is never justified (see discussion in Fogg, 2003; and O´Donoghue & Rabin, 2003), but it has not been part of our concern here.

4.3. Summary

We have provided a basic model and a host of examples of how the twin concepts of CME and distributed motivation can be applied and tailored to problems of self-control. Our scheme of classification is intended to provide an overview of the impact ubiquitous sensor and computing technology might have on the domain of self-control. It is our hope that the model and our discussion will provide a principled and useful way for designers of human–computer interfaces and context-aware systems to approach the domain of self-control, as well as to provoke further debate on the possible role of computing technology in matters of human motivation.

The technologies and theories proposed here are, we believe, well grounded, but need to be tested in an arena of real self-control problems and against a background of technological constraints. For the future, we envision precommitment technologies and tools of computer-mediated extrospection that can be configured by the users themselves, in ways and for purposes we cannot yet anticipate. It is in the ecology of devices, human needs and ingenuity that the field will take shape.

ACKNOWLEDGEMENTS

We would like to thank the participants of the Lund University Cognitive Science seminar for their input on an earlier draft of the paper. Special thanks go to Jens Månsson for his contributions at the early stages of this project.

REFERENCES

Ackerman, M., Darrell, T., & Weitzner, D. J. (2001). Privacy in context. *Human–Computer Interaction*, 16, pp. 167–176.

Ainslie, G. (1999). The dangers of willpower. In J. Elster and O-J. Skog (Eds.), *Getting hooked: Rationality and addiction* (pp. 65–92). Cambridge: Cambridge University Press.

Ainslie, G. (2001). *Breakdown of will*. Cambridge: Cambridge University Press.

Bechara, A., Ranel, D., & Damasio, A. R. (2002). The somatic marker hypothesis and decision-making. In F. Boller and J. Grafman (Eds.), *The handbook of neuropsychology*, vol. 7 (pp. 117–143). Amsterdam: Elsevier Science.

Beigl, M. (2000). MemoClip: A location-based remembrance appliance. *Personal Technologies*, 4, pp. 230–233.

Bellotti, V. M., & Edwards, W. K. (2001). Intelligibility and accountability: human considerations in context aware system. *Human–Computer Interaction*, 16, pp. 193–212.

Blanchard, E. B., Wulfert, E., Freidenberg, B. M., & Malta, L. S. (2000). Psychophysiological assessment of compulsive gamblers' arousal to gambling cues: A pilot study.

Applied Psychophysiology and Biofeedback, 25, 3, pp. 155–165.

Buckner, R. L., & Wheeler, M. E. (2001). The cognitive neuroscience of remembering. *Nature Reviews Neuroscience*, 2, pp. 624–634.

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., & McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin*, 126, 6, pp. 829–843.

Clark, A., & Thornton, C. (1997). Trading spaces: Computation, representation, and the limits of uniformed learning. *Behavioral and Brain Sciences*, 20, pp. 57–90.

Clarkson, B. (2002). *Life patterns: structure from wearable sensors*. Unpublished doctoral dissertation, MIT Media Lab.

Dehaene, S., & Naccache, L. (2000). Toward a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, 79, pp. 1–37.

Dennett, D. C. (1991). Real patterns. *Journal of Philosophy*, 89, pp. 27–51.

Dennett, D. C. (2003). *Freedom evolves*. London: Allen Lane.

Dey, A. K., Abowd, G. D., & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human–Computer Interaction,* 16, pp. 167–176.

Duchaine, B., Cosmides, L. & Tooby, J. (2001). Evolutionary psychology and the brain. *Current Opinion in Neurobiology*, 11, pp. 225–230.

Elster, J. (2000). *Ulysses unbound*. Cambridge: Cambridge University Press.

Everitt, B. J., Dickinson, A., & Robbins, T. W. (2001). The neuropsychological basis of addictive behavior. *Brain Research Reviews*, 36, pp. 129–138.

Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. San Francisco: Morgan Kaufmann Publishers.

Frederick, S., Loewenstein, G. & O'Donoghue, T. (2003). Time discounting and time preference: A critical review. In G. Loewenstein, D. Read and R. Baumester (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 13–86). New York: Russell Sage Foundation.

Gardner, G., & Halweil, B. (2000). Underfed and overfed: The global epidemic of malnutrition. Washington D. C: *Worldwatch Paper* 150.

Gazzaniga, M. S. (Ed.), (2000). *The new cognitive neurosciences*, 2 ed. Cambridge, MA: MIT Press.

Gellersen, H. W., Schmidt, A. & Beigl, M. (2002). Multisensor context-awareness in mobile devices and smart artifacts. *Mobile Networks and Applications*, 7, pp. 341–351.

Goldman, A. I. (1993). Consciousness, folk-psychology, and cognitive science. *Consciousness and Cognition*, 6, pp. 364–382.

Healey, J., & Picard, R. (1998). Digital processing of affective signals. MIT: Technical Report.

Hutchins, E. (1995a). *Cognition in the wild.* Cambridge, MA: MIT Press.

Hutchins, E. (1995b). How a cockpit remembers its speed. *Cognitive Science*, 19, pp. 265–288.

Kacelnik, A. (2003). The evolution of patience. In G. Loewenstein, D. Read and R. Baumester (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 115–138). New York: Russell Sage Foundation.

Kramer, A. F. & Weber, T. (2000). Applications of psychophysiology to human factors. In J. T. Cacioppo, L. G. Tassinary and G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 794–814). Cambridge: Cambridge University Press.

LeDoux, J. (2002). *The synaptic self.* New York: Viking.

Levin, D. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition*, 7, pp. 397–412.

Malle, B. F., Moses, L. J., & Baldwin, D. A. (Eds.), (2001). *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press.

Manly, T., Hawkins, K., Evans, J., Woldt, K., & Robertson, I. (2002). Rehabilitation of executive function: Facilitation of effective goal management on complex tasks using periodic auditory alerts. *Neuropsychologia*, 40, 3, pp. 271–281.

Mann, S. (2001). *Intelligent image processing*. New York; Chichester: Wiley.

Manuck, S. B., Flory, J. D., Muldoon, M. F., & Ferrell, R. E. (2003). A neurobiology of intertemporal choice. In G. Loewenstein, D. Read and R. Baumester (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 139–172). New York: Russell Sage Foundation.

Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review*, 108, 1, pp. 96–112.

Medina, J. F., Repa, J. C., Mauk, M. D., & LeDeoux, J. E. (2002). Parallels between cerebellum- and amygdala-dependent conditioning. *Nature Reviews Neuroscience*, 3, 2, pp. 122–131.

Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, 1, pp 59–65.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience,* 24, pp. 167–202.

Mokdad, A. H., Bowman, B. A., & Ford, E. S. (2001). The continuing epidemic of obesity and diabetes in the United States. *JAMA 286*, 10, pp. 1195–1200.

Monterosso, J., & Ainslie, G. (1999). Beyond discounting: Possible experimental models of impulse control. *Psychopharmacology*, 146, pp. 339–347.

Myers, K. M., & Davis, M. (2002). Behavioral and neural analysis of extinction. *Neuron*, 36:4, pp. 567–84.

Nelson, J. B., & Bouton, M. E. (2002). Extinction, inhibition, and emotional intelligence. In L. F. Barrett and P. Salovey (Eds.), *The wisdom in feeling: Psychological processes in emotional intelligence* (pp. 60–85). New York; London: The Guilford Press.

Nestler, E. J. (2001). Molecular basis of long-term plasticity underlying addiction. *Nature Reviews Neuroscience*, 2, pp. 119–128.

O´Donoghue, T., & Rabin, M. (2003). Self-awareness and self-control. In G. Loewenstein, D. Read and R. Baumester (Eds.), *Time and decision: Economic and psycho-*

logical perspectives on intertemporal choice (pp. 217–243). New York: Russell Sage Foundation.

O´Regan, J. K., & Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences,* 24, 5, pp. 939–1011.

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Science*, 6, 2, pp. 93–102.

Parson, R., & Hartig, T. (2000). Environmental psychophysiology. In J. T. Cacioppo, L. G. Tassinary and G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 815–846). Cambridge: Cambridge University Press.

Rachlin, H. (2000*). The science of self-control*. Boston: Harvard University Press.

Read, D. (2003). Subadditive intertemporal choice. In G. Loewenstein, D. Read and R. Baumester (Eds.), *Time and decision: Economic and psychological perspectives on intertemporal choice* (pp. 301–322). New York: Russell Sage Foundation.

Rhodes, B. J. (2000). Margin notes: Building a contextually aware associative memory. In *Proceedings of the 2000 international conference on intelligent user interfaces* (pp. 219–224). New Orleans, LA: ACM Press.

Robinson, T. E., & Berridge, K, C. (2003). Addiction. *Annual Review of Psychology*, 54, pp. 25–53.

Rogers, Y. & Ellis, J. (1994). Distributed cognition: an alternative framework for analysing and explaining collaborative working. *Journal of Information Technology*, 9, 2, pp. 119–128.

Rolls, E. T. (1999). *The brain and emotion*. Oxford: Oxford University Press.

Schacter, D. L. & Dodson, C. S. (2001). Misattribution, false recognition and the sins of memory. *Philosophical Transactions of the Royal Society*, 356, pp. 1385–1393.

Singletary, B., & Starner, T. (2001). Learning visual models of social engagement. In *Second international workshop on recognition, analysis and tracking of faces and gestures in real-time systems*, 2001, Vancouver, Canada.

Sally, D. (2000a). Confronting the sirens: Rational behavior in the face of changing preferences. *Journal of Institutional and Theoretical Economics*, 156:4, pp. 685–714.

Sally, D. (2000b). I, too, sail past: Odysseus and the logic of self-control. *Kyklos*, 53:2, pp. 173–200.

Tamada, J. A., Lesho, M., & Tierny, M. J. (2002). Keeping watch on glucose. *IEEE Spectrum Online*, April.

Warren, C. W. & McDonough, B. E. (1999). Event-related brain potentials as indicators of smoking cue-reactivity. *Clinical Neurophysiology*, 110, pp. 1570–1584.

WHO (2001). *The world health report*. Mental health: New understanding, new hope. Geneva: World Health Organization.

World Bank (1999). *Development in practice – Curbing the epidemic: governments and the economics of tobacco control*. Washington DC: World Bank.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, 21, pp. 179–217.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108, pp. 483–522.