Aalborg Universitet

AALBORG UNIVERSITY

DENMARK

HANDS Deliverable D4.1.1 Report on Test Methodology

Schärfe, Henrik

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University
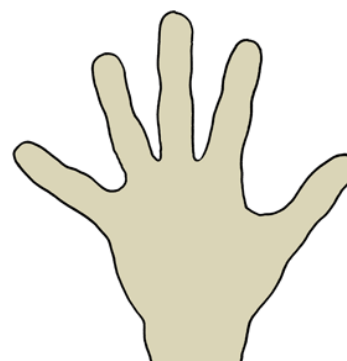
HANDS DELIVERABLE 4.1.1

Grant Agreement number:
224216

Project acronym:
HANDS

Project title:

Helping Autism-diagnosed teenagers
Navigate and Develop Socially

Funding Scheme:
Collaborative Project

Deliverable description

Delivarable no:	4.1.1
Delivarable name:	Requirements Report
WorkPackage No:	WP4
Lead beneficiary:	AAU
Nature:	Report
Dissimination level:	PU
Author	Henrik Schärfe, AAU

Public Summary(max 350 characters)

This deliverable contains theoretical and practical considerations regarding the evaluation of Persuasive Technology as it unfolds in the HANDS project. The document is structured as follows. In section 2, evaluative aspects of the term 'persuasion' are considered. In section 3, a brief description of principles in persuasive technologies is offered. In section 4, we consider Formal Concept Analysis as a tool for analyzing relationships between parts of the toolset and of persuasive principles. In section 5 we consider logs, quantitative, and qualitative approaches to evaluating the HANDS toolset. Section 6 addresses the differences in evaluating the two prototypes. Section 7 contains a brief plan for collecting analyzing and evaluating data. In section 8 principles for evaluating are summarized.

Project co-ordinator name, title and organisation:

Project Co-ordinator: Professor Peter Øhrstrøm
Organisation: Aalborg University
Tel: +45 9940 9015
Fax: +45 9815 9434
E-mail: poe@hum.aau.dk
Project website address: http://www.hands.hum.aau.dk

D4.1.1 REPORT ON TEST METHODOLOGY

CONTENT

1. EVALUATING PERSUASION

Objectives: The objective of this document of to chart out principles for evaluating the persuasive aspects of applied Persuasive Technology (PT) in the HANDS project. In the following section (section 2) we shall consider various aspects of persuasion, pertinent to the design of an evaluating strategy. We focus here on theoretical considerations concerning persuasion and response, the role of coercion, and intention and persuasion. In section 3 we consider the persuasive roles technology can assume. In section 4 we apply the principles of Persuasive Technology (as presented in section 3) to the HANDS tool set, using Formal Concept Analysis. In section 5 we consider kinds of evaluation, and in section 6 we address the differences in evaluation prototype 1 and prototype 2. Section 7 contains a brief plan for the

collecting and analyzing the material. This section also contains a table with the estimated a resources required on behalf of the schools. brief summery is found. In section 8 the principles for evaluating the persuasive effect and potential in the HANDS toolset are summarized.

2. ASPECTS OF PERSUASION

In this report we will take our point of departure to be the definition of Persuasive Technology (or Captology) as offered by B.J. Fogg:

> Briefly stated, captology focuses on the *design, research, and analysis of interactive computing products created for the purpose of changing people's attitudes or behaviors.* (Fogg 2003)

In particular, we shall utilize the framework for distinguishing between different aspects of Persuasive Technology (hereafter: PT) developed in Fogg's 2003 book: PT as a tool, as a social actor, and as a media/ simulator.

There are, however, some aspects of the term *Persuasion*, and some fundamental issues in the theoretical basis of PT that need to be addressed before we can turn in detail to the matter of developing a methodology for evaluating the persuasive potential and the persuasive effect in HANDS. These initial considerations can be seen as a critique of Fogg's defining framework, but are in fact intended as extension of the existing framework. Several such areas exist, but for the purpose of the report we shall only briefly examine three: Different responses to persuasion, the relation between coercion and persuasion, and the role of technology in relation to intended outcomes.

2.1 RESPONSES TO PERSUASION

Fogg's definition focus on the active purpose of changing attitudes or behaviors, and it is certainly within the scope of HANDS to be able to work with change in attitude and behavior. Changing a behavior, however, presumes that there already is a behavior, and to some extent that this behavior is stable; at least stable enough to be changeable. In his seminal paper from 1980 Gerald Miller suggests that we consider three behavioral outcomes of a persuasive process: Being persuaded as a *Response-Shaping Process*, as a *Response-Reinforcing Process*, and finally as a *Response-Changing Process* (Miller 1980).

Shaping responses means to form an appropriate pattern of responses in cases and situations where no such pattern exists; that is, it applies to cases where people or communities *learn* how to react to new situations or new information. This goes for children who are persuaded to behave in certain ways, or to adults who are adapting to new and maybe unforeseeable situations. This kind of persuasion is commonly employed by teachers and parents, but also by governments and other large organizations in shaping responses for instance to crises situations. This seems highly relevant to the context of HANDS in at least two capacities. We are, in the first place, going to implement hardware and software in contexts where there may not be an appropriate pattern of response. With respect to the teachers this means that they will have to adapt to a new situation where ICT plays a more active role in their work than before. The presence of these technologies and techniques may well, in turn, influence on the relationship between the affected teachers and their colleagues. For the students this also applies to the process of adopting new technology and the changes in routines that comes with that. Although we do avail ourselves of existing practices with respect to computers and mobile phones, the sheer facts that the HANDS software have different functionalities than

software used hitherto; and that PDAs have more functions than ordinary phones, does necessitate the need for shaping appropriate responses. It is also worth noticing in this context that the teachers will have (more) direct access to the student's phones, which means that a new practice most be established; one that rethinks the boundaries between private and shared information.

In the second place, we are not just concerned with implementing technologies, but in fact these technologies will alter certain aspects to the student's life, and it is very likely that new and unforeseen situations to which there are no appropriate responses will arise. In an earlier study, a conflict reportedly arose between teachers and a family in which the student presented the parents with a message on a PDA saying that after supper he had to go outside and smoke two cigarettes. The student had added this item on his own, but the parents initially held the teachers accountable. This is an example of an unforeseen situation to which there was no appropriate pattern of response. As a result of this situation, the phone was no longer used by this student at home. Looking at this situation from the point of view of human centered approaches to implementation, it seems unfortunate that it was not possible to establish an appropriate response. From this perspective, it is clear that the young man had gained a high level of understanding of the role of technology, although still a perspective that needs guidance. From a design perspective, it is easy think of a solution that takes multiple users into account. We could, for instance, color-code entries in the calendar, making it visible who added the entry. Other solutions are possible as well, but the important lesson to be learned from this situation is that we are likely to end up in circumstances where responses must be shaped rather than changed.

When we consider existing ICT practices, this notion of shaping (or creating) responses seems even more crucial to the implementation of HANDS. At this point, no complete picture of the actual ICT use and ICK skills at the schools exist. But we do know that the Helen Allison School, NAS, UK prior to this project had a rule prohibiting the use of mobile phones in school. We also know that several of the students at Svedanskolan, Sweden, do not at this point use mobile phones. The school has practices that involve related technologies, e.g., MP3 players and digital cameras, and in some cases also use of the Handi II. In cases (regarding individual students or entire schools) where mobile phones are not part of everyday ICT practice, special attention should be given to issues of implementation. Obviously, there is substantial difference between a situation in which a number of advanced functionalities are added to an existing practice; and a situation in which the technology in question must first be introduced. It is of great importance to clearly state that absence of existing practices regarding mobile phones in no way detracts from the importance of the schools, nor from the applicability of the HANDS solutions.

Reinforcing responses is essential to a vast number of persuasive situations. It may even be argued that that bulk of our persuasive endeavors aim at reinforcing existing attitudes and behaviors rather than changing them. Sermons, political speeches as well as parental guiding very often take on this form, and it is obvious that in the context of working with students with autism-diagnoses, persuasive strategies that aim at reinforcing certain behaviors should be considered. In particular, the proposed use of a Simple Safe Success Instructor falls in this category along with some intended uses of the Personal Trainer. The Personal Trainer may of course also be used to shape appropriate responses, e.g. in rehearsing for specific events to come. Also the use of interactive diaries and calendars are in many cases supposed to sustain already existing habits and practices, rather than to alter responses.

Situations where persuasion takes on the form of reinforcing are characterized by congruent intentions for the persuader and the persuadee. That is, at least partial congruent, as the persuadee may indeed agree with the

attitudes and values being communicated, but he or she may not necessarily regard the very act of persuading as important.

Changing responses, finally, deals with cases where existing behaviors and attitudes are converted into something different, or replaced by new practices and outlooks. Changing behavior then applies to situations where improving an existing response is insufficient. Instead the purpose may be to replace a violent behavior with a non-violent behavior. In many cases the change of behavior, attitude, and response requires a substantial effort on part of the persuadee. In other contexts than HANDS, good examples include converting from smoking behaviors to non-smoking behaviors, altering a diet or exercise habits and the like.

It should be noted that these three prototypical goals of persuasion are not thought to be mutual exclusive. Indeed overlaps may well occur, but in the context of designing a methodology, different persuasive goals should be taken into account. Evaluating whether the goals have been attained or not indeed depends on which strategies are employed. It may be easier to evaluate cases where we seek to motivate change than cases where we seek to influence behaviors or induce modifications in attitudes, but the latter may quite well be the most important outcomes of the effort. The evaluation should therefore operationalize the persuasive goals with respect to the intended outcome.

2.2 COERCION AND PERSUASION

Most readers will probably recognize the value and integrity associated with Fogg's definition of persuasion as "*an attempt to change attitudes or behaviors or both* (without using coercion or deception)" (Fogg 2003). But the definition also raises questions. In particular the pertinent question of what is meant by coercion. In many cases it is quite clear what coercion is, but it is far from clear how we can avoid it altogether; or even more problematic – if we should avoid it altogether.

Returning again to Miller, we find an important distinction between *direct* and *indirect coercion* that might prove useful in this context (Miller 1980). Direct coercion avails itself of force, be it physical, economical, emotional, or the like. So facing the reality of being grounded or having allowance reduced, a child might be more likely to adhere to instructions of, say, to clear the table. In such cases the success of the persuasive act depends on how credible the threat appears to be. The problem is that we find ourselves in all sorts of situations that are inherently infused with asymmetrical power structures and therefore prone with indirect coercion. It is in fact difficult to conceive an ordinary day where we do not partake in such communications at all. If a professor asks a student to revise a section of a paper (that happens frequently) the persuasive strategy would be non-coercive in the strict sense. An important question is, however: does the situation make any sense without considering the indirect coercion that takes place qua the roles we are given? It seems reasonable to assume that the persuadee would take into account that fact that the paper eventually will have to be graded by the persuader. As such, most interactions in teaching settings may be governed by indirect coercion. On a broader horizon this question touches of the relationship between knowledge and power. Most of us frequently encounter situations where the best idea doesn't seem to win – or get funded. Some have suggested that the better argument rarely is successful if it is not supported by adequate power structures (Flyvbjerg 1998).

In this context we may indeed ask weather coercive-free arguments exist at all? For the purpose of evaluating the persuasive potential and effect in HANDS, two immediate concerns arise.

We must consider the role of the organizations in which the prototypes are placed and tested. That is: evaluation must take into account how the technology is implemented with respect to all parties involved: teachers, students, parents and other stakeholders in the affected environments. This element includes evaluating decisions regarding the devises used in testing the tool set. After discussing the matter of selecting the devices, it was agreed to respect local customs and conditions such that any device in use should conform to a short list of technical requirements as well as be "appropriately cool" for teenagers in the individual countries. Such variations is choice of hardware may influence the results. On the other hand, it is highly probable that a decision to use the same model throughout the schools would pollute the results even more.

We must also consider the role of the teacher as interacting with the students. That is: the evaluation most take into account the interactions between the principal stakeholders with regard to introduction, daily use, suggested and mandatory activities etc.

For both of these areas, implicit power structures and similar factors should be considered in evaluation.

2.3 THE ROLE IN INTENTION AND TECHNOLOGY IN PERSUASION

One of the important theoretical complexes underlying our thought on PT is rhetoric (Hasle 2006; Hasle and Christensen 2008). This perspective highlights certain aspects of communicative acts, and is also one of the legacies considered by Fogg. The basic argument for this follows the line of: Most of our communications have persuasive purposes; rhetoric deals with persuasion; therefore most of our communications can be dealt with as rhetorical. This choice, however, comes at a cost. In particular it focuses on a very direct kind of intention being channeled through some computing technology, but is does not answer the question of how the intention is shaped by the technology.

In many instances it is quite clear who has the intention, and what the intention is. In the classical case of an online bookstore the intention resides in the owners of the store and the intention is to promote sale. The likewise classical example of PT that informs us that "customers who bought this, also bought…" is then to be understood within this frame. But since the information is given via use of computer technology, a certain distance is introduced between the interacting parties. This distance (and other factors that may be considered elsewhere) is in itself facilitating alternative uses of the advice, e.g. a customer may be persuaded to buy the suggested book – but in another store… In this case the intention is only partially fulfilled. We may call this **partial use**.

Consider another classical case: Many cars are equipped with instruments informing the driver of how far the present stock of fuel will last given that the recent circumstances for driving apply to the road ahead. This technology, consisting of a sensory devise, a computing algorithm, and an interface, offers a calculation intended to simply inform. It is, however, very likely that the information is use in slightly derived forms. It may, for instance, motivate the driver to drive more economical. This use, we may call it **derived use**, is not in conflict with the original intention, but something is added to original intended use.

Finally, let us consider yet another classical case: "Rocket Jumping" in computer games. In early computer games some users discovered that firing a rocket launcher into the ground would propel the avatar into the air – consistent with the programmed physical laws governing that fictional world. This "discovery" quickly became popular (maybe due to boredom or creativity) and made it possible to reach areas of the game world that would otherwise remain inaccessible. This development, in turn, led game designers to construct obstacles that could only (or more quickly) be overcome by use of rocket jumps. We may call this **alternative use**.

In the cases of partial use, derived use and alternative use, the success of the PT cannot be measured with respect to the designer's (or the teacher's) intention. Nonetheless such uses may indeed be valuable parts of the overall plan. For instance if a student is suggested a certain SSSI but then locates another that proves more helpful (partial use); if he uses a SSSI in another context than the one it was initially designed for (derived use); or if he uses it for an alternative purpose, e.g. in social interaction with other students (alternative use).

Evaluation should take Partial Use, Derived Use, and Alternative Use into account.

3. TECHNOLOGY IN PERSUASIVE ROLES

Following (Fogg, Lee et al. 2002; Fogg 2003), Persuasive Technologies (PT) takes on three different persuasive roles, namely a tool, a simulator, and a social actor. These roles, in turn, are described through 16 principles that together make up a significant portion of commonly accepted perspectives on PT. A recent attempt to apply the principles into a practice of design and evaluation is found in (Oinas-Kukkonen and Harjumaa 2008). Here the list of principles is slightly expanded; but for the present purposes the 16 principles are thought to suffice.

The principles are as follows:

PT AS PERSUASIVE TOOLS

Principle of Reduction: Reduce complex behavior to simple tasks

Principle of Tunneling: Guide through a process

Principle of Tailoring: Designing for individual needs, interests, personality and usage context

Principle of Suggestion: Offering suggestions at opportune moments

Principle of Self-Monitoring: Tracking performance or status

Principle of Surveillance: Observing other's behavior

Principle of Conditioning: Applying positive reinforcement

PT AS PERSUASIVE MEDIA: SIMULATION

Principle of Cause and Effect: Enabling to observe immediately causal relationships

Principle of Rehearsal: Providing motivating environments in which to rehearse

Principle of Virtual Rewards: Giving virtual rewards for target behavior

Principle of Simulation in Real-World Context: Simulating something in the context to which it belongs

PT AS PERSUASIVE SOCIAL ACTORS

Principle of Attractiveness: Providing visual attracting technology

Principle of Similarity: Resembling the user in some way

Principle of Praise: Providing praise in any modality

Principle of Reciprocity: Users tend to reciprocate favors

Principle of Authority: Assuming a role of authority

A further analysis of the principles renders the fact that in some of them, the designer of the communication takes on the most important role (e.g. reduction, tunneling), in other cases, the user is the more active (e.g. self-monitoring, virtual rehearsal), and finally, some principles places special emphasis on the context in which the persuasion takes place (e.g. tailoring, attractiveness).

Thus, the principles of persuasion forms a way of addressing pertinent questions in relation to the designers / senders, to the actual message and in context in which they occur, and to the users of persuasive technology. In terms of evaluating persuasion, two concerns immediately appear when seen through the PT principles: In the first place, the principles never appear alone. That is: they exist as distinct principles only in the analytical sense, and it follows that the stakeholders in the persuasive communication obtain the role exactly as roles reciprocally defined by one another. In the second place, since the principles addresses different aspects of the communicative situation, an evaluative approach to the principles must take the inter-relationship between the principles into account. Essentially, it is very likely that the most interesting parts of the evaluation arise precisely from combining the principles.

4. AN FCA MODEL OF THE HANDS TOOLSET

Taking a closer look at the HANDS tool set, it becomes clear that important aspects of elements can be described in terms of the PT principles. Principles such as surveillance and authority are important aspects of the Co-Me, and cause / effect simulation and virtual rehearsal are vital parts of the Personal Trainer and also of the SSSI, attractiveness and reciprocity relates to the Individualizer, and so on.

The principles seem to be an interesting way of operationalizing the evaluation also with respect to the technological aspects of the HANDS tool set.

Using a technique known as Formal Concept Analysis we may explore the relationship between different parts of the HANDS tool set, in particular the dependencies that seem to exist between different principles.

FCA is a mathematical theory based on set theory and lattice theory, originally devised by Rudolf Wille a few decades ago. Since then, a scientific community has emerged with conferences specifically dedicated to this topic – see for

instance: (Scharfe, Hitzler et al. 2006). In this context a very brief introduction will suffice. More thorough introductions to FCA can be found in (Ganter and Wille 1997) and (Burmeister 2003). A brief introduction to lattice theory is found in (Sowa 2001). The point of departure in FCA is a simple cross table in which objects are listed in the first column, attributes are listed in the top row, and a 'x' in the table marks the incident that a given object is associated with a given attribute. We call this a formal context. E next consider a set of formal objects. A formal object is a pair with an *intention* and an *extention*. The intention of a formal object is the set of attributes that apply to a given object and conversely, the extention is the set of objects that have precisely these attributes. Next, we can define an *order* over the formal objects and arrange the *formal objects* in a lattice structure. The formal context is next used to generate representations using the OpenSource program ConExp, written by (Yevtushenko 2003).

If we consider the 16 principles as objects and their roles in persuasion, their direction towards elements in the communication and the parts of the tool set as attributes, the following cross table emerges:

| PT principle	tool	simulation	soc actor	sender oriented	user oriented	contet oriented	HIPD	SSI	PersT	Tin	Spo	CoMe
Reduction	X			X			X	X	X			X
Tunneling	X			X			X	X	X		X	
Tailoring	X					X	X	X	X		X	X
Suggestion	X					X	X	X			X	X
Self-monitoring	X				X				X		X	
Surveillance	X			X			X					X
Conditioning	X			X								X
cause-effect		X				X		X	X			
Virtual rehearsal		X			X			X	X			
Rewards		X		X							X	
Real-world cont		X			X		X	X	X	X	X	
Attractivenes			X			X				X		
Similarity			X		X					X	X	
Praise			X	X								
Reciprocity			X		X					X	X	
Authority			X	X			X					X

Formal Context of HANDS modeled over PT principles, persuasion role, orientation and Hands tools

The Formal Context may then be converted into the lattice below. The nodes in the lattice may be read as follows: The objects (in white rectangles) are associated with all attributes (grey rectangles) above the node featuring the object. Conversely, an attribute applies to all objects below the node featuring the attribute.
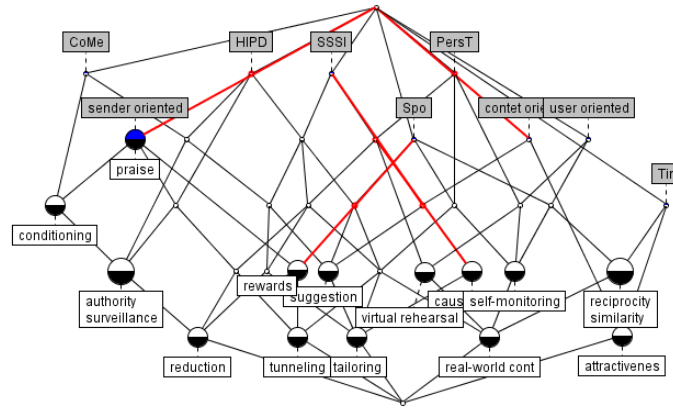
The HANDS Lattice

By selecting a node in the lattice, all paths to the top and bottom of the lattice will be highlighted. Viewing the information this may draws attention to the relations relevant to that particular item. In the lattice below, the Handy, Interactive, Persuasive Diary (HIPD) is selected and the relevant PT principles stand outwhile the remaining part of the lattice is dimmed.
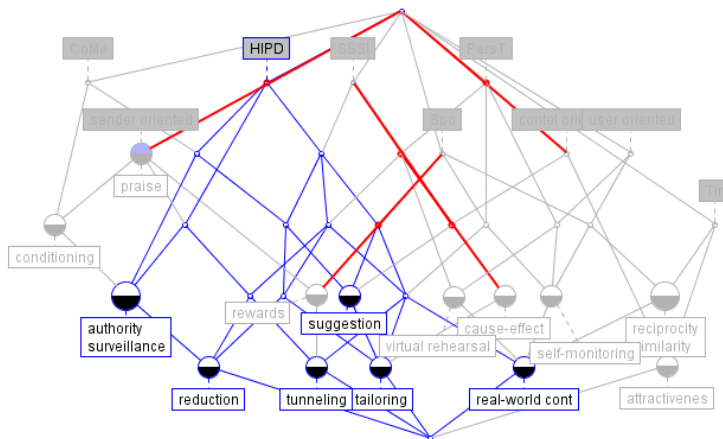
Selection of HIPD

Given the central position in the HANDS tool set it is not surprising that the HIPD is supposed to incorporate as many as 7 PT-principles. Using the lattice this way may serve to generate checklists to ensure that the PT-principles that are assumed to be relevant are indeed investigates.

Since the lattice allows for any node to be selected, it may also be used to reveal relationships between parts of the larger picture. Below, is the rendition of all paths above and below the node containing the PT-principle Reduction.
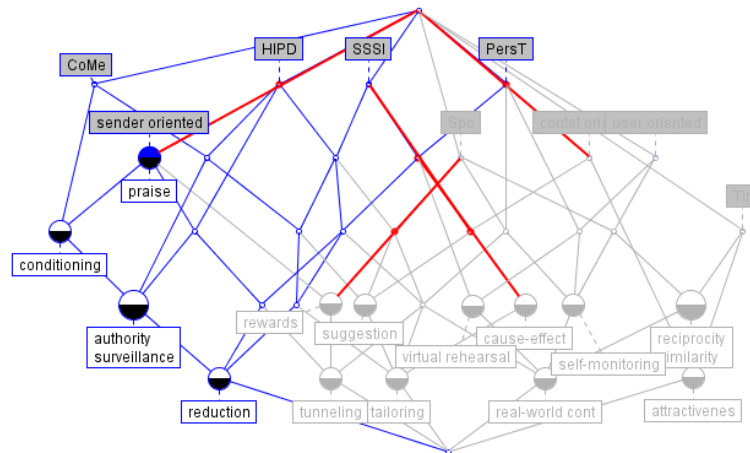
Selection of Reduction

It is easily seen that Reduction is thought to be relevant to most parts of the tool set: CoMe, HIPD, SSSI, and the PersT. The principle of Reduction is also believed to be oriented towards the sender: in order for the principle to function, a given context or scenario must be reduced properly by the sender of the communication. More interestingly, perhaps, is the observation that Reduction only occurs in conjunction with Praise, Conditioning, Authority, and Surveillance (these four PT-principles appear above Reduction in the lattice). Please note that this dependency not necessarily pertain to the PT-principles in general; the lattice only displays the interrelations between elements within the HANDS context.

5. KINDS OF EVALUATION

In order to evaluate the persuasive potential and effect of the tool set, we will consider three sets of techniques:

1. Log files harvested throughout the test period. This area has design implication and ethical implications.
2. Qualitative approaches, in particular observations of use and semi-structured interviews.
3. Quantitative approaches (explicitly harvested), in particular self-evaluation

All areas are heavily subjects to interpretation.

The subject of evaluation is also three-fold:

1. The teacher's interaction with technology and the ways in which this affects their interactions with the students
2. The student's interaction with technology and the ways in which this affects their interaction with their world
3. The context in which the text are carried out. This includes evaluating the actual devices and the qualities of the interfaces over time.

Finally we are especially interested in short and long term evaluation.

5.1 EVALUATION THROUGH LOG-FILES

Essential question here are frequency, duration, and location (for the handheld devices) of actual use. This yields at least the following questions in relation to the teachers:

- Which functionalities have been used?
- Which have not?
- Are the answers to the two above questions surprising?
- Which functionalities have proven useful?
- Have some of these rendered themselves superfluous? (e.g. what is the lifetime expectancy of a SSSI?)

And to the students:

- The same as above +
- Which functionalities were used? Where and when did the use take place?

The latter question requires gps tracking of the devices, which obviously requires permissions from the users. The results, in turn, may be used to validate the principles of suggestion (at the right time), self-monitoring (at the right time), and simulations in real-world contexts. Similar studies concerning the place of use have been used to shed light on cultural aspects of mobile media. One such study revealed that almost all use of mobile phones in Japan takes place within the home as opposed to in public places such as commuter trains.

5.2 QUALITATIVE APPROACHES

There are several reasons for placing emphasis on qualitative approaches. Observations are useful to circumvent the say-do problem: even among very skilled users of technology it is common to find a gap between what users actually do and what they say that they do. This is not a matter of trustworthiness, but rather a normal condition surrounding the use of advanced technology. In this case, the situation is furthermore complicated by the fact that although the teachers are very experienced users of such technologies, they should not be expected to be experts in persuasive techniques. In a sense this is about software prototyping in the second power. We are not just looking for usability traits, but for manifestations of PT principles.

When it comes to interpreting the results from the initial analysis of the logs, the teachers play an even more crucial role. Not only are the teachers domain experts and as such vital to triangulating the results, they are in fact an extremely important key to our understanding of the relationship between students and technology. Inasmuch as it is possible, observations and semi-structured interviews should also be conducted with the students. The physical surroundings of these activities must be negotiated between the teachers and the PT experts, but are likely to take place at school. This entails the problem of authenticity: it is often not feasible to in fact observe the use of technology where and when it is actually taking place, precisely because some of the functionalities are meant to be used in specific contexts in order for the student to remain calm and in control (e.g. the SSSI). These situations are to a large extend not predictable in the sense that observation can be scheduled. Other functionalities are meant to be used when the student is alone (e.g. the Personal Trainer) and here again, an observer would risk to disturb the actual use.

Interviews may benefit from the analysis of the logs, and should be semi-structured. The preliminary investigations as well as the habitus of the teachers all suggest that open-ended interviews are the most suitable for the job.

5.3 QUANTITATIVE APPROACHES

In addition to logging movement and other usage, a few other quantitative methods may be used in evaluating the persuasive effect the HANDS tool set. Logs are per definition covert, and although the existence of this technology is overtly stated and discussed in the information given to teachers, students, and parents, we tend to not think about logging as it occurs. In contrast, both teachers and students are already used to evaluate daily routines in some detail (scoreboards etc.) and it seems natural to incorporate this activity into the HANDS tool set.

Here, again, the teachers are vital in interpreting the results of self-evaluation. In fact, the role of the teachers can hardly be overestimated in evaluating persuasive effect with young students with autism diagnosis.

6. STEPWISE EVALUATION

The evaluation design must take the step-wise introduction of the HANDS tools into account for two important reasons:

1) The proposed functionalities of the tool set are most likely too many and too diverse to be implemented at once. This goes for the students as well as for the teachers, and is one of the reasons that we propose to emphasize the use of HIPD and SSSI in the first prototype.

2) There are bound to be significant differences between prototype 1 and prototype 2. A number of more advanced functions may be implemented in the second version, and may indeed come about as a result of the first round.

The implication is that the two rounds of testing and evaluation cannot be directly compared.

REGARDING PROTOTYPE 1

The first round of evaluation will naturally focus on matters regarding implementing the technology. In addition to testing factual matters such as use and response to use, this evaluation must also focus on what happens when you introduce new technology into the lives, especially of the students. As mentioned in section 2.1 the schools have different practices regarding the use of mobile phones prior to the HANDS project. It is therefore reasonable to assume that varies concerns related the common use of such technology will arise. It is essential that such issues are addressed appropriately in the evaluation at each school. For example, in cases where there is no existing practice with respect to mobile phones, it is likely that unintended use will occur, perhaps even usage that will disturb teaching. The mere presence of technological gadgets calls for attention, especially when it is relatively new, and introduced into environments of young people. There may also be very positive effects of introducing such technology, i.e., eagerness to share and compare uses of the devices. The news-value will inevitably wear off, at which point a more regular use may be observed. It is expected that these effects will differ from school to school. With respect to evaluation, three concerns arise.

1. Existing ICT practice should be accounted for as clearly as possible. Special attention should be given to areas where the phones may replace existing tools. The phones may for instance replace digital cameras and MP3 players and radios. This is not, however, necessarily the best solution.

2. Early stages of use may be infused with exaggerated enthusiasm and / or frustration. Accounts of this should be taken into account in the evaluation, but not necessarily attributed to the HANDS toolset.

3. As the news-value recedes, the more "real" use will emerge. This would typically imply a reduced use of certain features and an increased use of certain other features. To some extent, this should not necessarily be attributed to the HANDS toolset.

REGARDING PROTOTYPE 2

When the second prototype is to be evaluated, issues of implementation will be different. In the first place it is different because it is easier to integrate an upgraded version of something than to implement completely new solutions. And in the second place it is different because all parties involved have already been through the cycle once. It is reasonable to assume that both teachers and students will experience a higher level of reflective awareness regarding the technological solutions in the second version, simply because they have done it before, and at this time it is easier to build communicative bridges between developers, users and evaluators.

Since prototype 2 is not evaluated with direct improvement in mind, the situation is a little bit different. Suggestions and complaints in prototype 1 are likely to be addressed in the redesign of the software. This is not the case in prototype 2. Here, all suggestions and recommendations are given and collected for the purpose of other future projects.

7. PLAN OF EVALUATING

The evaluation of prototype 1 serves a dual purpose. In the first place, the evaluation is designed to investigate persuasive effect and potential. And in the second place, this evaluation is designed to systematically collect experiences in order to expand the functionality of the toolset in prototype 2.

A similar outlook can be said to apply to the evaluation of prototype 2. As well as investigating and verifying the persuasive effects of the second prototype, this round of evaluation is aimed at extracting design principles that may extend beyond the project, and into future research and development in comparable projects.

As indicated in the sections above, a natural flow of investigation will take its point of departure in the actual use in the HANDS toolset, and informed by these quantitative date move into qualitative methods in order to thoroughly interpret the data. The system log-files will deliver information about the teacher's use of the system as well as of the student's use. Since many important elements in the information ecology escapes quantification it is essential that the interview guides are not solely crafted from the hard data, but inspired by it. Social stories that are not longer in use may have achieved their goal, or they may not be used because of flaws in their making. Quantitative data will reveal this. Another interesting issue that can only be evaluated by qualitative methods, is the question of which elements are best served in digital form, which in analog form, and which in a combination of the two.

The evaluation will follow these steps:

------- Installment of Prototype 1 --------

- The log-files are collected and initially sorted.

- The logs are investigated for indications of aspects of persuasion (section 2 in this document), and with respect to the persuasive principles and their interrelations (section 3 and 4 in this document).

- Interview guides for semi-structured interviews are prepared based on these findings.

- Teachers are interviewed, and if possible and appropriate, observations are conducted.

- Students are interviewed, and if possible and appropriate, observations are conducted.

- Interviews are analyzed and triangulated with observations and logs.

- Results from interviews are verified by the interviewees in condensed form.

- The final evaluation report is written. This includes recommendations for prototype 2

------- Installment of Prototype 2 --------

- The log-files are collected and initially sorted. Special attention is given to changes in previous patterns.

- The logs are investigated for indications of aspects of persuasion (section 2 in this document), and with respect to the persuasive principles and their interrelations (section 3 and 4 in this document).

- Interview guides for semi-structured interviews are prepared based on these findings.

- Teachers are interviewed, and if possible and appropriate, observations are conducted.

- Students are interviewed, and if possible and appropriate, observations are conducted.

- Interviews are analyzed and triangulated with observations and logs.

- Results from interviews are verified by the interviewees in condensed form.

- The final evaluation report is written. This includes recommendations for future research and development.

During the spring of 2009, test protocols will de be crafted and presented to the Ethical Board for approval.

RESOURCES

The evaluation of the persuasive effect and potential will avail itself of log-files from all four schools. On sight Interviews and observations are conducted at Egebakken (Denmark) and Svedanskolan (Sweden), but not at the Helen Allison School and Autism Foundation. It is likely that the log-files and the interviews give rise to questions concerning the schools in the UK and in Hungary as well. To the extent that it is feasible, these matters will be settled through online communication with respect to prototype 1. On sight interview sessions will have a duration of approximately 1 hour per session. It should be noted that the validation of data collected through interviews also takes time.

The workload for the interviews for the teachers thus amounts to roughly 1.5 hours per child participating in the test. In prototype 1 this applies to Egebakken and Svedenskolan. In prototype 2, this applies to all four schools. Since observation should be as noninvasive as at all possible, it is required that the respective teachers assist in these sessions. It is also crucial that teachers assist in interpreting the data. Relatively few observations are needed, and the workload of the teachers is estimated to roughly 10 hours per school. In prototype 1 this applies to Egebakken and Svedenskolan. In prototype 2, this applies to all four schools. In addition, it is crucial that PD researchers to some extent will have access to minor discussions with the teachers, primarily online. The estimated workload will be relatively small, roughly in the area of 5 hours per teacher throughout the process.

Estimate of workload:

	Prototype 1	Prototype 2
Egebakken	30	30
Svedenskolan	30	30
Helen Allison School	15	30
Autism Foundation	15	30

8. SUMMARIZING PRINCIPLES FOR EVALUATING PT

The concerns and principles covered in this report can be summarized in the following way:

With respect to the organizations in which the tests are conducted the following must be taken into account:

- Elements of Indirect Coercion
- Variations in set-up.

With respect to the persuasive situation it should be considered whether:

- The intended outcome is:
  - Response-Shaping
  - Response-Reinforcing
  - Response-Changing

With respect to the actual use (by all stakeholders), it should be considered whether:

- The actual use is:
  - Partial Use
  - Derived Use
  - Alternative Use

With respect to the elements in the tool set it should be considered to what degree the elements (and their interrelations) fulfill their roles as:

- Persuasive Roles of PT
  - Tool
  - Social Actor
  - Simulation

To evaluate these aspects of Persuasive Technology in the context of HANDS, we place emphasis on:

- Log files (who, what, when, how often, and in some cases where)
- Qualitative approaches (observations and semi-structured interviews)
- Quantitative approaches (self-evaluation and in some cases questionnaires)

To evaluate the effect of HANDS with respect to the information ecology, we place emphasis on:

- Existing practices brought into the HANDS toolset
- Existing practices NOT brought into the HANDS toolset
- New practices brought about by the introduction of HANDS

REFERENCES

Burmeister, P. (2003). Formal Concept Analysis with ConImp: Introduction to the Basic Features, Available at: http://www.mathematik.tu-darmstadt.de/~burmeister/ConImpIntro.pdf.

Flyvbjerg, B. (1998). Rationality and Power: Democracy in Practice, University of Chicago Press.

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

Fogg, B. J., E. Lee, et al. (2002). Interactive Technology and Persuasion. The Persuasion Handbook: Development in theory and practice. J. P. Dillard and M. Pfau, Sage Publications: 765-788.

Ganter, B. and R. Wille (1997). Applied Lattice Theory: Formal Concept Analysis, Available at: http://www.math.tu-dresden.de/~ganter/psfiles/concept.ps.

Hasle, P. (2006). The Persuasive Expansion - Rhetoric, Information Architecture, and Conceptual Structures. Conceptual Structures: Inspiration and Application. H. Scharfe, P. Hitzler and P. Øhrstrøm. Berlin-Heidelberg, Springer Verlag: 2-21.

Hasle, P. and A.-K. K. Christensen (2008). Persuasive Design. Handbook of Research on Computer-Mediated Communication. S. Kelsey and K. St. Amant. Hershey, USA, IGI Publishing Group: 447-498.

Miller, G. R. (1980). On Being Persuaded: Some Basic Distinctions. Persusion: New Directions in Theory and Practice. M. Roloff and G. R. Miller, Sage Publications.

Oinas-Kukkonen, H. and M. Harjumaa (2008). A Systematic Framework for Designing and Evaluating Persuasive Systems. Persuasive Technology. Third International Conference, Oulu, Finland. H. Oinas-Kukkonen, P. Hasle, M. Harjumaa, K. Segerståhl and P. Øhrstrøm, Springer: 164-176.

Scharfe, H., P. Hitzler, et al., Eds. (2006). Conceptual Structures: Inspiration and Application. 14th Conference on Conceptual Structures, ICCS2006. Aalborg Denmark, Springer Verlag.

Sowa, J. (2001). Mathematical Background, http://users.bestweb.net/~sowa/misc/mathw.htm.

Yevtushenko, S. (2003). ConExp, Available from: http://sourceforge.net/projects/conexp.