Some Severe Deficiencies of the Input-output HCI-paradigm and Their Influence on Practical Design

Kai Tuuri
Department of Computer Science and Information Systems
FI-40014 University of Jyväskylä, Finland
krtuuri@jyu.fi

Antti Pirhonen
Department of Computer Science and Information Systems
FI-40014 University of Jyväskylä, Finland
pianta@jyu.fi

Eve Hoggan
Glasgow Interactive Systems Group, Department of Computing Science
University of Glasgow, Glasgow, G12 8QQ, UK
eve@dcs.gla.ac.uk

ABSTRACT

The Cartesian dichotomy of human mind and body has largely ruled the development of western thought. One effect of that Cartesian legacy is the tendency to conceive interaction between a user and smart device as being composed of different inputs and outputs. In many cases, this is a practical and highly appropriate approach to design interactive technology. We, however, argue that such an approach tends to put too much emphasis on the technical instrumentation that provides information for different senses, thus considering sensory modalities as independent 'receiver modules'. Perception is not directly created on the basis of the physical origin of the sensation. Rather, we argue that it is based on sensory-motorically integrated gestalts. For instance, a haptic feedback experience can even take place in the presence of only visual or audio cues that become coupled with interaction [1, 2]. If the concept of haptic feedback is merely understood in terms of the sense of touch and its usage with the help of, e.g. force actuator technology, it could be argued that the choice of options in user interface design is severely narrowed, and may result in the inappropriate use of available technology. By discussing the design of haptic feedback for touch screen applications, this paper illustrates the deficiencies of the input-output paradigm. It also stresses the close coupling between perception and action, which is realised in the course of interaction in a way that does not justify splitting them conceptually when striving towards a deeper understanding about human-computer interaction.

Keywords

interaction design, haptics, pseudo-haptics, multimodal interaction, embodied cognition

ACM Classification Keywords

H.5.2 [**Information Interfaces and Presentation (e.g., HCI)**] User Interfaces – *theory and methods*.

ECCE 2009, September 30 – October 2, 2009, Helsinki, Finland.

INTRODUCTION

To understand the current paradigm of human-computer interaction (HCI), it is beneficial to look at the origin of the discipline. In the era of centralised computers, interaction with them was a highly professional activity, and people dealing with computers could be educated to cope with them. However, as soon as computers were spread to the hands of everyone, it became essential to pay attention to interaction with computers. The aim was to make a computer and its user a seamless whole [3]. In the conceptualisation of HCI, basic concepts were largely derived from a technical context, paralleling the function of the computer and its user. Thus, the whole paradigm of HCI is often seen as constituting inputs and outputs between a user and a computer.

Under the umbrella of the *cognitivist* paradigm, the adoption of the computer metaphor of human cognition [4] in HCI has played an important role in the development of the discipline ever since. The computer metaphor was easily adopted on the basis of the traditional Cartesian dichotomy of mind and body, which has ruled the development of western though for centuries. Although cognitivism has been gradually rejected by researchers of mainstream cognitive science [5], this recent development of cognitive science and the philosophy of mind has never really overdriven the temptingly simple and easy-to-understand computer metaphor in HCI. In human computer studies, the influence of the computer metaphor is salient. Even in today's vocabulary of HCI practitioners, interaction is largely conceptualised in terms of technical devices which represent input and output modalities of interaction. The term *multimodality* thus often means the use of different technical devices, as channels of information transmission [6], in the design of HCI applications. For example, visual displays, speakers and motion actuators for corresponding output channels, or keyboards, microphones and motion sensors for corresponding input channels. Such an approach is indeed appealingly practical, and it makes the analysis and development of HCI applications straightforward.

As long as HCI mainly meant working with a desktop computer by entering information with a keyboard, and receiving information through a visual display unit, the input-output paradigm was not questioned. However, the addition of computational power to various products in our everyday life has not only broadened the scope of HCI but also challenged the appropriateness of this model of interaction. New technologies and use cultures opened a new perspective to HCI, whether it is a question of traditional desktop setting or some newer concept of utilising digital technology.

The ongoing shift towards user-centered design merits an investigation of the appropriateness of treating HCI in terms of the technology employed in interaction. From the human point-of-view, our interaction modalities are not channels of information transmission. None of the contemporary models of human cognition suggest that we have separate sensory and motor systems for each available user-interface technology. Even though we can see screen events with our eyes, our minds do not handle the visual perception in isolation from other cognitive processes. Rather, recent studies in neuroscience suggest that our cognitive processes are based on extensive sensory-motor integrations at a neural level [7].

In this paper, we challenge the prevailing HCI input-output paradigm, by indicating clear deficiencies in its appropriateness for the conceptualisation of multimodal interaction. We present examples which illustrate design issues at a practical level, and relate these to the traditional HCI paradigm. We also discuss alternative ways to conceive multimodal interaction in terms of embodied cognition.

SAMPLE DESIGN: A VIRTUAL SLIDER

For about five years, we have conducted research projects in which the potential of non-speech sounds in HCI has been studied[22]. Recently, we have broadened the scope to cover haptics as well. The projects have been carried out in close cooperation with the Finnish manufacturing industry. The application areas have been extremely diverse; from wrist computers to control room technology, from tiny portable applications to heavy industrial machinery. The problems and focus have varied, but there have also been a great deal of common factors among the different applications and contexts of use. In this paper, we do not present any of these industrial cases in detail, but present an illustrative sample about the issues of conceptual design for a 'virtual slider' for touch screen applications. Although the discussed sample design is closely related to an ongoing case study, it also draws upon past experiences concerning many kinds of products of our partners.

Background

Touch screens are an interesting piece of interaction technology, in many senses. Their obvious advantage over mediated pointing devices, like the mouse and roller-ball, is that users can physically point at objects in graphical user-interface (GUI). Touch screens have also been successfully used without visual displays, utilising their touch sensitivity in mobile applications by substituting the visual display with an auditory one (see, e.g. [8]).

GUIs, especially when implemented in touch screen, provide a tempting opportunity to substitute tangible hardware with similar looking simulations. From the very beginning of GUIs, the appearance of basic tools has imitated their real counterparts. Word processing applications, for instance, have inherited much of what they look like and even the vocabulary, from mechanical typewriters. Sometimes, this strategy of substituting a real-world object with a somewhat similar digital application, are referred to as metaphors. It is more appropriate, however, to talk about simulation or imitation [9].

In terms of usability, the rule-of-thumb is that each function has a dedicated, hardware based control [10]. Virtualised controls, or virtual substitutes, can hardly ever be better than the real-world counterparts. However, when reasoning beyond pure substitution, virtualised controls have advantages. First of all, in mass products in particular, their production is practically free of cost. This is probably the foremost motivation for switching from hardware based UIs to software based. Second, the virtual control can provide levels of flexibility that a real thing cannot. For instance, the UI is no longer dependent on controlling schemes based on hardware controllers. It can also be changed in terms of personal preferences or use mode.

The clearest disadvantage of virtual control elements, compared with hardware-based ones, is the lack of physical interaction with the ongoing process. For instance, with a physical slider users can feel the current position and the resistance of the slider as well as adjust it accurately. An ideal slider would combine the strengths of the physical and virtual. The design challenge of a virtual slider would thus be how to reproduce the physical experience of moving a slider. In the application which was the foremost motivation for this discussion, it was also necessary to provide such an unambiguous experience, that the slider could be adjusted without looking at it.

Desired Feedback Technology vs. Desired Interaction Experience

In interaction design, expertise from various disciplines is necessary. We have been in many brainstorming sessions, in which novel interaction ideas have been discussed among the experts of technology, human behaviour and the experts of context of use. In such sessions, visions are encouraged to be freely expressed. However, in concrete visions, technical issues tend to be the driving force.

The traditional input-output paradigm is a perfect match to the technical approach to design. In order to cause a haptic experience, some tangible, force-emitting elements are needed in the user interface. Such apparatus can thus represent the channel through which haptic information is provided to user. In the case of touch screens, a screen with an enormous number of up and down moving pins might represent the ideal scenario in enabling the haptic feedback. The issues which an engineer would be mostly interested in, would be the number of pins per inch, the available force, range and speed of movement of each pin, and the latency in milliseconds. On the basis of these specifications, the appropriateness of the technology in the given context could then be assessed.

Since it is not currently economically realistic to implement the display described above, designers must usually conform to some kind of 'light' version of the haptic feedback technology, such as simple vibro-tactile actuators. However, the motivation for going for such a technology, which has very limited resolution for presenting haptic information, has not been any model of human perception, but purely economical. The initial idea is just reduced to a level which is feasible to implement. That underlying idea is to physically cause events which can be perceived by the sense of touch.

In the case of the virtual slider, feedback which could somehow simulate the touch-related properties of real physical interaction would certainly be beneficial. But must we confine our design considerations to the seemingly obvious 'haptic technology' for information presentation? While not having the credible technology to implement physical events other than visual and audio, we were thus forced to consider our design task the other way round. We can either strive to use motorised tactile/haptic effects as part of the user-interface, or we can try to provide an appropriate overall experience for interaction with the available means. If we choose the first one, we will need to apply the given technology as appropriately as

[22] GEAR, GEAR2 and GEAR3 projects funded by the Finnish Funding Agency for Technology and Innovation (TEKES).

possible. It cannot be assumed, however, that using e.g. vibro-tactile actuators would result in desired interaction-relevant haptic experience. Therefore, the preferable option is to start the considerations from the other end, the desired experience. Only after having specified the qualities of the desired experience, should the technical issues be considered.

Designing Pseudo-haptics

There are a number of studies which show that designers do not need to be limited to existing, fairly primitive actuator technologies when aiming to provide a haptic experience. For instance, sounds as well as visuals have the potential to produce a haptic perception, as the coupling between a user's motor-movements in touch screen interaction and the related sonic/visual feedback can result in *pseudo-haptic* [1] illusions. Thus, auditory or visual sensations, as direct feedback for actions of the user, can modulate touch-related perceptions about the physical features of an object, like surface properties [11, 12, 13] or inertial properties such as resistance to motion [2]. The designer has the option to consider, for example, what kind of sonic feedback would modulate the haptic perception towards the desired experience, i.e., how it should sound when the screen surface is touched.

Touch screen interaction is an example of direct manipulation, in which feedback techniques can give touch-related impressions that some kind of *surrogate object* is interacted with instead of touch screen surface itself. We started conceptualising the virtual slider by exploring the propositional flow of interaction: what is being controlled with the slider and what kind of feedback experiences would be desired for such activity? In other words, we needed to determine what kind of surrogate object the user would perceive while using the virtual slider – how it should feel, and how that 'object' would function as a controller. Using the ecological approach to perception [14], we can thus utilise the directly meaningful attributes of the active object – meanings which are engaged with the user's purposeful actions, and which are based on the experiential history of interacting with the environment.

The surrogate virtual object can be directly based on the subject matter which is being manipulated via the controller. When controlling operations remotely, such as rock drilling, physical invariants of the actual drilling operation can be used as feedback cues – thus supporting feedback experiences analogous to the non-remote drilling. However, as a surrogate object, we can also utilise many kinds of virtual tool-objects which are based on a suitable metaphor that matches the real situation of manipulation. Moreover, we do not have to fully simulate any real object. We can just use certain *action-relevant* attributes of active objects. These attributions are action-relevant because they propose the occurrence or afforded potential of some activity, and are also indexed to contextual activity.

Let us assume that the function of a virtual slider is related to the manipulation of a parameter, which is to be increased until the predetermined threshold-value is reached. For such a scheme, we can use, for example, the action 'rotating an increasingly resisting wheel clockwise until it snaps into a retainer that holds it' as a model for determining what kind of feedback, mapped to the circular movement of a finger on touch screen, would relate to sensations of appropriate resistance and finally the removal of the restraint. We do not have to accurately re-create such a wheel, but use it as a propositional mental model for conceptualising the functionality of the controller. Based on that functionality of the 'wheel' and the adequate modelling of its inertial and material properties, we can discover invariant attributes that can be utilised as action-relevant feedback cues. As physical invariants, such action-relevant cues can be potentially encoded into various modes of presentation; vibrations, sound or visuals.

In the case of the virtual slider, the general design intention was to provide a feedback experience that would indicate functional 'resisting friction' during the control process. This is possible, for example, by attributing surface roughness as a continuous, action-relevant parameter to the feedback of virtual slider. The current design is still in a conceptual phase, in which different touch screen widgets and corresponding interaction schemes or surrogate tool-objects are being considered. Pseudo-haptic feedback has offered a viable substitution to the use of force actuator technology. Because the use context of the virtual slider does not allow the gaze of the user to be focused on display all the time, audio is clearly the preferred form of feedback over visual. Audio feedback has been already proved to be effective in providing and modulating haptic experiences [2, 11, 12].

Haptics as a Perceptual System

Although the construction of the virtual slider is still in an early stage, we have already learned a lesson or two about interaction when going through different options for design approaches and technical issues. First of all, we needed to critically examine the current interaction paradigm, in which interaction is usually seen as a combination of inputs and outputs. If the hypothesis about pseudo-haptic perception holds and perceived haptic feedback can be based on sounds only, it is difficult to justify the current divide to input and output. From the point of view of the user, what is 'input' if it is physically a sound but results in a haptic experience? Moreover, in the case of haptic experiences, the user's exploratory activity is the key element in 'picking up' the information about, for example, the roughness of the surface. In other words, this kind of haptic experience is *obtained* through the active role of the user in interaction. This issue further undermines the input-output dichotomy. The finger, which is moved around the touch screen surface, could be seen as a *perceptual subsystem* because it participates in the exploratory effort of obtaining the information [15]. So, when the perception of the information also requires muscle activity and body movement, it is not easy anymore, or even appropriate, to make a distinction of which is 'input' and which is 'output'.

Of course, these presented issues, are not new findings at all. In the mid 1960's, J. J. Gibson [15] discussed them in his book 'The senses considered as perceptual systems'. He asserts that inputs for perception are not equivalent to inputs for sensation. A haptic experience is not merely based on the sense of skin pressure, or even the sum of the sensations of skin pressure and kinaesthesia. Gibson sees haptic systems as a much broader apparatus, by which the subject picks up information about both the environment and his body. He admitted that 'the simple, neat easily-remembered contrast between *receptors* and *effectors*, between *sensory* and *motor*, will have to be abandoned.' Despite this, he did not completely reject the input-output dichotomy, but tried to reformulate it to take the complexities of sensory-motor integration into an account. However, more recent, related accounts of *enactive perception* [16] and *embodied simulations* [7] have further blurred the boundaries between input and output (or perception and action), by arguing that all perception is being intrinsically 'acted out', i.e., the acquiring of perceptual experience requires the perceiver's skilful activity. Recent findings in neuroscience support these approaches by highlighting the integration of perception and motor action at the neural level [7].

The concept of interaction modalities is another issue, closely related to the discussion of this section. When talking about multimodal interaction, the design example of a virtual slider forces the question, what kind of interaction is *not* multimodal? We next consider the themes which became topical in the design of virtual slider: the nature of multimodal interaction, and multimodal approach to interaction design.

THE NATURE OF MULTIMODAL INTERACTION

In the bulk of HCI literature, interaction modalities are handled as 'channels' between human cognition and the environment. Whether we talk about channels or pipelines, which is another popular, related metaphor, it is undeniable, that the underlying concept is closely related to the computer metaphor (see introduction). In it, our senses and motor-systems are conceived as being directly and independently connected to the metaphorical central processing unit.

In the simple pipeline metaphor, different modalities are handled as independent resources. This basic assumption can be seen in numerous HCI studies. For example, the problem of human computer interaction in mobile contexts is typically 'solved' on the basis of pipeline metaphor. While on the move, users are often unable to fully focus their visual sense on the mobile device when using it. Also, mobile devices are usually so small in size that the interaction cannot be based on a tiny screen alone. Due to these limitations, more 'bandwidth' for human computer interaction is searched for from other modalities than vision. Hearing and touch, for instance, have been identified as free resources and therefore excellent opportunities for enhancing interaction with a mobile device.

However, it has been shown for quite some time, that interaction modalities are far from independent from each other. Extensive research work in this area has taken place in attention studies since 1950's. The development of theories of human attention has started from simple pipeline models [17] gradually to more and more complicated theories.

The development process of the models of attention is typical in the history of science: we first have a simple theory, which is easy to adopt. When we gradually get results of empirical studies which are in conflict with the model, the model is revised to match the empirical evidence. As a result, the revised model matches better with the existing evidence, but is probably more complicated than the previous version. After a number of this kind of revisions, the theory is so complicated that it is hard to communicate to the scientific community which is supposed to apply and further develop it [18].

In the models of multimodal human-computer interaction, the idea of independent interaction channels has been challenged; not only by previous attention studies, but also by user studies with e.g. mobile applications. A prominent example is the use of mobile phones while driving. The channel metaphor suggested that the use of hands free equipment would solve the safety problem of talking on the phone on the road; since the driver does not need to talk or listen in order to control a car, those modalities were declared as free resources which could be used for phone calls. While using hands free equipment instead of a handset, the same resources for driving were supposed to be available whether there was a phone call on or not. However, the observed deteriorating of driving performance while talking in the phone, whether using a handset or hands free equipment [19], questions the channel model. Apparently, concentrating in conversation is distracting regardless of the available telephone user-interface.

The above example is obviously closely related to the studies about human attention, the paradigm of divided attention in particular. It appears, however, that the cognitivist approach used in such studies has quite limited power to support the design of multimodal interaction. The recent studies in *embodied cognition* are a promising option for conceptualising interaction. In the embodied approach [5], the human mind is inseparable from the sensory-motor experiencing of the physical world and cognition is thus best described in terms of embodied interaction with the world. Human understanding is thus seen as arising inherently from embodied couplings with the environment. The emergence of these couplings is based on an experiential background of constant encounters and interaction with the world by using our bodies. Quite in the same line with the ideas of ecological perception, the embodied approach stresses an action-based understanding of the world, (see e.g. [14, 5, 7, 20]). From such an action-orientated perspective, multimodality and the multimodal experience appear as an inseparable characteristics of interaction. That is because the understanding of actions, as a subjective experience, is not characterised necessarily by sonic, visual or tactile qualities. In fact, actions presented in different modalities seem to produce a very similar neural basis for the understanding of an action [21]. This agrees with other findings of extensive sensory-motor integrations in neural mechanisms, which are hypothesised to contribute as a basis for action understanding [7].

The viewpoint of embodied action understanding further challenges the existence of isolated domains for sensory input and motor output. We must presume that perception is not directly created on the basis of the physical origin of the sensation. Rather, we argue that it is based on *sensory-motorically integrated gestalts*[23], which are based on intrinsic human capabilities of action understanding and recurrent patterns of embodied experiences of interacting with the environment Such multimodal gestalts would indeed explain the phenomenon of pseudo-haptics, i.e., why a stimulus in auditory or visual sensory modality has the potential to effectively modulate haptic perception. It would also explain why concurrent stimuli in different modes of presentation result in a fused perceptual content [23], or why even stimuli of a single type of modality, for example music, can trigger multimodal completions that include imagery of body kinaesthesia, touch and even visuals [24]. Such a gestalt completion results in multimodal experiences, even when the stimulus is unimodal. In a sense this could be called *amodal* completion. But unlike the traditional cognitivist view [25], we propose that amodality is *not* symbolic in nature but inseparably bound with our sensory-motor system. Thus the resulting meaningful experience is not seen here as modality independent, but essentially as engaged with the human body and all of its modalities of interaction. Therefore it might make sense to call these mental completions multimodal, rather than 'amodal'.

On the basis of what we have discussed so far, it appears that all interaction is inherently multimodal. Therefore multimodality is not something that designers can implement in applications. Rather, multimodality is the nature of interaction that designers must take as a starting point and acknowledge the embodied situation as a whole. Hence, such a multimodal approach to interaction design [26] should not focus on different presentation (or input) modalities or on any communication technologies in themselves.

[23] Such gestalts refer to the theories of, for instance, embodied image schemas [22, 20] and embodied simulations [7].

THE CONCEPT OF CROSSMODAL INTERCHANGEABILITY

One traditional rationale for using multiple modalities in user interfaces is to provide choices for the user. It has been proposed that when a message is coded in the user interface in multiple forms, the user should be given the opportunity to receive the message in a form most suitable for her given the situation or context [27]. It has been argued that, by allowing the user to choose from a selection of modalities, different capabilities or cognitive styles can be taken into account. In this kind of design strategy, redundancy is a central concept: information is presented redundantly in various communication channels in order to assure a correct reception of a message.

The strategy described above, often presupposes an unambiguous segmentation of form and the content. In other words, this presupposition means that there would be interchangeable interaction modalities, i.e., a given piece of information should be able to be presented equally in different modalities. For example, rhythms can be created identically in both audio and vibrotactile feedback. Previous work in crossmodal transfer has shown that, if trained to understand multidimensional audio alerts, a user can then also understand the corresponding tactile alerts with no additional training and *vice versa* [27]. In this case the mappings between information and feedback are abstract and have to be learned.

In some other real-world situations, crossmodal presentation of information is also somewhat successful; for instance, in pedestrian traffic lights in which the message is either ‘go’ or ‘don’t go’, this simple message can probably quite appropriately be argued to be communicated equally by a colour, and an icon, frequently by a sound as well (at least in Scandinavian style lights). However, it is as well common sense as a result of theoretical reasoning, that different forms of presentation deliver qualitatively different kind of information [28]. Therefore, the idea that modalities would be universally interchangeable is perhaps not so appropriate in certain situations. In the example of traffic lights, the encoding and decoding of the messages requires a great deal of conventionality and habit-based knowledge. Without that backlog of knowledge, in the case of the traffic lights, it could be hard to describe what a ‘red colour’, a ‘figure of a standing man’ and ‘short, repetitive beeps’ have in common. As mentioned in the previous work, to use crossmodal feedback effectively, some user training is required [27]. This, in particular, applies to symbolic/arbitrary mappings between information and feedback. However, the utilisation of *ecolocigally valid* mappings, designed to agree with our existing experiential knowledge, should diminish the need for training.

The underlying idea of interchangeable modalities is that there first exists an amodal meaning. Communication of such a meaning would then involve the implementation of certain amodal, i.e., modality independent attributes in the chosen presentation modality. The traditional cognitivist approach considers amodal information as essentially symbolic [25], thus implying an arbitrary relationship between amodal meaning and its form of expression. However, in the light of embodied approach to human cognition and the recent neuroscientific findings, such a view on amodality has been strongly criticised [7]. Designers can successfully utilise symbolic attributes (such as information encoded in rhythms) in the design of crossmodal feedback [27], but it does not mean that crossmodal interchangeability (e.g., crossmodal recognition of rhythms) would be based on symbolic or abstract processes. As discussed briefly in the previous section, we propose that amodality should be seen in relation to embodied sensory-motoric integrations that takes place in the course of perception and thinking. These integrations are realised in neural mechanisms – such as mirror neurons [29] – which have a role in contributing action understanding. For example, crossmodal rhythm perception can arguably be based on motoric attuning to the observed rhythmic impulses [30]. As these enactive perceptual processes should occur in a sensory-motorically integrated substrate [7], it does not matter whether it was the sonic, tactile or visual features of the observed feedback that allowed the ‘motor-mirrored’ amodal understanding of the rhythm. But can we consider such ‘amodal’ understanding modality independent then? Not entirely, because it would still be dependent on the motoric perception and its intergration to the senses.

There is no any reason why we should not consider ‘amodality’ as being engaged to the body and its means to interact with the environment. For example, in our case of designing pseudo-haptics, we could call the perception of surface roughness or friction amodal. Indeed it seems that the attribution of surface texture can be equally based on touch, sound or visuals. We argue, however, that even in this case the desired meaning is highly dependent on interaction, in particular, the ecological validity of the presented artificial feedback events. Therefore the key element in the haptic experience is the intentional activity of the perceiver, who acquires the information by moving her finger on the touch screen surface. It is still true, though, that the certain invariant properties of a modelled object, which indicate the roughness during the interaction, can be equally manifested through a variety of modalities. But perhaps the next step in crossmodal interaction research is to acknowledge the inevitable dependencies of such crossmodal interchanges on the nature of multimodal interaction. As Gibsonian approaches have already underlined, it is almost impossible to keep perception fully separated from interaction. Thus, the perceptual content is unavoidably affected by our intentional activity.

In terms of embodied cognition, it is difficult to argue that content can be fully separated from its form. According to the embodied perspective, the conception of mind is bound to the physical body, or even beyond [31]. For instance, writing with a pen on paper is not just a reflection of human thought, resulting in words. Rather, the act of writing itself is essential part of both expression and thinking. As an articulatory process, writing also results in nuances of form that cannot be understood in terms of verbal semantics. Some theorists go even further by conceptualising the physical result of e.g. writing, as part of the human mind [31].

CONCLUSION

Interaction between users and devices is traditionally viewed as a combination of inputs and outputs, in accordance with the computer metaphor of the human mind. This paper has suggested that such an approach leads to an emphasis on technical instrumentation with each sensory modality simply viewed as a receiver module. The argument presented in this paper focuses on sensory-motorically integrated gestalts. In other words, there is a close relationship between perception and action, which occurs in the course of interaction and therefore, there is no justification in splitting these into two separate concepts. The case of pseudo haptics highlights this issue because it has been shown that a haptic feedback experience can take place without the presence of physical haptic cues but in the presence of only visual or audio cues that become coupled with interaction. It may even be more appropriate to replace the term *pseudo-haptics* simply with *haptics*, given that the main difference between this and 'real' haptics is related to the available sensory stimuli, not the resulting experience.

In order to demonstrate that our arguments are not just academic 'hairsplitting' about concepts, we wanted to show that the prevailing input-output paradigm has impact on practical designs. Established dispositions towards conceptualising interaction in oversimplified technical terms can effectively hinder the design potential for multimodal interaction, as the sample design of virtual slider revealed.

The underlying idea of this paper is to stress the human-centeredness in the cost of technology. The strength of the input-output paradigm is its consistency with the structure of technical devices. By conceptualising interaction in technical terms, it is easy to analyse a system as a whole consisting of a technical system and its user. From this point of view, the aim of user interface design is easily biased to view the user as a part of the system, and make her behave as effectively as possible in terms of the task. However, different products which have very similar interfaces in terms of usability and functionality can still differ dramatically in terms of use experience [32]. Perhaps, a design approach genuinely focusing on the users point of view, would lead to the successful design of a desired experience – which, of course, does not rule out effective task completion.

The discrepancy between the application of traditional information-centered models of human cognition and how users experience the use of technology has already been widely acknowledged. Recent trends in HCI research stress the role of user experience in understanding HCI, and the importance of the embodied interaction in understanding the human mind [33]. The shift of focus from the design of efficient HCI onto sketching experiences should inevitably change the way we conceptualise interaction. By questioning the boundaries between input and output and boundaries among different modalities, we hope to contribute to the efforts of adopting more appropriate concepts for interaction design.

ACKNOWLEDGMENTS

This work is funded by Finnish Funding Agency for Technology and Innovation, and the following partners: GE Healthcare Finland Ltd., Suunto Ltd., Sandvik Mining and Construction Ltd. and Bronto Skylift Ltd.

REFERENCES

[1] Lécuyer, A., Coquillart, S., Kheddar, A., Richard, P. and Coiffet, P. 2000. Pseudo-haptic feedback: Can isometric input devices simulate force feedback? In Proc. IEEE Int. Conf. on Virtual Reality, 2000. Pp. 83–90.

[2] Avanzini, F. Serafin S. and Rocchesso, D. 2004. Friction sounds for sensory substitution. In: Proceedings of ICAD 2004, Sydney.

[3] Card, S. K., Moran, T. P. and Newell, A. 1983. The psychology of human-computer interaction. Hillsdale, NJ: Lawrence Erlbaum Associates.

[4] Gardner, H. G. 1985. The mind's new science. New York: Howard Gardner.

[5] Varela, F., Thompson, E. and Rosch, E. 1991. The Embodied Mind. Cambridge, MA: MIT Press.

[6] Shannon, C. E. and Weaver, W. 1949. The Mathematical Theory of Communication. Urbana, IL: The University of Illinois Press.

[7] Gallese, V. and Lakoff, G. 2005. The Brain's Concepts: The Role of the Sensory-Motor System in Reason and Language. Cognitive Neuropsychology, 22, 455–479.

[8] Pirhonen, A., Brewster, S. and Holguin, C. 2002. Gestural and audio metaphors as a means of control for mobile devices. In: Proceedings of CHI2002, 20-25 April 2002, Minneapolis, Minnesota. Pp. 291–298.

[9] Pirhonen, A. 2005. To simulate or to stimulate? In search of the power of metaphor in design. In Pirhonen, A., Isomäki, H., Roast, C. and Saariluoma, P. Future Interaction Design. London: Springer Verlag. Pp. 105–123.

[10] Norman, D. A. 2002. The design of everyday things. New York: Basic Books.

[11] Jousmäki, V. and Hari, R. 1998. Parchment-skin illusion: sound-biased touch, Curr. Biol. 8, 190.

[12] Guest, S., Catmur, C., Lloyd, D. and Spence, C. 2002. Audiotactile interactions in roughness perception, Exp. Brain Res. 146, 161–171.

[13] Mensvoort van, K., Hermes, D. J. and Montfort van, M. 2008. Usability of optically simulated haptic feedback. Int. J. Human-Computer Studies, 66, 438 – 451.

[14] Gibson, J. J. 1979. The ecological approach to visual perception. Boston, MA: Houghton Mifflin.

[15] Gibson, J. J. 1966. The senses considered as perceptual systems. Boston, MA: Houghton Mifflin.

[16] Noë, A. 2004. Action in Perception. Cambridge, MA: MIT Press.

[17] Broadbent, D. E. 1958. Perception and Communication. London: Pergamon.

[18] Kuhn, T. S. 1970. The structure of scientific revolution. Chicago, IL: University of Chicago Press.

[19] Horrey, W. J. and Wickens, C. D. 2006. Examining the impact of cell phone conversations on driving using meta-analytic techniques. Human Factors 48(1), 196 – 205.

[20] Johnson, M. and Rohrer, T. 2007. We Are Live Creatures: Embodiment, American Pragmatism and the Cognitive Organism. In: Zlatev, J., Ziemke, T., Frank, R. and Dirven, R. (Eds.). Body, Language, and Mind, Vol. 1, Berlin: Mouton de Gruyter. Pp. 17–54.

[21] Kohler E., Keysers C., Umiltá M.A., Fogassi L., Gallese V. and Rizzolatti G. 2002. Hearing sounds, understanding actions: action representation in mirror neurons. Science 297, 846–848.

[22] Johnson, M. 1987. The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. Chicago, IL: University of Chicago.

[23] McGurk, H., and MacDonald , J. 1976. Hearing lips and seeing voices. Nature, 264, 746–748.

[24] Godøy, R. I. 2003. Motor-Mimetic Music Cognition. Leonardo, 36(4), 317–319.

[25] Fodor, J. A. 1975. The Language of Thought. Cambridge, MA: Harvard University Press.

[26] Tuuri, K., Pirhonen, A. and Välkkynen, P. Bodily Engagement in Multimodal Interaction: A Basis For a New Design Paradigm? In: Kurkovsky, S. (Ed.). Multimodality in Mobile Computing and Mobile Devices: Methods for Adaptable Usability. Hershey, PA: IGI Global. (In press).

[27] Hoggan, E. and Brewster, S. A. 2007. Designing audio and tactile crossmodal icons for mobile devices. In: Proceedings of ACM International Conference on Multimodal Interfaces 2007 (Nagoya, Japan). ACM Press. Pp. 162 – 169.

[28] McLuhan, M. 1966. Understanding Media: the Extensions of Man. New York, NY: New American Library.

[29] Rizzolatti, G. and Craighero, L. 2004. The mirror-neuron system. Annual Review of Neuroscience, 27, 169 – 192.

[30] Leman, M. 2008. Embodied Music Cognition and Mediation Technology. Cambridge, MA: MIT Press.

[31] Clark, A. Supersizing the mind. 2008. New York: Oxford University Press.

[32] Buxton, B. 2007. Sketching User Experiences. San Francisco, CA: Elsevier.

[33] Dourish, P. 2001. Where the Action is: The Foundations of Embodied Interaction. Cambridge, MA: MIT Press.