Being There: Concepts, effects and measurement of user presence in synthetic environments
G. Riva, F. Davide, W.A IJsselsteijn (Eds.)
Ios Press, 2003, Amsterdam, The Netherlands

20 Interaction Manifestations in Multi-player Games

Tony MANNINEN

Abstract. This chapter describes the findings of ethnographical research which elaborates and analyses the interaction forms of a contemporary multi-player game. The motivation for the research originates from the issue that the lack of intuitive and non-intrusive interaction cues is one of the distinctive features separating shared virtual environment settings from face-to-face encounters. The analysis of the interaction forms in a multi-player game session indicates that the participants of collaborative virtual environment can use various forms of non-verbal communication and perceivable actions to enhance the social presence. However, players tend to communicate outside the game system and they try to overcome the limitations of the systems by inventing various imaginative ways to communicate, co-ordinate and co-operate. This indicates that there is a need for additional interaction support.

Contents

20.1 Introduction

The aim of this chapter is to describe and analyze the interaction forms that exist in a contemporary multi-player 3D game. Interaction forms are perceivable actions that act as manifestations of the user-user and user-environment interaction. These forms are used to convey the actions of the user to oneself and to others. They enable awareness of actions and social presence by offering mutually perceivable *visualizations* and *auralisations* within the environment. In addition, the feeling of presence and the level of psychological immersion are increased due to the communication, coordination and collaboration aspects these forms bring forward.

Multi-player 3D games include interaction forms that support, either directly or indirectly, teamwork and interpersonal communication. By analyzing these forms against the rich interaction concept model, which has been constructed as a framework, it is possible to delineate the specific interaction forms that help the creation of social presence.

The explicit description of the interaction manifestations provides additional perspective on the interaction in desktop VR systems.

The research methods used in the work are qualitative. The players were interviewed using open-ended and semi-structured interviews and the game playing sessions were observed using an ethnographical approach. The main objective of the interviews and observations was to acquire explicit descriptions of interaction forms in order to find ways to apply this knowledge to desktop VR applications.

This chapter focuses on analyzing interaction forms in multimedia games. The main research questions can be formulated as follows:

1. How are the interaction forms supported and used in a multi-player game setting?
2. How do the existing interaction forms enable the creation of shared presence?

In order to understand the scope and effect of interaction forms in virtual environments, modern networked games have been studied. One reason for this approach is the generally different nature of game settings in comparison with traditional VR applications. The relevant question at this point is whether the highly focused, entertaining and problem-solving oriented games domain provides different perspective to shared presence that is increased by perceivable interaction forms.

The starting point for this work is the various games that contain a degree of teamwork, either forced by the game plot, or arranged voluntarily by a group of players. Traditionally, role-playing games have been falling to this category, although many of them provide fellow team members as computer controlled constructs only. A more recent games domain is the networked 3D-first-person-view fighting games, such as Quake and Unreal, which provide somewhat natural possibilities for teaming-up with fellow players.

20.2 Shared presence through computer-mediation

While the early prototypes of multimedia games suffered from both user interface and communication limitations, they still were able to provide meaningful interaction.

Improved graphics rendering and other techniques reduced the effort required to approach other players for interaction. Still, there commonly exist limitations that prevent the full expression of interaction. For example, if avatars are not unique to each user, it is difficult to distinguish between strangers and familiar users [1].

The modern multimedia 3D games are very similar to the desktop VR systems and Virtual Environments (VEs) developed for other application domains (e.g., research, military, and training). VEs, just like games, provide another means of simulating real or imaginary world places and activities. They are used to create a computer-generated simulated space with which an individual interacts [2]. According to Singhal and Zyda [3] Networked Virtual Environment (Net-VE) is distinguished by the following five common features: a shared sense of space (illusion of being located in the same place), a shared sense of presence (avatars of participants), a shared sense of time (real-time interaction possible), a way to communicate (various interaction methods), and a way to share (dynamic environment that can be interacted with).

Contemporary multi-player 3D games can be considered as applications and extensions of virtual reality technologies. Riva [4] defines the 'soul' of VR as a mental experience, which makes the user believe that 'he or she is there', that he or she is present in the virtual world. When interacting with a virtual environment, or with other users, the user is no longer a mere observer of that which is happening on the screen. Instead, the user feels immersed in that world and can participate in it, in spite of the fact that these worlds are spaces and objects existing only in the memory of the computer and in the user's mind.

Computer games emphasize compelling content, fast pace of action and aspects of fun, and thus, they reduce the limitations in interfaces with creative design and artistic selectivity.

Interactivity is the extent to which the user feels convinced of the mutual effect that he or she and the environment have on one another. Level of interactivity is a function of the speed of response, the range of possible user interactions and the mapping of controls [5].

Rapid response time promotes interactivity by providing immediate feedback to operations performed by the users. The number of choices that a user can make at any point in time corresponds greatly to the level of communication technology. The third dimension, mapping of controls, is the relationship between the user's physical interactions with the input mechanisms and the changes to the virtual environment.

One of the problems with contemporary multi-player games is the fact that human beings are highly sensitive to seemingly insignificant variations in the facial expressions, gestures, and posture of their conversational partners [6]. This reduces the feeling of presence among the participants and, thus makes the interaction seem unnatural. In order to enhance the sense of presence within virtual environments, it is important to increase the quality of the physical or social interaction of the participants with each other and with the environment. Furthermore, as argued by Thalmann [7], more complex virtual human embodiments will increase the natural interaction within the environment. The more natural perception of each other increases the sense of being together, and thus enhances the overall sense of shared presence in the environment.

20.3 Rich interaction model as framework for analysis

The perceivable interaction forms have been analyzed and the results have been compared with the rich interaction concept model. The model has been constructed by collecting theoretical knowledge (e.g. communication theories) and empirical material (video recordings, interviews, and observations) from networked games and game events and from self-arranged gaming sessions.

A total of twenty games have been studied and the material has been expanded with heuristic evaluations. The aim of the model was to get as exhaustive a set of interaction forms as possible. The forms have been categorized into 12 classes, each consisting of a number of sub-concepts.

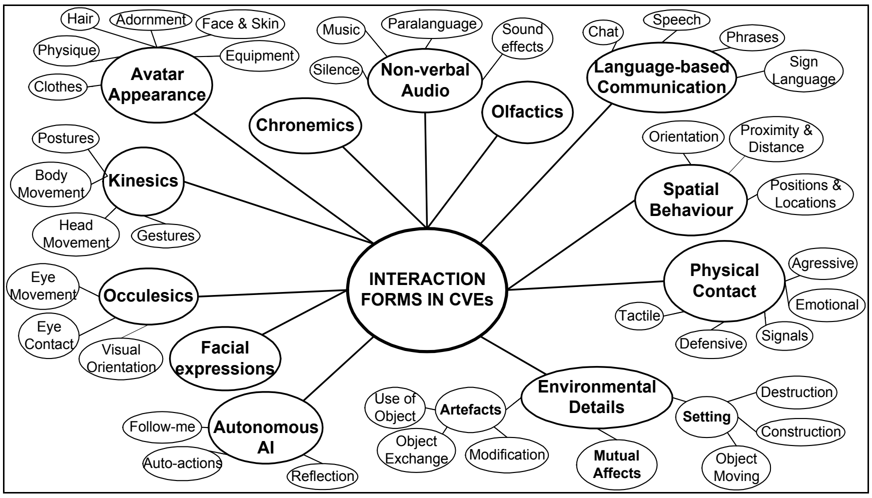

Figure 20.1 Rich interaction concept model in terms of top-level categories.

The conceptual model is a tool for analysis, and thus, is not meant for hard-coding the categories. The borders of individual concepts can be very vague, and some of the concepts can be categorized differently.

For example, the placing of individual forms varies according to different theoretical sources, so the model is more or less a compromise of several previous models. The aim of this work is to explicitly illustrate the occurrences of these forms, and, therefore, the actual taxonomy is not the most important issue.

Figure 20.1 represents the first layers of the decomposition that illustrates the proposed concept model of rich interaction forms. Additional details concerning the model can be found from other publications by this author (cf. [8]). The model illustrates the main interaction types that can be found within the current desktop VR systems. The basis for this taxonomy is the categorization of various interaction forms in terms of communication channel, context, and acting entities (e.g., body parts, environment, fellow team members, etc.). The concept model of interaction forms acts as a concrete set of examples and categories of interaction manifestations. The model consists of 12 main categories: (1) avatar appearance, (2) facial expressions, (3) kinesics, (4) occulesics, (5) autonomous / AI & automatic, (6) non-verbal audio, (7) language-based communication, (8) spatial behavior, (9) physical contact, (10) environmental details, (11) chronemics, and (12) olfactics.

While some of the aforementioned categories are self-explanatory, some of them require brief explanations. For example, the movement of the head and body (kinesics) in space (spatial behavior), in order to re-orient oneself (spatial behavior) and focus on a fellow player (occulesics) for presenting a winning triumph (facial expression and non-verbal audio) can be decomposed into various interaction form categories.

The emphasis of the model is in perceivable actions of the users and the environment. However, the model does not try to replicate the physical world interaction forms. The relation to the physical world is taken into account when applicable, but the potential of virtuality is also harnessed. The model is not meant for dictating how the participants are allowed to act and function within a VE. On the contrary, it tries to bring forward the scope and scale of interaction forms and their potential in desktop VR applications.

Figure 20.2 Screenshots illustrating scenes from Counter-Strike.

20.4 Ethnographical field research

The empirical material for this research was collected during the Vectorama 2000 LAN gaming event held in Oulu during June 2000. The players with their personal computers were all gathered within the same hall.

Furthermore, the players were using headphones to increase their sensitivity to contextual audio information and to suppress the constant outpour of background music provided by the organizers.

The games played during the event were selected according to players' wishes. The Half-Life multi-player modification Counter-Strike (CS) was the top choice, thus, confirming the fact that CS was one of the most played action games on the Internet at the time. The game is very realistic when compared to the regular Unreal/Quake "*deathmatch*" games, and players really have to work in teams and think what they are doing to achieve their goals. Counter-Strike can be best described as a "light-weight tactical combat simulation". It modifies the multi-player aspects of Half-Life towards more team-oriented game play. Each team has access to different guns and equipment, as well as different abilities. Game locations have different goals such as: hostage rescue, bomb defusing, terrorist escape, etc. Weapons include the usual assortment of pistols, shotguns, rifles, grenades, etc.

Figure 20.2 illustrates an example game environment setting with a team of players advancing towards the enemy territory (left) and the appearance of two characters with slightly different outfits (right).

Since Counter-Strike is a quite realistic military action game, it was interesting to see how much players tried to implement real-life tactics in the game as seen from the films and TV. However, it was not expected to see too serious tactical meetings before the games – that is not the general idea of fun in a gaming event like this – but some coordinated team efforts were anticipated since CS is a team game. Another interesting area of research was the communication during the games: how much text communication was used and did it have an affect in the game play? Also, since players were in the same room during the game, an amount of vocal orders and instructions were observed.

20.5 Analysis

The material collected during the observation session was limited to one game application, because most of the players stayed within the same game system during the event. On the other hand, the research provided plenty of material related to Counter-Strike. With the future work focusing also on different type of games, the material collected provides an adequate starting point for the work. The total amount of material videotaped during the event was 17 hours consisting of descriptive general pictures and more detailed screen captures.

20.5.1 Examples of perceivable interaction forms

The following section describes some of the sample findings in relation to virtual team interactions during the observed gaming sessions. Some of the examples are selected from the video recordings while others have been formulated from player interviews. The cases have been selected from the raw material based on their occurrence, importance, and level of interest in terms of this study. The descriptions are categorized according to the components of the rich interaction model.

Avatar Appearance, in this game, seems to play a sort of double role – on the one hand it provides visual information to other players, and on the other it allows the players to have different roles in the game. The *physique* of the avatar is limited in terms of size and body structure. Furthermore, team members have similar type of *clothes*, but there is some amount of variation available for the players to select (e.g., type of camouflage). The active weapon and possible protective armor (*equipment*) are also visible to others, and thus it is possible to estimate the range, speed of fire, and other characteristics of the fellow team members and opponents.

Apart from the weapon and other equipment, the appearance is relatively static and offers mainly support for identification, such as, team and role (e.g., hostage, bomb carrier, etc.). However, the level of perceived details is not always adequate to enable accurate identification. The avatars are most of the time too far to be clearly recognized.

Facial expressions are not largely supported by the system and they are not directly controllable by the players. The only perceivable expressions are combined with specific animation sequences, such as getting injured or killed (e.g., grins, heavy breathing, etc.)

Supposedly, due to the fast action nature of the game, the players are not too interested in having stronger expressions and much rather treat the avatars as pawns in the game.

The fast pace of the game may render the effectiveness of facial expressions useless, because there may not be enough time for the players to perceive these relatively small-scale cues in any case. The importance of facial expressions may be stronger in more socially oriented games, such as role-playing.

Kinesics seems to be rather insignificant in the current implementation of the game with the most common occurrences being merely slight additions to the overall atmosphere. The pace of the game does not leave too much time to observe these non-verbal cues.

Moreover, with the already complex key mappings, it could be impossible to provide adequate means to users for controlling the expressions of their avatar. However, one area where body language is widely implemented is the group of injury related expressions such as limping. In addition, *posture* changes of the avatar, during various actions, convey the act to others within vicinity. This interaction form is usually automatically represented alongside tactical action, such as jumping or kneeling to take cover.

Perhaps due to the current limitations in the game system, *gestures* are one of the most requested types of interactions among the action gaming community. The players would like to have more possibilities in signaling various messages with their avatars. Currently, the game system does not support these types of actions adequately, so the players have invented their own workarounds to overcome the communication barrier. One very clear example of gesturing is pointing to a certain direction with weapon while moving back and forth with rapid movements. This action can be used, for example, to tell a fellow team member to divert and check the location within the pointed region. In addition to this type of gesturing, there are several occurrences of using grenades for representing a direction of concern, or just for indicating one's position to a fellow team member.

The automatic gestures, which have been combined with various actions, provide mutually perceptive interaction while not increasing the control load of the player. For example, changing the weapon, reloading it, or throwing a grenade or other item, all execute specific gesture animations that are visible to oneself and to others.

Occulesics, or the eye movement and direction are not specifically modeled. The eyes stay fixed to the general direction of the face, and there are no perceivable differences.

This group of interaction forms requires higher level of detail and slower pace of action. Only the close encounters, especially with the team member or unknown character, could provide enough level of detail for accurate perception of eye movement.

Autonomous /AI & Automatic Actions play an important role in terms of game-controlled features offering interaction cues. The most common actions revolve around injuries – the overall performance of the avatar is affected by the level and location of his or her injuries, which cause certain amount of bleeding (visible to others) and restrictions to movement (limping, etc.). When the amount of injuries exceeds the critical level, the avatar falls down and the player is transferred to the observation mode and is, thus, out of the game.

The reactive actions that are stimulated by some other actions can be considered as sub-conscious or non-voluntary actions. For example, the backward movement when taking a shot or the loss of sight when affected by the flashbang are actions that cannot be interrupted. These forcefully implied interaction cues enable the simulation of occurrences that might be impossible to model in the physical world.

Non-verbal audio includes the sound effects of various actions, such as shooting, walking, running, jumping, reloading, etc. These effects seemed to be one of the major signals for others to comprehend what is going on in the game. The awareness of actions was thus greatly increased by the voluntary and involuntary sounds of action.

Silence is one important aspect of a fighting game. The clearest examples of this interaction form were the quiet walking (sneaking) and restraining of reloading while ambushing the opponent. Furthermore, certain surfaces and objects, such as doors and ladders, cause specific sound effects, which in turn can be clear signals to the nearby opponent.

Because there were no in-game support for voice communication at the time, the paralanguage was modelled only in terms of pre-programmed phrases. The game session itself contained a large amount of paralanguage that was occurring out of the game, i.e., over the monitors across the room. Various non-verbal utterances, sights and sounds were constantly perceivable. This indicates that the players react to the game in much larger sense than can be conveyed with the current game systems.

Language-based Communication is mainly managed with predefined voice messages and text-based chat channels. The cries for medic, need for backup, enemy sighted, and team status reports are the most common messages filling the audio space of the game environment. The textual chatting is an additional form of exchanging information, but due to the typing effort, seems to be less used, at least during hectic action. The contextual messaging is mostly used to provide information not supported by the predefined voice messages, or to add some relevant data to these.

One major issue observed during the gaming event was the amount of out-of-game communication over the computer screens. The players were able to hear each other even with the headphones on, so there were plenty of examples showing the importance of the social aspect in games. The players got excited, were frustrated, tried to get their acts together, or just laughed out loud during the game rounds. In this form, computer gaming seems to come very close to the traditional games with their strong social importance and feeling of togetherness. Of course, this type of interacting is only possible when the players are actually in the same physical location, so when considering pure networked gaming with isolated players, there are quite big differences. The lack of natural communication has been approached with voice-over-IP messaging, but still the player community favors the LAN parties in terms of making best out of the games.

Spatial behaviour involves mainly the character control within the environment. The *orientation* seems to be both an important cue for others to see where the player is looking, as well as, a way to overcome the restrictions of the limited field of vision. The players tend to constantly run around with necessary side-glances to see what is going on outside their field of view.

The position changes of the avatar during various actions, both conveys the act to others within the vicinity, and also provides possibilities to access, for example, an advantage over the opponent. The kneeling behind a crate, or other object within the environment, is one typical way of seeking cover and gaining an element of surprise.

The aggressive nature of the game tends to force the players to have a large distance between them and the opponent. However, there are examples of close encounters that occur due to the element of surprise and due to the narrow passages and small rooms in the environment.

Physical Contact is clearly an important type of interacting in this kind of fast action fighting game. The level of violence in general is relatively high due to the mainly *aggressive* interactions. Shooting, stabbing, and blowing-up are the main actions executed to get rid of the opposing team. However, the game is not a chaotic deathmatch, because the added winning factors bring in another aspect of team playing.

There are no direct alternatives to violence, except running away or hiding. The game system does not support any form of blocking, or otherwise *defensive* contact-oriented actions, except in the form of protective armor. The few examples of non-violent contacts include freeing the hostages by touching them, and standing in the way of an other player and thus blocking his or her route.

Environmental details include the use, modification and affect of the environment and the objects within. The environment limits the playing area in the form of boundaries, it affects the player performance (e.g., darkness), and it offers protection for the players by representing additional scenery (e.g., colors, textures, and objects).

Creating a diversion is tactically one of the most frequent form of interactions. Diversion is created, for example, when a major force attacks straight on while other team members circle around to surprise the opponent from behind. This type of activity is strongly related to the layout and familiarity of the environment. If the teams are experienced players, the element of surprise seems to be rather non-existent, because every player knows all the advantage points of the environment.

Artifact related interaction consists mainly of handling various weapons, ammunition, armor, and other equipment. The combination of different weapons and corresponding player roles (e.g., sniper, close-range, rapid fire, grenades) bring a sense of role-playing to the game, and also provides more variations in terms of tactics and playing styles.

Tactically the selection of weapon, the time of reloading, and also the use of grenades affect strongly on the performance, so in this sense the artifacts and their management play an important role in this game.

Dynamic environmental details, such as *construction* and *destruction*, are generally not supported by this game, although some of the environments contain windows and doors that can be destroyed. The main reasons for this type of activities include getting an access to another place, or increasing the level of visibility towards the presumed attacking direction of the enemy. Of course, the changes in the environment can also give a hint to other players suggesting there is, or there has been, someone near the broken element. In addition to breaking, there are also doors that can be opened and closed, and even locked.

Chronemics is the form of interaction that relates to the use and perception of time. One clear example of programmed time control is the feature, which makes the individual game rounds start at a pre-set time, so the players have time to equip themselves. The end game is usually determined by one of the several winning factors, such as rescuing all of the hostages, defusing the bombs, loosing every team member, or encountering the time out. If the player gets killed before the round ends, he or she has to wait the start of the next round as an observer.

Olfactics includes the use of scents and odors. Because these are not modeled in *Counter Strike*, this group is not applicable.

20.5.2 Discussion and additional findings

The individual players seemed to coordinate their actions mostly by adapting to the actions of others - both friends and foes. The game rounds were usually won by the teams who utilized some amount of coordination. In this sense, the shared presence is highly important factor in the overall performance. The actions of other players need to be perceivable if the team is trying to coordinate their behavior.

According to the analysis, the players were relatively satisfied with the game and the team interaction support. When playing for fun, without any greater goals, they gave no or very little instructions outside the game. The players simply played observing what others did and built their own strategy to meet the needs of the team. No strategies before the game, no analyzing afterwards.

An additional interesting finding was the increase of interactions when the significance of the game increased. There was clearly a strong correlation between the amount of team interactions and importance of the game. In the team playing tournament the players really tried to help their teams by organizing their actions and by communicating over the screens. This indicates that even though the players say that they do not need any more support for team interaction, they would use it if it were available.

The players tend to invent their own form of communicating, even when the game genre in question is not a very strongly communicative one. Especially the LAN gaming sessions indicate the need for strong social togetherness, and thus, are often the venues for the strongest experiences. It seems to bring great satisfaction for a player when he or she has successfully out-performed the opponent, and can actually see the opponent's reaction in the real world.

It was relatively surprising to see how little the team members actually communicate in the middle of the game rounds. The players have obviously learnt to observe other players in order to know what would be the best place to be in.

This can be a result of an adaptation process – in a way the feeling of presence has been enforced by the familiarity of the environment and the missions.

Furthermore, since it is impossible to write, control your avatar, and shoot at the same time in the current games, people decide to keep themselves alive instead of chatting with each other. The limited hand signals and reduced non-verbal communication available in today’s 3D games are severely restricting the communicative possibilities of the multi-player gaming.

20.6 Conclusions

The nature of the application domain seems to help the players' adaptation process in terms of high motivation levels. Furthermore, the suspension of disbelief, being a general phenomenon in entertainment domain, creates enhanced feelings of presence for the players. In this type of setting the users do not necessarily require realistic implementations, but instead they are willing to believe to be part of the imaginary world.

According to the material collected during this research, the support for team interaction in action games is adequate. However, the players tend to communicate by giving orders and information outside the game, over their screens or by using third-party software. This indicates that the players would like to have some additional interaction support as well.

Majority of the players observed during the research were experienced players, who generally tend to adapt to the complexities and limitations of the systems. The findings suggest that especially less experienced and novice users would benefit from the richer set of interaction possibilities.

The main implications for design, provided by this work, are not necessarily technical ones. The increase in bandwidth and communication channels can enhance the interaction, but the main task is to balance the system using artistic selectivity and principles of game design while creating team-oriented applications. The designers of these environments could really learn from the field of entertainment and game industry, but also from the conceptual work and theoretical models of interaction. The need to understand the concept of interaction in these desktop VR environments seems to be highly relevant if we want to construct rich experiences for the users.

The results of this research are significant for VE designers as they illustrate the importance and possibilities of rich interaction in networked settings. It is possible to reduce the limitations and restrictions of computer mediation by enabling more flexible and natural interaction. Although the naturalness and intuitiveness of face-to-face communication is hard to achieve, the virtual environments provide additional and novel ways to enhance the weak areas of interaction.

The rich interaction concept model can aid the designers by directing the development to support the important areas of interaction. The freedom of choice and possibilities to combine various interaction forms can enhance the communication, co-ordination and co-operation aspects of the VR applications.

20.7 References

[1] B. Damer, C. Kekenes & T. Hoffman, Inhabited Digital Spaces, *Computer Human Interaction*, Proceedings, **1996**.

[2] B. G. Witmer, J. H. Bailey, & B. W. Knerr, Virtual spaces and real world places: transfer of route knowledge, *International Journal of Human-Computer Studies*, **1996**, 45, 413-428.

[3] S. Singhal & M. Zyda, *Networked Virtual Environments: Design and Implementation*, ACM Press, **1999**.

[4] G. Riva, From Technology to Communication: Psycho-social Issues in Developing Virtual Environments, *Journal of Visual Languages and Computing*, **1999**, 10, 87-97.

[5] J. F. Nunamaker, Future research in group support systems: needs, some questions and possible directions, *International Journal of Human-Computer Studies*, **1997**, 47, 357-385.

[6] H. J. W. Spoelder, E. M. Petriu, T. Whalen, D. C. Petriu, & M. Cordea, Knowledge-based animation of articulated anthropomorphic models for virtual reality applications, *Instrumentation and Measurement Technology Conference*, Proceedings, **1999**, 690-695.

[7] D. Thalmann, The Role of Virtual Humans in Virtual Environment Technology and Interfaces. *In Frontiers of Human-Centred Computing, Online Communities and Virtual Environments* (Eds.: R. Earnshaw, R. Guejd, A. van Dam, & J. Vince), Springer-Verlag, London, **2001**, 27-38.

[8] T. Manninen, Rich Interaction in the Context of Networked Virtual Environments - Experiences Gained from the Multi-player Games Domain, *In Joint Proceedings of HCI 2001 and IHM 2001 Conference* (Eds.: A. Blanford, J. Vanderdonckt & P. Gray), Springer-Verlag, **2001**, 383-398.