

Chapter 15

Risk Issues in Developing Novel User Interfaces for Human-Computer Interaction

Gudrun Klinker, Manuel Huber, and Marcus Tönnis

When new user interfaces or information visualization schemes are developed for complex information processing systems, it is not readily clear how much they do, in fact, support and improve users' understanding and use of such systems. Is a new interface better than an older one? In what respect, and in which situations? To provide answers to such questions, user testing schemes are employed. This chapter reports on a range of risks pertaining to the design and implementation of user interfaces in general, and to newly emerging interfaces (3-dimensionally, immersive, mobile) in particular.

Keywords Human-computer interaction · Design-risk · Miscommunication risk · Augmented reality · Usability testing

The Facts

- In our modern society, much of what we do is at least partially supported, guided or influenced by information stored, simulated or analyzed in computers.
- It is important that everybody is able to use and understand such virtual information without mental or physical barriers.
- Research in human-computer interaction strives towards finding suitable interaction paradigms that support people in using computer information in their daily activities—for both personal and professional use.
- Research on suitable interaction metaphors analyzes risks of miscommunication between humans and computers.
- Potential miscommunication can be a challenge (e.g., in games), a nuisance (in uncritical situations) or a physical danger (in life-critical situations).

G. Klinker (✉) · M. Huber · M. Tönnis

Fachgebiet Augmented Reality, Department of Informatics, Technische Universität München, Boltzmannstr. 3, 85748 Garching bei München, Germany
e-mail: klinker@in.tum.de

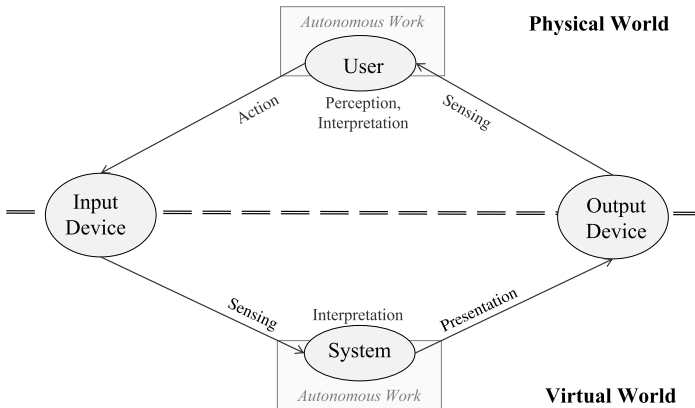


Fig. 1 Cycle of interaction based on human and computer-based perception, interpretation and action/presentation (adapted and extended from [2])

1 Introduction

When users interact with computer systems, they, as well as their real, physical environment, get in contact with the virtual world in the computer, as shown in the interaction cycle adapted from Bowman et al. [2] in Fig. 1. At the interface between the physical and virtual worlds are input and output devices that sense human actions via dedicated sensors as input signals, interpret and act upon them before rendering suitable output signals on displays. These, in turn, are perceived by the users and interpreted before a new interaction cycle starts.

Risks of misperception, misinterpretation and mispresentation exist at all stages in this interaction cycle. Users may not know well enough what actions they have to perform and how carefully they need to act them out such that the system can decipher them unambiguously. Sensor systems suffer from noise and various physical limitations. Furthermore, interpretation algorithms may be lacking some of the physical-world context when they analyze their input data, resulting in false positive and negative decisions in their command recognitions. The subsequently generated visualizations may fall short of representing the wealth of available information with appropriate clarity and detail on the available display hardware. Users may overlook important issues in the visualizations, or they may draw wrong conclusions because they are not familiar with the metaphors that were used.

For these reasons, human-computer interfaces need to be tested thoroughly and repeatedly to minimize the risk of miscommunication. In user-centered approaches, various different testing methods are applied throughout the entire product design and development life cycle. Yet, such testing has its own set of risky fallacies. The subsequent sections address each of these issues in detail. We begin with a brief description of a few current developments.

2 Examples

In recent years, user interfaces have progressed rapidly. They move away from the well-established WIMP¹ style of the Desktop metaphor that provides direct manipulation on a raster display, as described in the seminal text book by Shneiderman and Plaisant [13] towards highly immersive, multi-modal and multi-media, ubiquitous or mobile multi-touch-based interfaces (see, for example, Myers, Hudson, and Pausch [43] for further reading). Technological advances, regarding speed, resolution and accuracy of sensing devices, have recently triggered a number of novel user interface schemes to find their way into commodity devices, such as smartphones and game consoles. This section presents a few examples of such novel, post-WIMP user interfaces, as described by van Dam [16], and briefly glimpses at associated current user interaction issues.

2.1 Multi-touch

Very prominently, novel devices, such as smartphones, tablets^{2,3} and larger surfaces⁴ provide (*multi-*)*touch* input facilities: one or more users can jointly manipulate several virtual objects on small or large screens by touching them with one or more fingers.

The left picture in Fig. 2 shows three users collaboratively solving a Sudoku game on a large tabletop surface, presented by Echtler [30]. In the right picture, an ambulant incident officer uses a multi-touch tablet PC to monitor and organize the actions of a medical relief unit during the triage process of a catastrophic event, discussed in Nestler [44] (see also Iserson and Moskop [37]).

Issues: Some interaction schemes, such as a pinching gesture to resize an object, are becoming commonly understood. Yet, beyond such basic schemes, there is not yet a generally accepted way of moving, grouping and manipulating objects via multi-touch. We investigate suitable multi-touch use on a heavy, rugged device while a user is holding it in two hands (Coskun et al. [5]).

2.2 Mobility, Augmented Reality

By tracking users, *mobile* location-based services or ubiquitous computing (Weiser [17]) and *augmented reality* (AR) (Azuma et al. [1]) provide users with computer

¹Windows, Icons, Menus, Pointers.

²<http://www.apple.com/iphone/> (accessed 2012-02-26).

³<http://www.android.com> (accessed 2012-02-26).

⁴<http://www.microsoft.com/surface> (accesses 2012-02-26).



Fig. 2 Multi-touch interaction. (a) A collaborative sudoku game. (b) Coordinating support during a catastrophic event on a tablet PC

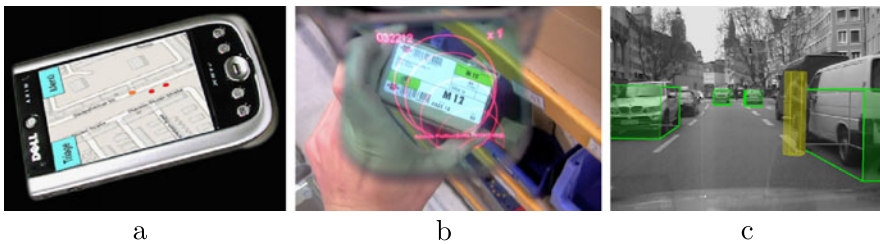


Fig. 3 Navigation assistance on mobile devices. (a) A bird's eye view on a PDA. (b) Logistics: ego-centric tunnel in a head-mounted display, leading to an object. (c) Ego-centric driver assistance in a car

information directly based on where they currently are and what they do. With AR, users see such information three-dimensionally embedded into their physical environment.

The left picture in Fig. 3 shows a PDA-based (2D) navigation assistant for each member of a rescue team in catastrophic events (Nestler [44]). The red dots indicate injured patients who need help urgently. Assignments of patients to rescuers are coordinated by the ambulant incident officer (right picture of Fig. 2), as well as collaboratively in the rescue center on a multi-touch table, such as in the left picture Fig. 2. The central picture of Fig. 3 shows an AR-based (3D) navigation assistant for commissioning tasks in large warehouses (Schwerdtfeger [52]). The logistics workers wear a head-mounted display which shows a tunnel (pink rings) that reaches from the display to the shelf. The right picture comes from a car driver assistance application. It indicates the locations of potential obstacles in the car's drive path, as detected by the on-board sensors (Tönnis et al. [58]). Current research investigates how such information can be presented to the driver: in a central information display, by warning sounds, vibrations, or potentially directly in the driver's view in a head-up display.

Mobile user interfaces raise several critical *issues*. Since users are seeing such information while they also participate in activities of their physical environment, they must not be distracted from looming dangers. Human-computer interaction

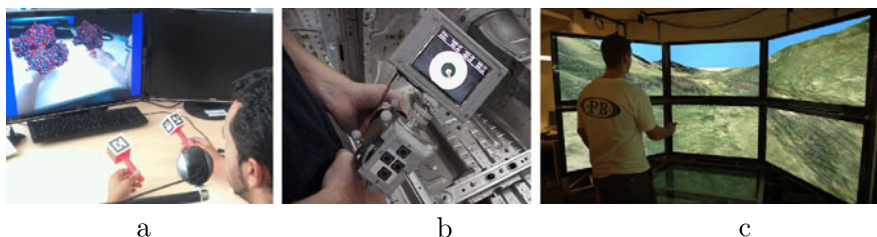


Fig. 4 Tangible interaction. (a) Augmented chemical reactions. (b) Intelligent welding gun. (c) Phone-based terrain exploration in a flexibly reconfigurable virtual environment

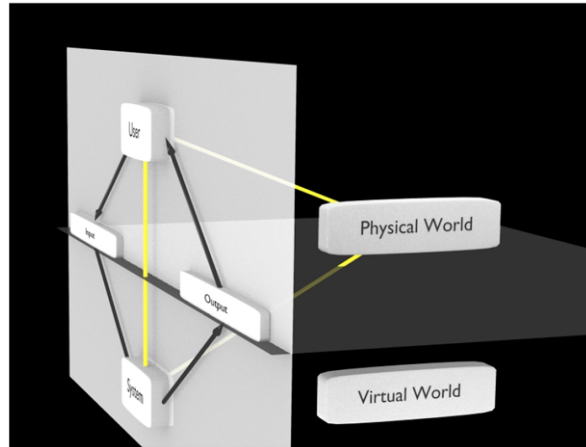
in time-critical or dangerous settings must ensure that secondary tasks, such as responding to computer information systems, do not overwhelm users such that they ignore primary tasks, such as attending to a patient (Nestler [44]) or evading physical obstacles (Tönnis [15]). Such issues will become even more urgent when AR is used in mobile settings and users have to operate in complicated physical settings, such as adapting their motion to uneven or slippery surfaces.

2.3 Tangible Interaction, Three-Dimensional User Interaction

By tracking not only users but also physical objects in a three-dimensional physical environment, *three-dimensional user interaction (3DUI)* beyond mouse, keyboard or multi-touch surface has become possible (Bowman et al. [2]). *Tangible user interfaces (TUIs)* (Ishii and Ullmer [9]) allow users to affect virtual worlds in AR and *virtual reality (VR)* applications by manipulating physical objects (Burdea and Coiffet [4]).

Examples of such tangible interaction are shown in Fig. 4. In the left picture, a user investigates and controls bonding activity between atoms of two molecules in a chemical simulation by rotating and moving the molecules via two sticks, one in each hand. The molecules change shape depending on their proximity to one another and the exerted energy fields. Special user gestures, such as holding both hands still for some time, establish and finalize proposed bonds between the molecules (Maier et al. [42]). In the central picture, a user holds a welding gun to attach a number of studs to a car frame. Using a notch and bead metaphor, a display on top of the gun indicates by an arrow inside several concentric rings where the next welding position is. When the user moves the gun to this location, a virtual ball becomes visible; when it fills the center ring completely, the gun is in perfect welding position (Echtler et al. [31]). In this case again, the (tangible) gun is tracked. Extra commands such as the welding itself and the selection of the next stud from a list are activated by special triggers and buttons on the gun. In the right picture, a user flies through a large virtual terrain by moving a smartphone in his hands like a toy airplane. Steering

Fig. 5 Interaction triangle between human, computer and physical environment (extension of Fig. 1: the *white plane* shows Bowman's interaction cycle, the perpendicular *dark plane* separates the physical world from the virtual world, with the user, other people, objects and input/output devices residing on the physical side)



commands are derived from the accelerometers in the phone and thumb gestures on the built-in small multi-touch display. This is not a one-to-one mapping from the phone motion to the virtual flight path since the user can move only slightly in front of the screen. Current research investigates what motion gestures are most useful for users to navigate as quickly or precisely as possible along an intended path in a large virtual environment (Benzina et al. [24], Tönnis, Benzina, and Klinker [57]).

Research on three-dimensional human-computer interaction needs to determine the most suitable combination of interaction facilities, such as object or user motion (gestures), buttons, voice commands and more (Bowman et al. [2], Sandor [50]). Furthermore, research needs to determine where in the vicinity of a user these facilities exist (fixed within the environment or attached to the user or to an object) (Feiner et al. [32]).

An important *issue* regards the question how users can immerse deeply into exploring a high-dimensional space of simulated or measured data without being distracted by the human-computer interfaces. Visualization, simulation, virtual animation and interaction need to be so intuitive that the computer becomes virtually invisible (Norman [46]). The computer and human become partners in exploring and analyzing the information, with the computer amplifying human intelligence (Brooks [3]).

This human-computer partnership, embedded within a physical setting, is the overarching issue across all areas of multi-touch, mobile, AR-related or virtual *human-computer interaction (HCI)*. *Human users*, the *physical world* including sensors and displays, and the *computer system* (including a virtual world full of simulations and animations), form an intricate triangular relationship (Fig. 5). Each corner of this triangle has its own set of errors or risks, all of which need to be dealt with in order to determine good human-computer interfaces.

3 Computer-Related Risks

The development and enhancement of human-computer interfaces, such as those presented in Sect. 2, has various sources of uncertainty, as represented by edges and nodes in Fig. 1. Uncertainties in measurements, interpretations, and presentations or actions result in the risk of miscommunication between humans and machines, which, in turn bear the risk of harmful consequences, if the user interfaces control computer programs with significant impact on our lives. This and the following section describe issues related to these risks. Tables 1–6 relate them to the application examples of Sect. 2.

Figure 1 shows the full interaction cycle between humans and computers (Bowman et al. [2]). This section presents issues pertaining to the uncertainties on the computer side, i.e., issues a computer system designer has to take into account when conceiving, building and testing the system hardware, the computer algorithms, and the underlying concepts. These are represented in the lower part of Fig. 1.

3.1 Sensing

Sensing is the first technical step in human-computer interaction. It is represented by the lower left arrow in Fig. 1. It receives the original input from human users and provides it to the computer system. Table 1 describes sensing requirements and impacts for the exemplary post-WIMP interfaces that were presented in Sect. 2.

To determine user input, suitable sensors need to be installed and registered, and their sensing properties need to be calibrated. Even though this is an issue with any kind of user interface, the following section focuses on issues pertaining to trackers for the multi-touch and tangible interfaces that have been presented in Sect. 2. Position $\mathbf{p} = (x, y, z)^T$ and orientation $\mathbf{r} = (\theta, \phi, \psi)^T$ of a user or an object are generally described as a pose $\mathbf{X} = (\mathbf{p}, \mathbf{q})^T$ with six degrees of freedom in a three-dimensional environment.⁵ Orientation corresponds to rotations around three axes that can be provided as Euler angles, in matrix notation or as quaternions. Different fields refer to the rotation angles in different terms, such as yaw, pitch and roll for aircrafts or azimuth, elevation and tilt in astronomy.

Several physical principles can be used to determine and track object poses: optical, inertial, electro-magnetic, acoustic, radio-based, or mechanical tracking (Welch and Foxlin [18]). Each such principle suffers from errors that are generally classified and handled in a number of ways. To some extent, sensor error may be characterized as white noise $\mathcal{N}(0, \Sigma)$, following a Gaussian distribution with mean value 0 and covariance matrix Σ that depends on sensor-internal imprecision with respect to all six degrees of freedom. This is the accumulation of many unknown physical sources and is summarized according to the central limit theorem. Yet, not all

⁵Two-dimensional multi-touch surfaces require three degrees of freedom with $\mathbf{p} = (x, y)^T$ and maximally one rotation angle θ .

Table 1 Computer sensing issues in exemplary applications

Applications	Interfaces	Issues
Sudoku game	Multi-touch, TUIs (Figs. 2a, 4c)	Sensing and system reaction must be immediate and very robust and reliable such that users are able to build up an intuition for proper system control [30].
Catastrophic events	Multi-touch, Mobile (Figs. 2b, 3a)	System reliability is very important and devices thus need special protection (ruggedization), reducing e.g. the sensitivity of the touchscreen. Input signals are thus rather noisy [5, 44].
Logistics	Mobile, AR (Fig. 3b)	The position of the logistics worker (esp. the head pose) must be determined both precisely and robustly across a wide area of a warehouse, requiring complex tracking setups [48]. Further input devices (such as push buttons on a belt) are needed for system control. They must be usable intuitively and blindly [52].
Driver assistance	Mobile, AR (Fig. 3c)	Apart from user input, a large amount to environmental sensor data is required [58]. Relevant measurements include for example the friction coefficient of the street at the relevant section in order to correctly compute and display the breaking distance. For driver assistance, the state of the driver also has to be taken into account. To this end, eye and head tracking are becoming integrated into driver cabins.
Augmented chemical reactions	AR, TUIs (Fig. 4a)	The application uses a “desktop-AR” setup, consisting of tangible objects with markers, in front of a desktop monitor and a camera close to the user’s head [42]. The optical marker tracking algorithm must be fast, precise and robust against partial occlusions of the markers. The camera must be close to the user’s eyes in order to minimize discrepancies between the fields of view of the camera and the user.
Intelligent welding gun	Mobile, TUIs (Fig. 4b)	Stud welding requires sub-millimeter precision and thus very precise tracking. Further input methods for system control must be accessible during the (mobile) welding process [31].
Terrain exploration	TUIs, VR (Fig. 4c)	Sensing of flying gestures requires suitably accurate tracking data, either from a stationary tracker in a fixed VR setup, or from mobile inertial or touch sensors, e.g. built into mobile phones or tangible objects [24]. Thus this represents a tradeoff between high tracking fidelity with high setup costs and mediocre tracking quality without setup costs.

influences average out that way. Some, rather specific errors contribute systematic deviations from the true mean pose of an object, resulting in an inaccurate pose estimate with a systematic offset in position and/or orientation. This may stem from misaligned, or imprecisely placed sensors in an environment, such as a camera after someone has bumped into it. It can also stem from inaccurate depth measurements, if, for instance, a camera possesses an automatic zooming function. A third cause of inaccurate pose estimations is temporal measurement lag. Careful calibration and registration procedures both in the spatial and in the temporal domain are required in order to obtain precise and accurate input data (Huber [8], Keitler [39]).

Further problems arise from the physical limitations of sensors. Camera-based tracking fails when the direct line of sight to a tracked object is lost, e.g. because the object is temporarily occluded by another object due to current object or user motion or when an object leaves the field of view of a camera. Inertial sensors suffer from drift. Field-sensing devices, such as electro-magnetic trackers, compasses or radio-based trackers, loose precision when unforeseen further sources, such as magnetic objects, are added to the environment.

Due to individual sensor limitations, hybrid combinations of sensors are investigated. Many concepts of sensor fusion exist in probabilistic robotics (Thrun, Burgard, and Fox [56]), including Kalman filtering and particle filters. An important aspect involves the construction of robust, redundant sensor networks that combine mobile and stationary sensors (Pustka et al. [48]). In such networks, pose estimations from different sensors are transformed back and forth between different sensor coordinate systems, using forward and backward propagation, with respect to both geometry and sensor errors (Bauer [21]).

Furthermore, diligent calibrations and registrations of sensors and physical reference targets are required, to be performed by a *tracking engineer*. The degree of quality of this work, as well as of the sensors involved, has a serious impact on the performance of the entire human-computer interaction system, to the extent that poor quality may render the system dysfunctional. There is the risk that quality can vary over time, with users not being aware of the current quality level. In applications, such as medical surgery (Bauernschmitt et al. [22]) or high precision metrology (Keitler [39], Luhmann [40]), the current quality has to be checked frequently.

3.2 Interpretation

Interpretation is the central computational step in human-computer interaction. It is represented by the bottom circle in Fig. 1. Table 2 describes computer interpretation issues for the exemplary post-WIMP interfaces that were presented in Sect. 2.

Interpretation receives low-level information, such as a continuous flow of pose data, from the sensors and analyzes it in order to derive higher-level interpretations of users' intended commands to the system, as well as the current state of a changing physical environment. In the background, the computer system then does, what it is best at. According to the received commands, it accesses and analyzes data, computes new results and/or simulates situations within the constraints of a given model and according to further sensor input that monitors aspects of the physical world. Finally, the computer system then provides its analysis to the computer output component to generate appropriate output (visualizations and/or actions).

The complexity of interpreting user input depends on the number of different commands or steering controls that need to be distinguished. In principle, commands can be discrete events or they can relate to continuous control. For continuous control, the stream of tracked user or object poses is transformed into manipulation

Table 2 Computer interpretation issues and solutions in exemplary applications

Applications	Interfaces	Issues
Sudoku game	Multi-touch, TUIs	Due to similarities between multi-touch gestures, these are occasionally misinterpreted. E.g., there can be confusions between two-finger zooming and rotation.
Catastrophic events	Multi-touch, Mobile	Gestures are sometimes not sensed well on a ruggedized device and thus not recognized. Continuous signals can thus be interrupted. Extrapolations from history into a limited future are able to bridge small sensor gaps [5].
Logistics	Mobile, AR	To reduce the risk of wrong interpretations, control input is collected via a rotary dial with a push button [52].
Driver assistance	Mobile, AR	Car traffic scenes—especially at large intersections or in varying weather conditions are extremely complex and variable. The analysis of such sensor data is a long-standing research issue in robotics. For driver assistance, head tracking [58] and the determination of the driver's current state of mind (e.g., from glancing behavior) is added.
Augmented chemical reactions	AR, TUIs	Several different gestures are necessary to select and confirm one of many potential bonds between molecule models. These may differ only minimally from each other. Accordingly, the recognition may misinterpret these gestures [42].
Intelligent welding gun	Mobile, TUIs	The main input device is the welding gun itself. The risk of wrong gesture interpretation is avoided by using menu-based interaction.
Terrain exploration	TUIs, VR	There is no unique mapping from 6 DoF pose tracking to the control of the flight path of an airplane—with only 4 parameters. Proper transfer functions need to be defined [24].

of virtual objects according to predefined transfer functions. Steering results are directly related to the tracking quality of the sensors. Yet, the transfer functions may provide some filtering such as damping to reduce jitter that is due to sensor noise.

For recognizing discrete commands, the system designer needs to know how many different commands exist and how they can be distinguished. This means that a vocabulary of gestures needs to be considered, with each gesture being associated with distinctive features. Using machine learning and pattern recognition techniques, sequences of user poses (i.e.: user gestures) need to be compared and clearly separated from each other. For each gesture, the features form a cluster in some measurement space. Clusters from different gestures should not overlap—better: there should be a wide gap between clusters such that they can be distinguished even under the presence of noise. To this end, recognition algorithms need to be designed that derive appropriately distinguishable properties (the measurement space) from pose sequences. In addition to distinguishing between gestures, also false alarms (false positives) need to be considered and discarded.

The risk of misinterpreting gestures arises from non-unique situations, i.e. situations for which noisy pose sequences cannot be associated clearly with exactly one command. Another risk comes from the fact that the underlying world model for the design of appropriate gestures and the reaction to measurements of physical events

may not have been complete: in real-world physical settings, situations may arise that lead to unintended gestures that are interpreted inappropriately in the limited scope of a computer application because the overall context was not fully modeled and understood (Neumann [11]).

3.3 Presentation

Presentation is the last computational step in human-computer interaction. It is represented by the lower right arrow in Fig. 1. It receives interpretations of the current user input and of the state of the physical environment from the interpretation component as well as the results of the background work, such as the current state of a simulation. It generates appropriate visualizations on output devices, to be perceived by the users. Table 3 describes computer interpretation issues for the exemplary post-WIMP interfaces that were presented in Sect. 2.

The amount of information that is acquired, generated and processed inside computers can be tremendous. The information presentation component is concerned with issues (1) what to show, (2) how to show it and (3) where to show it. Information presentation may involve hundreds of different attributes in a high-dimensional property space. Objects can have very intricate relations to one another, forming clusters, correlations, anti-correlations etc.

The first question is, *what to show*. Data reduction schemes such as projections from high-dimensional spaces to lower dimensions, as well as selections and combinations or rearrangements of dimensions using, e.g. principal component analysis are employed. There is a risk that important information is omitted or hidden in accumulations or projections along an attribute dimension. Interactive data exploration schemes and automatic data mining are part of an answer to such risks, allowing users to poke at the data and massage it until they are convinced that they have observed and explored all relevant aspects. Yet, short of performing an exhaustive search, little guarantee can be given that the entire body of information has been presented in all possible combinations of and along all attribute dimensions. An emerging concept of information selection for mobile applications concerns context dependency. It cannot be formulated as succinctly in mathematical terms but rather depends strongly on the interpretation of sensor data and the assumed world (context) model—bearing the risk that such model may not be complete and interpretations thus deficient (see Sect. 3.2).

The next question is, *how to show* the selected information. Information visualization and scientific visualization are concerned with developing schemes to present a wealth of information to users, bringing out the essential details without losing the overview of the general context (Spence [14], Bederson and Shneiderman [23], Nielson, Hagen, and Müller [45], Tufte [59]). Quite a number of concepts exist on how to represent information in perceivable (visual, aural or tactile) form, by mapping attribute dimensions to the dimensions of a representation scheme. Information can be represented both for individual objects (object visualization) and

Table 3 Computer presentation issues

Applications	Interfaces	Issues
Sudoku game	Multi-touch, TUIs	Since the display is partially occluded by the hand during a multi-touch interaction, users may not be aware of all provided feedback. In programs that make heavy use of motion-sensors as input, showing detailed information could lead the user to either a bad performance or ignoring the information because the user needs to stop the motion of the display to be able to read the information properly.
Catastrophic events	Multi-touch, Mobile	The system shows a map plus icons of victims, rescue personell, ambulances etc. Issues involve how to provide both an overview and detail and how to arrange the icons in dense areas where there is not enough space to show them side-by-side. Another issue is the suitable presentation of aggregate information. The entire presentation is shown on a large multi-touch table. Specialized views are also shown on a handheld tablet PC and on mobile phones to support and coordinate activities on-site [20].
Logistics	Mobile, AR	The logistics application shows a tunnel that directs the user to the picking target. The curvature of the tunnel reflects the distance and the turning angle to the target. Additionally, the application identifies the relevant target at the destination [52].
Driver assistance	Mobile, AR	Information related to neighboring or approaching traffic participants can be shown in a variety of modalities and at several places in a car, e.g. in a central display, a head-up display, as well as via sound or a vibrating seat, steering wheel or gas pedal. Exemplary information includes driving directions, augmented onto the street, required breaking distances, the drive path, as well as the direction of looming dangers that the driver should attend to [15].
Augmented chemical reactions	AR, TUIs	Aside from showing the current state of two molecules, it is important to show some or all potential further bonds, as well as the status of current gesture analysis and recognition (e.g. the upcoming timeout for a “holding still” gesture) [42].
Intelligent welding gun	Mobile, TUIs	To guide the welder with the required accuracy, not only navigational aids are displayed, but also accuracy indicators that show the deviation of the current position of the tip of the welding gun from the targeted welding spot. Furthermore, the target position is presented textually at the top of the display—as a global orientation aid [31].
Terrain exploration	TUIs, VR	To immerse viewers into the terrain data, it is spread across an arrangement of 3 walls of a fully recoverable cave, FRAVE [57]. Each wall consists of two large stereo displays. Two further displays are placed on the floor, equipped with multitouch sensing facilities.

for statistical aggregations with respect to attribute dimensions (attribute visualization), such as histograms, scatter plots, and parallel coordinates plots. To this end, two or three spatial dimensions and the temporal dimension (animations or inter-

active, iterative steering) can be used to layout data geometrically as framing dimension that span the spatial layout of a representation scheme. These spatial dimensions can also be recursively sub-arranged (nested) to show blocks of data from further dimensions—contained dimensions representing information in each cell of the framing dimensions. Visualization schemes for contained dimensions are, for example, color, semi-transparent presentations, texts and glyphs with special shapes, orientations etc. The use of special structures such as trees or graphs leads to further well-established options. In many mobile and AR-based applications, the huge amount of data is not as much an issue as the question how to find suitable three-dimensional metaphors to relate the virtual data to the physical world of the user without occluding too much of the environment. Should the information be represented in a first-person perspective (ego-centric view) or in a bird's-eye perspective (exocentric view) (Bowman et al. [2])? What are suitable metaphors to indicate information behind a physical object—i.e., information that is currently occluded (Dey, Cunningham, and Sandor [29])? The x-ray metaphor is not immediately intuitive since it is inconsistent with physical reality. Other questions discuss how to represent operational information, social information about groups of objects or people or abstract background information that does not have a unique spatial connotation.

The final question is, *on what physical displays and where in the environment to show* the information. Representational layouts of information depend on the device, as well as on the available compute power and network bandwidth. Current presentation schemes vary from large, detailed presentations on combinations of multiple wide screens, such as a CAVE in VR over large single screens to desktop systems, tablet solutions and tiny displays on smartphones, with and without audio support (Artinger et al. [20], MacWilliams et al. [41], Sandor and Klinker [51]). Display characteristics such as the resolution, the dynamic range and color gamut of a display, the field of view and field of regard that it subtains in front of a user, and its current pose play an important part in devising an information presentation concept for the human-computer-interaction aspects of a computer application (Bowman et al. [2]). Providing interactivity, e.g. via WIMP-based devices, multi-touch, or tracking also influences the information presentation schemes since the UI also needs to have visual representations on the display in form of GUIs, virtual hands, icons or avatars.

Design decisions with respect to these issues bear many risks—yet those are generally not directly related to the technical issues that are presented in this section but rather to human issues and thus will be discussed in the following sections.

4 Human Issues

Following the discussion on uncertainties on the computer side, this section presents issues pertaining to the uncertainties on the human side: human sensing, perception and interpretation, and action. These are represented in the upper part of Fig. 1.

Human issues are not risks in themselves. Yet, they need to be considered as human factors when designing the computer interaction schemes of Sect. 3. To this end, they become the focus of user-centered design and testing schemes that are presented in Sect. 5.

4.1 Human Sensing

Human sensing is the first step on the human side of human-computer interaction. It is represented by the upper right arrow in Fig. 1. It describes the instant when humans sense computer output with their innate sensory organs (Eysenck and Keane [6]). Table 4 describes human sensing issues for the exemplary post-WIMP interfaces that were presented in Sect. 2.

Human sensing has general properties and limitations, as well as special limitations of individuals, depending on age, health and other factors. In the following, the discussion is restricted to visual sensing. The human retina has two kinds of sensory cells: cones and rods. They respond to light stimuli in the spectrum of “visible light”. Cones have three different pigmentations that make them sensitive to different wavelengths and allow humans to see colors. They need significant amounts of light in order to respond. That’s why color vision works well in broad daylight, but not at night. Color blindness is caused by deficient pigmentation, e.g., when green pigments are missing. Rods, on the other hand, work at low light levels. Yet, they respond to the entire visible spectrum rather than to subranges of wavelengths. Thus, they allow humans to see at night time—in grayscale rather than in color (Eysenck and Keane [6], Gregory [7]).

An important property of human eyes is foveal acuity. The retinal focus of the eye, the fovea, contains cones with very high density. Humans can see very acutely with this part of the eye, whereas vision in the remaining areas of the retina (peripheral vision) is decreasingly acute with increasing distance from the fovea. Yet, the peripheral area is known to help humans in detecting object motion. It also provides a wider, coarse overview of the physical environment for a field of view of about 175 degrees.

A further important issue of human vision is depth perception. A large number of monoscopic and stereoscopic depth cues exist. One of the most important ones is stereopsis. With two eyes, humans see objects in front of them twice, with a horizontal offset (disparity) on each retina. The disparity depends on the distance between the eyes and the distance of an object from the eyes. The smaller the distance the larger the disparity. By triangulating, the human brain is able to estimate depth from disparity, up to a distance of about 6 meters (Gregory [7]).

Further important cues are summarized under the term adaption, describing oculomotor (muscle) activity in the eye. Humans converge their eyes inward such that the most important object is seen in the foveal areas of both eyes with the highest acuity. This converging eye rotation is used by the brain as an additional source of

Table 4 Human sensing issues

Applications	Interfaces	Issues
Sudoku game	Multi-touch, TUIs	Multi-touch tables can be rather large; inspecting objects at the other side of the table is less precise than at close distance. Objects/text can also be upside-down relative to the current viewpoint.
Catastrophic events	Multi-touch, Mobile	This application takes place in a very stressful physical environment. Users are possibly bombarded by many physical sensations, such as bright or very dim illumination, loud sounds, and physical obstacles. Virtual presentations need to be shown with the right amount of contrast such that they are neither too strong nor too weak.
Logistics	Mobile, AR	Virtual navigation information is shown in an optically see-through head mounted display (HMD). Such displays present information at a specific virtual distance from the user's eyes—typically approximately at arms' length. If the object is not close to the virtual presentation distance, users cannot focus simultaneously on a physical object and on its annotation. Furthermore, if the current background is similar to the presentation in the HMD, the virtual information cannot be perceived well.
Driver assistance	Mobile, AR	When information is displayed inside the car, users need to adapt their eyes, both for brightness and for focal distance when they look at the information and also when they go back to the road. This takes valuable time that can be critical when reacting to physical dangers. Head-up displays show information outside the car, thereby reducing the adaptation time.
Augmented chemical reactions	AR, TUIs	The molecules are shown on a common desktop monitor, providing only monoscopic depth cues and motion parallax. In this case, people have shown problems seeing the 3D structure of the molecule well. There can also be a hand-eye coordination problem.
Intelligent welding gun	Mobile, TUIs	The welder needs to look at color images on a small screen. The navigational arrows are shown in 3D. Depth perception might be an issue. Yet, motion parallax is a dominant cue since the welder and the gun are mobile. Furthermore, at welding time, the physical car frame provides strong haptic cues.
Terrain exploration	TUIs, VR	To trigger the human ability to see stereoscopically in three dimensions, each screen of the FRAVE shows two versions of the scene, one for each eye of a tracked user. He/she has to wear shutter glasses that are synchronized to all displays simultaneously such that each eye only sees the version dedicated to its viewpoint.

depth information. Accommodation is a further depth cue. Muscles contract or dilate the lens in the human eye to allow it to focus on objects at different distances. Accommodation also contributes to the brain's estimation of object distances.

In normal physical settings, stereopsis and oculomotor cues all contribute to a consistent depth perception of objects in front of a person's eyes.

In human-computer interaction, *computer displays present representations of virtual information that are subject to human sensing capabilities and limitations*. It is

important to account for potential color blindness, as well as for the fact that presentations can only be seen sharply in the very small foveal area of each eye. Thus, most parts of a computer presentation are seen without high resolution. It takes humans time to actively move their eyes across important areas of a display in order to see each area with the foveal area of their eyes. Eyes jump in saccades between different display areas. This may be in conflict with computer animations that expect users to focus on a certain display area at a particular instant in time. Thus, there is a risk that parts of a presentation are not seen with sufficient acuity by a user, since these areas were not within the foveal area at that moment.

Peripheral vision must also be considered with care (Jones et al. [38]). In many cases, it is not integrated into information presentation schemes and devices. Since peripheral vision provides humans with overview and advance notice of potential hazards in their environment, lack of peripheral input can result in risky or tiring situations: if a head-mounted display is closed around the eyes it shuts off users' peripheral view of the physical world. If open, but not covered by the display, there is a visible seam between the virtual and physical world. Furthermore, users have to rotate their heads continuously from side to side to see the information that is geometrically related to areas that are not covered by a small field of view in a wide geometric range (Rolland and Fuchs [49]).

Another issue is a sensory mismatch of depth cues for three-dimensional presentations of virtual objects in a stereoscopic display (Bowman et al. [2]). Here, convergence and stereopsis—induced by unnatural viewing conditions involving shutters, polarized or red-green glasses in front of a user's eyes—provide a depth impression that is inconsistent with accommodation: the eyes focus on the display surface rather than on the simulated depth of the virtual object. Such sensory mismatch is a problem for VR and also for AR, using head-mounted displays. Depending on the situation and the physical constitution of the user, one cue may dominate over another, thereby inducing the respective depth impression. Yet, there is the risk of users getting head aches or suffering from simulator sickness (especially, if motion cues are involved). Furthermore, there is a risk of potential after effects, i.e. the brain may adjust to this sensory mismatch and remain in this state even when the user deals with physical objects—with thus reduced sensory ability.

If, on the other hand, a video-based presentation scheme is used that replaces the *optically see-through* direct integration of virtual information into video streams of the real world (e.g. in a mobile phone on a stationary display, or on opaque head-mounted glasses), humans suffer from reduced hand-eye coordination since the viewpoint of the camera does not coincide perfectly with their eyes.

4.2 Perception

Human perception is the central step on the human side of human-computer interaction. It is represented by the top circle in Fig. 1. It describes the process when humans attend to a sensed computer output, become aware of it and thus perceive

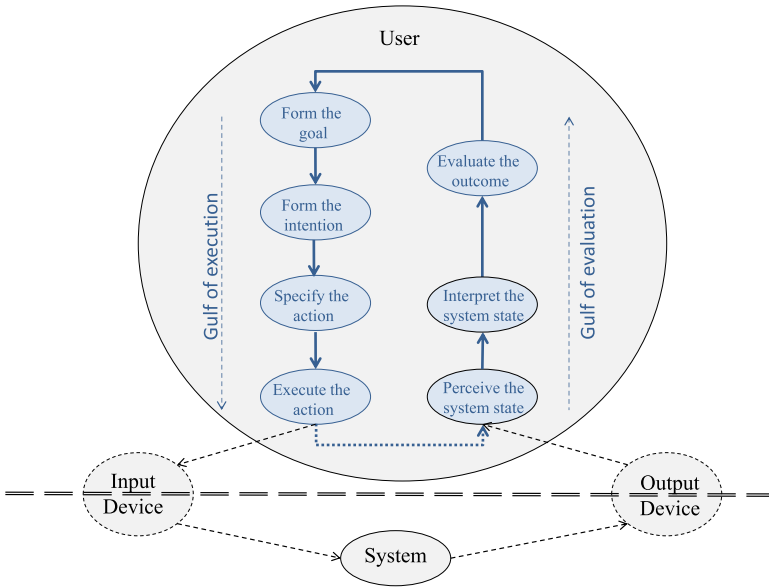


Fig. 6 Stages of action as a dynamic process of execution and evaluation during human-computer interaction (adapted and extended from [12])

it (Eysenck and Keane [6]). According to the perceived computer output, they reason about the world. They analyze the situation and draw conclusions, forming a goal and intention towards performing the next action. In the background, humans may meanwhile perform a number of independent autonomous tasks that are related to the project but do not need computer support. Eventually, they return to the human-computer interaction cycle, with the intention to provide input to the computer. Table 5 describes human perception issues for the exemplary post-WIMP interfaces that were presented in Sect. 2.

A seminal schematic of human cognitive work during human-computer-interaction has been presented by Norman [12]. Figure 6 shows Norman’s action cycle, integrated into the cycle of interaction of Fig. 1. Norman’s cycle spells out human cognitive activity in more detail, focusing on perception and interpretation as the central part. Norman connotes the transitions both from the sensing phase (Sect. 4.1) and to the action phase (Sect. 4.3) with metaphorical gulfs that users have to cross and that might overwhelm them with great or even unsurmountable challenges: the gulf of evaluation, i.e., perceiving and fully understanding the current state of a computer program from its current and past output, and the gulf of execution, i.e.: deciding what next step to take and how to convey this to the computer via its input system.

Perception requires humans to attend to stimuli. By doing so, humans may pay less attention to other stimuli. This situation is known as *perceptual tunneling* (Wickens and Hollands [19]) or *inattentional blindness*. Furthermore, experiments have demonstrated that humans can be completely oblivious to changes in parts of

Table 5 Human perception and interpretation issues

Applications	Interfaces	Issues
Sudoku game	Multi-touch, TUIs	On a large multitouch table, players may not be able to perceive all simultaneous changes on the board, if multiple players are interacting with the game (local focus, change blindness, cognitive capture).
Catastrophic events	Multi-touch, Mobile	Important information events can be missed by the rescue workers e.g. while treating the patient (change blindness). To avoid wrong decisions, each change should be clearly visible continuously, even if the rescuer missed the change when it occurred.
Logistics	Mobile, AR	The visual augmentations in an HMD may overwhelm commissioners such that they do not pay enough attention to the physical environment and run into physical obstacles (inattention blindness).
Driver assistance	Mobile, AR	Information presentation in cars needs to be kept to a minimum to ensure that drivers are not overly distracted from the physical world. In tests, drivers have shown symptoms of perceptual tunnelling from as little as a single polygonal presentation of the expected drive path “contact analog” in a head-up display: in an experiment in a driving simulator, they drove significantly faster when seeing the drive path than when seeing only a line indicating the current breaking distance [15]. Furthermore, many users also suffer from simulator sickness, due to a very strong visual stereo impression inconsistent with their haptic sense (balance, proprioception).
Augmented chemical reactions	AR, TUIs	When interacting with large molecules, users may be overwhelmed by the amount of information and simulated activity. While a user is holding still and watching the simulation, a bonding activity may get activated accidentally (change blindness, cognitive capture). The system must ensure that the user is properly informed suitably.
Intelligent welding gun	Mobile, TUIs	Welders may be so fascinated by the AR-based guidance on the display that they ignore their physical environment and bump into obstacles (perceptual tunneling).
Terrain exploration	TUIs, VR	Due to a very strong stereo impression that does not match haptic sensing, users often suffer from simulator sickness.

their surroundings that are not within their current focus—a problem called *change blindness* (Steinicke et al. [54]). In addition to ignoring stimuli, humans here also exhibit problems memorizing recent images with the level of detail that is required to compare them against newly incoming stimuli. Another well-known problem is *cognitive capture*. In this case, humans are so absorbed by a cognitive task that they ignore new, unexpected stimuli, as shown by the stunning video experiment of a gorilla walking through a basket ball game without being perceived by a large number of test persons.⁶ Even further, humans may suffer from conceptual or *cognitive*

⁶D. Simonis and C. Chabris. Selective Attention Test. <http://www.youtube.com/watch?v=vJG698U2Mvo>, 1999. (Accessed 2012-03-02.)

overload, i.e. they receive so much information that they are unable to deal with it. The result can be *cognitive tunneling* (Wickens and Hollands [19]): humans become unable to make decisions due to information overload. They may then restrict themselves to pursue only a very limited subset of available options.

These are serious *concerns for information visualization and human-computer interaction schemes* since the mere fact that information has been presented cannot be taken as a guarantee that users have actually perceived and understood it. This covers some of the aspects of Norman's gulf of evaluation. Beyond being overwhelmed by too much information, users may also experience a gulf of evaluation due to poor, misleading representation schemes. Reasons could lie in sensing difficulties (color blindness, poor resolution), or in the choice of confusing presentation metaphors.

AR and VR have different strategies and goals towards dealing with human perceptual limitations. VR strives towards generating a virtual immersive experience by exploiting human limitations, such as perceptual tunnelling, change blindness, and cognitive capture—just as magicians when they confuse spectators with their tricks. Users are expected to overlook or suppress the perception of cues that tell them that the physical reality is different from the virtual experience. VR faces the danger of simulator sickness or a non-perfect sense of presence when sensing mismatches are not sufficiently strongly overwhelmed by the sense that is intended to be dominant. AR, on the other hand, needs to ensure that users co-exist safely with their physical environment. To this end, virtual information must not overwhelm the user's senses to an extent that physical reality is ignored. Virtual distractions may lead to a lack of situation awareness due to perceptual tunnelling, information overload or cognitive capture—with potential physical harm to the user. In evaluations of AR-related applications, users need to be assessed regarding their level of distraction, e.g. by requiring them to simultaneously perform activities related to the physical environment while also interacting with virtual information. The amount of distraction is determined via eye-tracking, analysis of the response time, and the amount of errors.

After interpreting sensor input, users face the gulf of execution when planning the next action. To this end, users need to be aware of the options that the input devices of the user interface offer. These options either need to be learned and memorized from manuals or from trial and error experiences, or the interface must be flexible enough to allow natural, spontaneous human input (such as natural speech or natural gestures). It is crucial that the user interface allows users to interact with as little contemplation of available options as possible. Users may get lost and confused in poor, unclear and inconsistent input schemes.

4.3 Action

The third, final step on the human side of the human-computer interaction cycle involves executing/performing the planned action. It is represented in Fig. 1 by the

Table 6 Human action issues

Applications	Interfaces	Issues
Sudoku game	Multi-touch, TUIs	At a large multi-touch table or in a large virtual 3D space, players may not always be able to directly touch and manipulate an object on the far side of where they are. They either have to move, or the system needs to provide metaphors for non-linear motion or manipulation via indirect pointing or selection, using a prop [2].
Catastrophic events	Multi-touch, Mobile	Collaborative control via a multi-touch table raises issues similar to the games scenario above. For the mobile part, the ruggedized tablet PC is very heavy and needs to be carried with two hands. Thus, only the thumbs are able to touch the multi-touch area of the tablet—along the vertical rims of the device. The application takes these limitations into account to design special “thumbs-only” metaphors to select and manipulate icons on the map while also scrolling and zooming the map [5].
Logistics	Mobile, AR	The user interface of the logistics application was deliberately kept simple and uses only a rotary encoder and a push-button. Still, picking incorrect items cannot be prevented by the system. Also a certain amount of training is necessary to adapt to the system.
Driver assistance	Mobile, AR	Car drivers need to be able to reach all computer control elements with their limbs while sitting in their seat. Suitable arrangements and semantic grouping of knobs and dials around the primary control area involving the steering wheel and the gas, break and clutch pedal are major issues of modern car design.
Augmented chemical reactions	AR, TUIs	A common problem is that different users perceive different gestures as natural or intuitive. For example shaking a tangible object may be interpreted as removing the current molecule from the tangible by one user, or as moving to the next selected atom in the molecule by another user.
Intelligent welding gun	Mobile, TUIs	Apart from welding the studs, the only user actions concern setup and selection of welding scenarios as well as the inspection of previously recorded data. For deployment in an industrial environment, this interface was kept simple and robust. It has been in industrial use for several years; the welder(s) did not report problems with the required human action.
Terrain exploration	TUIs, VR	Flying a toy airplane is, in principle, well understood by most people. Yet, the relationship between the physical actions (recorded in a 6 DoF space) to the 4 DoF flight control of an airplane is not easy to grasp. Experiments show that it requires significant physical talent and gaming experience to isolate the important parameters from the redundant ones.

upper left arrow. It takes a user’s planned action and transforms it into a physically measurable action, suitable to the input devices of the system. Table 6 describes human action issues for the exemplary post-WIMP interfaces that were presented in Sect. 2.

Even when users have decided what action to take, it requires skill, experience and dexterity to actually perform the necessary physical action, depending on the

input devices. The action may need to be executed with specific speed or precision, e.g. when typing on a keyboard, pointing (double clicking) with a mouse, pen or finger, speaking a command, looking at an object (glance control), or performing a free-form 3D gesture. Proper execution may require high reactive skills, good gross or fine motor control or even a well-developed sense of balance or rhythm (e.g. when interacting via a balance board in a sports game).

Users may not be fully aware of the assumptions and requirements of the input system, regarding the speed and precision for the intended actions to be recognizable by the system. Furthermore, even if the requirements are clear, it is not always easy for users to act according to the input specifications—or it may be uncomfortable or straining for them to perform the actions.

Depending on the application, this may be a thrilling and interesting challenge (games of skill), a frustrating hinderance (e.g. in office applications), or a potential source of danger (e.g. in safety-critical situations). Such issues are topics in ergonomics and human factors research (Shneiderman and Plaisant [13]).

Keyboard layouts and pointing devices have been analyzed, regarding users' ability to produce fast and/or precise input. As a prominent example, Fitts' law describes a relationship between the size of a target and its distance from a user's current pointing position (Fitts [33]): the larger a target, the faster can users move across a long distance to hit it easily. Vice versa, the shorter the distance to the target, the smaller can the target be—an essential aspect for designing layouts of icons on desktop-style graphical user interfaces. The GOMS model (Card, Moran, and Newell [27]) describes user interaction as an interplay of goals, operators, methods, selection rules. It was designed to help dividing interactive tasks into series of small actions in order to predict the time required to perform complex tasks. For example, this has been done for typing, using the keystroke-level model (KLM) (Card, Moran, and Newell [26]).

The so-called QWERTY keyboard⁷ is a negative example: more than a century ago, the arrangement of keys was not designed to improve humans' typing speed but rather to keep the physical hammers from jamming.

There are also many evaluations of pointing devices. Critical distinctions exist between the concepts of direct pointing/touching versus indirect pointing. Direct pointing and touching, e.g. with a pen, with one's fingers on a multi-touch surface, or in augmented reality and tangible interaction, provides users with a direct association between their action and the visual object/icon that they are manipulating. In its purest form, the performed action has a one-to-one mapping to the intended manipulation, such as moving, rotating or enlarging a virtual photo that is shown on an interactive table, or manipulating a physical object. Yet, fingers or pens may not provide sufficient precision and accuracy when selecting very small objects—probably within a densely populated neighborhood of further objects—and/or when intending to perform minuscule manipulations. Indirect pointing e.g. with a computer mouse, on the other hand, allows much more precise selection and

⁷Called so due to its arrangement of keys in the upper row: Q-W-E-R-T-Y (English version).

manipulation—especially when they can be performed in conjunction with sufficiently large widgets such as scroll bars and dials. Yet, the direct association between user action and resulting object manipulation is missing. Users have to familiarize themselves with a mismatch of the position and direction of their physical manipulation with respect to consequences in the virtual computer world. For example, novice users of a computer mouse (such as very young children) have been observed lifting the mouse vertically upwards (rather than pushing it horizontally on a table) when trying to control upward cursor motion on the vertical screen of a desktop monitor.

Another critical issue are the dimensions of the physical interaction space versus the virtual world. Direct manipulation such as multi-touch interaction cannot work in locations (e.g. on very high screens on a wall) that users cannot reach. Furthermore, even if a location can be reached, users may not always want to move across extended distances. To this avail, rate control and non-uniform mappings between user action and virtual interpretation have been established (Bowman et al. [2], Shneiderman and Plaisant [13]).

5 Testing Issues

The previous Sects. 3 and 4 have presented and discussed a large number of issues and uncertainties pertaining to the design and the implementation of suitable interfaces for human-computer interaction. There is a huge parameter space of design options with many alternatives. At the outset, it is not clear which design choice is better than another one—or even optimal with respect to some criterion. Moreover, criteria and options may change over time, due to improving computational, sensing and presentation facilities of computers, as well as due to evolving cultural backgrounds on the human side regarding the ease of understanding upcoming interaction metaphors.

For each interaction concept, there exists the risk of misunderstanding and misinterpretation. Depending on the application, such miscommunication may be a challenge, a nuisance, or a source of danger, bearing potential harm to the user and/or the environment. Independently of the severity of the consequences, it is mandatory for the design of human-computer interaction systems to be accompanied by dedicated evaluation procedures, from project conception to product delivery in a *user-centered design process*. The evaluations are typically conducted empirically, using hypothesis-based testing procedures with a specified level of significance and associated alpha and beta errors (Sirkin [53], Swan, Ellis, and Adelstein [55]).

5.1 Evaluation Design

During the entire process of conceiving, building and finalizing a human-computer interaction concept, the current state of the design and implementation needs to undergo continuous evaluation. This is not a one-step task. Rather, design, prototypical

implementation, evaluation and re-design build upon one another in ever-continuing circles (Bowman et al. [2], Shneiderman and Plaisant [13], Chandler and Chandler [28]).

Yet, the test designs may change over time, depending on the maturity of the system, as well as on the urgency of obtaining a preliminary appreciation of vague ideas vs. an in-depth comparison of well-thought-through metaphors or devices.

5.1.1 Strategies and Methods for Different Process Phases

A number of different evaluation approaches exist that are appropriate in different phases of a project and/or serve different evaluation strategies (Bowman et al. [2], Shneiderman and Plaisant [13]). Evaluation designers use a palette of different approaches during the course of the project.

In the very early, conceptual phase of a project and while the very first prototypes are being built, first evaluations and feedback are often acquired via *expert reviews*: the ideas and concepts are presented to a small number of experts at the example of use cases, e.g. by walking them mentally through the intended interaction processes or by demonstrating a rudimentary prototype. Feedback is gathered via questionnaires or interviews. If possible, experts will also be asked for heuristic evaluations, relating the current ideas to known guidelines and cases of best practice in the field. Such early feedback is valuable in cutting back on ideas that experts can quickly identify as unsuitable, based on their background expertise.

When early prototypes become available, *usability testing* becomes an option. Initial tests are typically conducted as *formative evaluations*, involving only few test persons and investigating only a small, well-selected subset of issues to form the base for a consistent interface design. As with expert reviews, designers can retrieve quick and very valuable feedback from such small evaluations: typically, very few test runs suffice to indicate the initial, major issues that need to be improved (Schwerdtfeger [52]). At later project phases—especially shortly before release, more substantial *summative evaluations* are conducted to sum up thorough comparisons of all options. The next Sect. 5.2 presents usability testing in detail. Usability tests are typically surrounded by demographic and subjective questionnaires, as well as closing interviews. Those are the topic of Sects. 5.3 and 5.4. In principle, the entire design space needs to be evaluated at this point. Yet, some simplifications are typically made for the sake of reduced complexity (Chandler and Chandler [28]).

At a later, more mature phase, larger target groups of users are also increasingly involved via *user surveys* and *acceptance tests*. A good example is the early release of beta-versions of computer systems, e.g. before the roll-out of a new game (Chandler and Chandler [28]).

Finally, after product release, feedback is gathered online, e.g. in newsgroups, as well as via telephone call centers, further acceptance tests, and user surveys.

5.1.2 Evaluation Criteria

A number of different metrics can be used to compare and evaluate the quality of different designs (Bowman et al. [2]).

For computer systems, system performance, such as the average frame rate or latency, or the network delay is of utmost importance. For visual systems, optical distortions of cameras and displays, the provided field of view and the resolution are further important evaluation criteria. These can be measured and compared without much user involvement. Yet, they also need to be considered in the context of user-centered evaluations since different performance metrics may have a large impact on the user-based test results.

The next set of evaluation criteria focuses on task performance: how fast can a user reach a specific location? What accuracy is being achieved? How many errors do users make when selecting or manipulating an object? Further criteria are the speed of learning a concept, the spatial awareness a user has gained when interacting with objects in a three-dimensional space and the degree of distraction induced by the system (e.g. by analyzing users' eye movements: when did they look where?). These are metrics that can be measured objectively, using automatic procedures. Data is collected during a test run and stored for subsequent statistical analysis. These criteria are describing the pragmatic quality (PQ) of a user interface, i.e. its effectiveness and efficiency.

The final set of evaluation criteria deals with subjective metrics involving user satisfaction—the so-called hedonic quality (HQ) of a system (Hassenzahl, Kekez, and Burmester [36]). In questionnaires, users are asked to describe their perceived ease of use, ease of learning and their satisfaction during the interaction process on a given scale. Further parameters, related to novel three-dimensional user interfaces are related to users' sense of presence in a virtual environment, and their degree of comfort (simulator sickness), pertaining to the elaborations in Sect. 4.1 on user accommodation, adaptation, and potential after effects. How long does it take users when they subjected to optical illusions (sensory mismatches) during an experiment to re-adapt to the true physical interpretation of their senses after an experiment?

5.2 Usability Testing

Usability testing has gained much attention and importance. It represents the attempt to parameterize all important issues of a human-computer interaction approach systematically, describing them as a set of factors (dimensions). If these factors are independent, and if there are no additional, confounding factors, different approaches can be compared by letting a sufficiently large group of representative users interact with the computer in all different variants.

In order for such testing to be successful (i.e. to produce significant results), great care has to be taken to design a good test plan. In the following, several aspects of the physical environment, the underlying concept, and the established process and experimental structure are presented.

5.2.1 Test Setup in a Usability Lab

A usability lab typically consists of two areas that should be separated from each other as much as possible.

The first area is set up for the test person to interact with the system, as designed in the test plan. The environment may be as simple as a desktop monitor with WIMP-style interaction devices, or as complicated as an immersive, multimedia, three-dimensional driving or flight simulator, possibly even integrated into a motion platform. It may also be mobile, e.g. integrated into a real car driving in real traffic, or a mobile phone in pedestrian applications, such as in an augmented reality context. In all cases, the setup should be as realistic as possible, and the test person should be disturbed as little as possible while the experiment is running. The test area should be instrumented with extra recording equipment, such as microphones, cameras, eye trackers etc—in order to store as much information as possible about the course of the experiments and especially about the users' actions, reactions, gestures, mimics and side remarks. Such data can be invaluable during the post-analysis step when questions arise because a particular experiment has unusual results (i.e., outliers).

A second area is arranged for the person in charge of running the experiment. The experimenter should not influence the test person. Thus, the areas should be separated—ideally by a wall with a semi-transparent window.

Contrary to these well-established standards, the evaluation of novel user interfaces (e.g. for augmented reality) may require arrangements that conflict with prior guidelines of best practise. Schwerdtfeger argued in his dissertation that, for AR-based user interfaces in a logistics application, it was more reasonable to interrupt test persons when they consistently went astray than to let them fail during the entire experiment—since the reason for such errors was often related to poor calibrations of the optical-see-through display on their heads or to a basic misunderstanding of some aspect of the very novel hardware and interaction metaphors they were exposed to Schwerdtfeger [52].

5.2.2 Process of Collecting the Data

When preparing and conducting user tests, utmost care has to be taken to apply proper procedure—such that the results are not unnecessarily tainted and thereby rendered unusable. It is extremely difficult and costly to rerun an experiment: test persons will not react the same when they are exposed to the same interface a second time. Acquiring new test persons is time consuming and difficult.

Thus, much care must be taken during the planning phase of the experiment. All potential aspects that might have an influence have to be identified and either explicitly discarded from the test design or accounted for as one of the parameters under evaluation (see Sect. 5.2.3).

During the experimental part, proper procedure has to be set up and executed for each test person. A well-established procedure consists of greeting and introducing each newly arriving test person to the test setup in a predefined way (possibly

using rehearsed sentences) such as not to bias persons at this stage. Typically test persons fill out a demographic questionnaire requesting information about general human factors (age, sex, ...) as well as special factors (color-blindness, familiarity with novel user interfaces or games, ...). The test in itself may consist of one or several parts, exposing test persons to different kinds of user interfaces or to different scenarios. Inbetween, further questionnaires may ask people about subjective impressions (see Sect. 5.3). The experimental session closes after the last set of experiments—possibly with a further subjective questionnaire and/or with a standardized or open interview (see Sect. 5.4).

After all test persons have participated in the experiment, the collected data is analyzed, filtered, and subjected to statistical analysis tools for hypothesis testing (Sirkin [53]). The results are compiled into a report.

5.2.3 Multi-factorial Design

As stated in Sect. 5.2.2, the design of a usability test has to account for all parameters that have a potential impact on the results. Multiple parameters are modeled by *multi-factorial design* approaches (Mukerjee and Wu [10]).

One or more measurement functions $t_d = f_d(x, y, z, \dots)$ are established that describe criteria listed in Sect. 5.1.2, t_d , as functions f_d of parameters x, y, z, \dots . The metrics t_d are *dependent variables* since they are the result of running testing with respect to varying x, y, z, \dots . The parameters x, y, z, \dots are independent variables—so-called *factors*. Each can assume values—also called *levels*—within a predefined range.

When testing a user interface with respect to metric t_d , all independent variables need to be checked with respect to all their levels. Thus, the design space of the user interface, with respect to the given evaluation criterion, is the cross product of all factors. Its cardinality is the product of the cardinality of all level ranges: $\|T_d\| = \|X\| \times \|Y\| \times \|Z\| \times \dots$. In a practical example, this means: test designers want to compare a novel multi-touch interface to a traditional mouse-based interface. This results in an independent variable UI-TYPE with levels MOUSE and MULTI-TOUCH. At the same time, the designers want to explore the benefit of sound and thus introduce an independent variable SOUND with levels SOUND-ON and SOUND-OFF. This creates $2 \times 2 = 4$ variants of user interfaces that need to be compared to one another in a statistical test procedure. If the designers were to include one more UI-TYPE level, PEN, the space of UI variants would extend to $2 \times 3 = 6$ variants. Evaluations have to test each of these variants in their experiments.

In addition to these *planned factors*, experiments may also be subject to unwanted—yet unavoidable—further factors, so-called *confounding factors*. Examples are learning effects, user fatigue or simulator sickness. This means that, if test persons are requested to participate in experiments for more than one variant, the sequencing of the variants may have an impact on the results since test persons may learn something about the scenario in the first test run (e.g. about the traffic situation in a driving simulation) that they can exploit during the second test run

with a second variant of the user interface. They may thus perform better in the second test—not due to a superior user interface but due to learning effects. Conversely, fatigue or simulator sickness may have a negative impact on the results that also needs to be discounted: test persons may perform better in the first run than in subsequent runs.

To discount the effects of such confounding factors, the sequencing of the evaluations for different variants needs to be permuted between different test persons—requiring $n!$ different sequencing plans for n variants. For the given example of 6 variants, this means that $6! = 720$ test persons are needed for the evaluations, one for each permutation of the 6 variants. Without level PEN for factor UI-TYPE, only $4! = 24$ permutations need to be compared. This small example exemplifies the critical impact of introducing more levels to a factor: the resulting design space rapidly explodes to unmanageable numbers of user interface variants that all need to be compared systematically in the test design, requiring rapidly increasing numbers of test persons.

5.2.4 Experimental Structure

Critical to successful evaluations is the proper selection of test persons. These should correspond to the population of the targeted final users of an application. If the new user interface is expected to be helpful across many applications, test users must be drawn from a wide, diverse background. In most cases, it is not reasonable to recruit test persons only from the immediate, close circle of friends and co-workers since such group might be rather homogeneous regarding age, education, sex and experience with computers. On the other hand, some initial problems with a novel user interface might be so universal that they are criticized by nearly everybody—except for the developer of the interface. In such cases, first formative usability tests may be conducted with colleagues and friends who are more easily accessible than a non-biased, well-balanced broad group of representative test persons. When reporting on an evaluation, it is critical to describe the demographic constitution of the selected test group and to present the rationale why these people were selected.

As discussed in the previous Sect. 5.2.3, several variants of a user interface need to be compared. To this end, all variants have to be tested with the same depth, i.e.: test persons have to be organized in groups for each such variant. Two approaches exist for organizing such groups of test persons.

In a so-called *between-subject test design*, test persons are assigned to different groups, with each group testing exactly one user interface. This approach has the advantage that no learning effects of fatigue can occur since users participate in only one evaluation. A disadvantage of this approach is the need to balance all test groups such they all have a demographically similar distribution with respect to age, sex, etc. To ensure well-balanced (i.e., *unbiased*) test groups, a large number of test persons are required.

Alternatively, in a so-called *within-subject test design*, each test person is asked to work with all variants. In this setup, demographic bias is not as much of an issue—especially for initial formative evaluations. Yet, confounding factors such as learning or fatigue are a considerable problem (see Sect. 5.2.3). In order to discount biases due to confounding factors, the variants need to be presented to different test persons in permuted order. Still, the problem remains that test persons may be overly strained and exasperated from very long series of experiments (possibly with interleaved subjective questionnaires). Thus, the design of the test procedure per variant should be kept as short as possible.

5.3 Subjective Evaluations

When working with a particular user interface, *objective* measurements of user performance are generally accompanied by questionnaires. Increasingly, such *subjective* user feedback, is becoming more and more essential to the success of a novel product or computer system. Concepts such as user satisfaction and user experience are becoming central issues in user interface design.

Yet, it is not easy to measure subjective feedback from test persons. Research in psychology and human factors has established a number of standardized questionnaires that have been the result of thorough investigations how to pose questions such that individually differing degrees of emotions can be discounted.

The *NASA Task Load Index (TLX)* (Hart and Staveland [34]) measures mental workload. To this end, users are asked to grade the mental, physical and temporal demand of the system, as well as their appreciation of their own performance, and the amount of effort and frustration they experienced during the test. For each of these six criteria, test persons are asked to indicate their rating on a 20 point scale ranging from VERY LOW to VERY HIGH.

The *System Usability Scale (SUS)* (Brooke [25]) determines the effectiveness, efficiency and satisfaction of test persons working with a particular user interface. Users are requested to comment on ten standardized statements on a 5 point scale ranging from STRONGLY DISAGREE to STRONGLY AGREE, resulting in a score in the range from 0 to 100.

Finally, a method using a *Semantic Differential* (Osgood, Suci, and Tannenbaum [47]) and the *AttrakDiff* (Hassenzahl, Burmester, and Koller [35]) test use a list of opposing attribute pairs with a 5 point scale to elicit indications from the test persons, regarding ideas and affective attitudes which are associated with an interface.

5.4 Interviews and Anecdotal Use

Despite all attempts towards gathering objective or subjective measurements from test persons that can be quantitatively evaluated, verbal feedback and anecdotal usage are invaluable forms of in-depth information. Especially in the early phases of

developing a novel human-computer interaction concept, a large number of issues are undefined. It is generally impossible to properly design a test setup that covers all of these issues. Careful inspection and recording of the test persons' every move and interrogating them about any observed moment of confusion or irritation sheds light on a sizeable number of essential problems that need to be mastered before large-scale systematic tests lead to conclusive results.

To this end, interviews are conducted in all stages of interface design and implementation. They can be associated with presentations, demonstrations and exhibits during expert reviews, small-scale usability studies, or random quests for user feedback. Interviews can be completely unstructured. Yet, already known issues are also cast into a systematic sequence of questions to be answered as part of the interviews by every interviewee.

6 Food for Thought

This chapter has reported on a large number of things that can go wrong in human-computer communication. Just as in human to human communication, there is much potential for misunderstanding. Humans can misinterpret computer presentations and animations, due to misleading metaphors or simply due to misled attention and overload. Conversely, machines can misinterpret human input commands, due to noisy sensor data or due to unprecisely performed human actions.

In conversations between humans, we are aware—to some extent—of potential misunderstandings, and we thus also communicate on a meta-level about the course of conversation with one another. We question, ascertain and reassure that the important issues have been clearly conveyed. We also express level of completeness when we simplify complicated matters for didactic matters such that the communication partner can comprehend an issue gradually over time. How can computer interfaces communicate on such a meta level, in parallel to conducting the principle exchange of information? This a topic of increasing importance, pertaining to issues of uncertainty analysis and uncertainty visualization.

Another important issue covers conscientious resource management—both on the human side and for the computer. Neither one has unlimited resources to perceive, interpret and act/present. There may be shortages of sensing/perception power, as well as memory and processing capacity. Human-computer interaction systems need to be aware of such resource limitations. Communication processes need to take explicitly into account that a communication partner may be currently overwhelmed by information and that it is, thus, better, to slow down and maybe even to keep quiet for a while. How can computers monitor resource shortages and adapt their communication strategy accordingly? This in another topic that is requires increasing attention.

7 Summary

This chapter has presented risks and issues of potential miscommunication on the basis of the interaction cycle by Bowman et al. [2]. For each step along this cycle, the chapter has discussed a number of critical issues and related them in associated tables to experiences that were made in the FAR-lab for when building and evaluating novel user interfaces for a number of applications. The chapter closes with a presentation of the most critical issues for planning proper test designs to evaluate novel human-computer interaction concepts in a human-centered approach.

Acknowledgements The discussion on user interfaces in this chapter is the result of nearly two decades of work on novel user interfaces related in particular to Augmented Reality. Many people have participated in the work and have influenced the authors in many ways. The authors (particularly the first author) would like to thank the many current and past members of the Fachgebiet Augmented Reality (FAR), as well as the labs at the Fraunhofer Institute for Computer Graphics (IGD) and the European Computer-industry Research Centre (ECRC) prior to FAR for sharing and jointly developing many of the ideas. The authors are grateful to many collaborators across a large number of publicly or privately funded research projects, such as TUMMIC, Presencia, KAUST, Crumbs, FORLOG, Speedup, IGSSE, Trackframe, the Chair for Computer-Aided Medical Procedures (CAMP) and the Chair for Applied Software Engineering.

References

Selected Bibliography

1. R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, B. MacIntyre, Recent advances in augmented reality. *IEEE Comput. Graph. Appl.* **21**, 34–47 (2001). doi:[10.1109/38.963459](https://doi.org/10.1109/38.963459). <http://portal.acm.org/citation.cfm?id=616073.618862>
2. D.A. Bowman, E. Kruijff, J.J. LaViola, I. Poupyrev, *3D User Interfaces: Theory and Practice* (Addison-Wesley/Longman, Redwood City, 2004). ISBN 0201758679
3. F.P. Brooks Jr., The computer scientist as toolsmith II. *Commun. ACM* **39**, 61–68 (1996). <http://doi.acm.org/10.1145/227234.227243>
4. G.C. Burdea, P. Coiffet, *Virtual Reality Technology*, 2nd edn. (Wiley, Hoboken, 2003). ISBN 0471360899
5. T. Coskun, A. Benzina, E. Artinger, C. Binder, G. Klinker, User-centered development of UI elements for selecting items on a digital map designed for heavy rugged tablet PCs in mass casualty incidents, in *Proceedings of ACM SIGHIT International Health Informatics Symposium (IHI 2012)* (ACM, New York, 2012)
6. M.W. Eysenck, M.T. Keane, *Cognitive Psychology*, 6th edn. (Psychology Press, New York, 2010). ISBN 978-1-84169-539-6
7. R.L. Gregory, *Eye and Brain, the Psychology of Seeing*, 5th edn. (Princeton University Press, Princeton, 1997)
8. M. Huber, Parasitic tracking for augmented reality. Dissertation, Technische Universität München, München (Nov. 2011)
9. H. Ishii, B. Ullmer, Tangible bits: towards seamless interfaces between people, bits and atoms, in *CHI 97* (1997), pp. 234–241. citeseer.ist.psu.edu/ishii97tangible.html
10. R. Mukerjee, C.F.J. Wu, *A Modern Theory of Factorial Design*. Springer Series in Statistics (Springer, Heidelberg, 2006). ISBN 978-0387319919

11. P.G. Neumann, *Computer Related Risks* (ACM Press/Addison-Wesley, New York, 1995). ISBN 0-201-55805-X
12. D.A. Norman, *The Design of Everyday Things*, 2nd edn. First Basic (Perseus Books Group, Jackson, 2002). ISBN 0465067107
13. B. Shneiderman, C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 4th edn. (Pearson, Boston, 2005). ISBN 0321197860. <http://www.gbv.de/dms/ilmenua/toc/492668051.PDF>
14. R. Spence, *Information Visualization: Design for Interaction*, 2nd edn. (Prentice-Hall, Upper Saddle River, 2007). ISBN 0132065509
15. M. Tönnis, Towards automotive augmented reality. Dissertation, Technische Universität München, München (Nov. 2008)
16. A. van Dam, Post-WIMP user interfaces. *Commun. ACM* **40**, 63–67 (1997). <http://doi.acm.org/10.1145/253671.253708>
17. M. Weiser, Ubiquitous computing. *Computer* **26**, 71–72 (1993). <http://doi.ieeecomputersociety.org/10.1109/2.237456>
18. G. Welch, E. Foxlin, Motion tracking: no silver bullet, but a respectable arsenal. *IEEE Comput. Graph. Appl.* **22**(6), 24–38 (2002)
19. C.D. Wickens, J.G. Hollands, *Engineering Psychology and Human Performance*, 3rd edn. (Prentice-Hall, Upper Saddle River, 2000)

Additional Literature

20. E. Artinger, T. Coskun, S. Nestler, M. Mähler, Y. Yildirim-Krannig, F. Wucholt, F. Echlter, G. Klinker, Creating a common operation picture in realtime with user-centered interfaces for mass casualty incidents, in *Proceedings of the 4th International Workshop for Situation Recognition and Medical Data Analysis in Pervasive Health Environments (PervaSense), PervaSense'12* (2012). ICST.org
21. M. Bauer, Tracking errors in augmented reality. Dissertation, Technische Universität München, München (Sept. 2007)
22. R. Bauernschmitt, M. Feuerstein, J. Traub, E.U. Schirmbeck, G. Klinker, R. Lange, Optimal port placement and enhanced guidance in robotically assisted cardiac surgery. *Surg. Endosc.* **21**(4), 684–687 (2007). doi:[10.1007/s00464-006-9057-z](https://doi.org/10.1007/s00464-006-9057-z)
23. B.B. Bederson, B. Shneiderman, *The Craft of Information Visualization: Readings and Reflections* (Morgan Kaufmann, San Francisco, 2003). ISBN 1558609156
24. A. Benzina, M. Tönnis, G. Klinker, M. Ashry, Phone-based motion control in VR: analysis of degrees of freedom, in *CHI Annual Conference on Human Factors in Computing Systems. Extended Abstracts (CHI EA)'11* (ACM, New York, 2011), pp. 1519–1524. ISBN 978-1-4503-0268-5. <http://doi.acm.org/10.1145/1979742.1979801>
25. J. Brooke, System usability scale (SUS): a quick-and-dirty method of system evaluation user information. Technical Report, Digital Equipment Corporation Ltd., Reading, UK (1986)
26. S.K. Card, T.P. Moran, A. Newell, The keystroke-level model for user performance time with interactive systems. *Commun. ACM* **23**(7), 396–410 (1980)
27. S.K. Card, T.P. Moran, A. Newell (eds.), *The Psychology of Human-Computer Interaction* (CDC Press, San Francisco, 1983)
28. H.M. Chandler, R. Chandler, *Fundamentals of Game Development* (Jones and Bartlett, Boston, 2009)
29. A. Dey, A. Cunningham, C. Sandor, Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments, in *3DUI* (2010), pp. 127–128
30. F. Echlter, Tangible information displays. Dissertation, Technische Universität München, München (Nov. 2009)

31. F. Echter, F. Sturm, K. Kindermann, G. Klinker, J. Stilla, J. Trilk, H. Najafi, The intelligent welding gun: augmented reality for experimental vehicle construction, in *Virtual and Augmented Reality Applications in Manufacturing*, ed. by S. Ong, A. Nee (Springer, New York, 2003), Chap. 17
32. S. Feiner, B. MacIntyre, M. Haupt, E. Solomon, Windows on the world: 2D windows for 3D augmented reality, in *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology, UIST '93*, New York, NY, USA (ACM, New York, 1993), pp. 145–155. ISBN 0-89791-628-X. <http://doi.acm.org/10.1145/168642.168657>
33. P.M. Fitts, The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* **47**(6), 381–391 (1954)
34. S.G. Hart, L.E. Staveland, Development of NASA-TLX task load index: results of empirical and theoretical research, in *Human Mental Workload*, ed. by P.A. Hancock, N. Meshkati (North Holland, Amsterdam, 1988)
35. M. Hassenzahl, M. Burmester, F. Koller, AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität, in *Mensch und Computer 2003: Interaktion in Bewegung*, ed. by G. Szwillus, J. Ziegler (Teubner, Leipzig, 2003)
36. M. Hassenzahl, R. Kekez, M. Burmester, The importance of a software's pragmatic quality depends on usage modes, in *Proceedings of the 6th International Conference on Work with Display Units (WWDU'02)*, Berlin, Germany (ERGONOMIC Institut fuer Arbeits- und Sozialforschung, Berlin, 2003), pp. 275–276
37. K.V. Iserson, J.C. Moskop, Triage in medicine, part I: concept, history, and types. *Ann. Emerg. Med.* **49**(3), 275–281 (2007)
38. J.A. Jones, J.E.I. Swan, G. Singh, S.R. Ellis, Peripheral visual information and its effect on the perception of egocentric depth in virtual and augmented environments, in *VR (2011)*, pp. 215–216
39. P. Keitler, Management of tracking and tracking accuracy in industrial augmented reality environments. Dissertation, Technische Universität München, München (Apr. 2011)
40. T. Luhmann, Accuracy limits in photogrammetry, in *Proceedings of the Workshop Traceability in Large Scale Metrology (2006)*
41. A. MacWilliams, C. Sandor, M. Wagner, M. Bauer, G. Klinker, B. Brügge, Herding sheep: live system development for distributed augmented reality, in *ISMAR (2003)*, pp. 123–132
42. P. Maier, M. Tönnis, G. Klinker, A. Raith, M. Drees, F. Kühn, What do you do when two hands are not enough? Interactive selection of bonds between pairs of tangible molecules, in *Proceedings of the 5th IEEE Symposium on 3D User Interfaces (3D UI)* (2010), pp. 83–90
43. B. Myers, S.E. Hudson, R. Pausch, Past, present, and future of user interface software tools. *ACM Trans. Comput.-Hum. Interact.* **7**, 3–28 (2000). <http://doi.acm.org/10.1145/344949.344959>
44. S. Nestler, Konzeption, Implementierung und Evaluierung von Benutzerschnittstellen für lebensbedrohliche, zeitkritische und instabile Situationen. Dissertation, Technische Universität München, München (July 2010)
45. G.M. Nielson, H. Hagen, H. Müller, *Scientific Visualization: Overviews, Methodologies, and Techniques* (IEEE Comput. Soc., Los Alamitos, 1997)
46. D.A. Norman, *The Invisible Computer* (MIT Press, Cambridge, 1998). ISBN 0262140659
47. C.E. Osgood, G.J. Suci, P.H. Tannenbaum, *The Measurement of Meaning* (University of Illinois Press, Champaign, 1957)
48. D. Pustka, M. Huber, C. Waechter, F. Echter, P. Keitler, J. Newman, D. Schmalstieg, G. Klinker, Automatic configuration of pervasive sensor networks for augmented reality. *IEEE Pervasive Comput.* **10**(3), 68–79 (2011). <http://doi.ieeecomputersociety.org/10.1109/MPRV.2010.50>
49. J.P. Rolland, H. Fuchs, Optical versus video see-through head-mounted displays in medical visualization. *Presence* **9**(3), 287–309 (2000)
50. C. Sandor, A software toolkit and authoring tools for user interfaces in ubiquitous augmented reality. Dissertation, Technische Universität München, München (Oct. 2005)

51. C. Sandor, G. Klinker, A rapid prototyping software infrastructure for user interfaces in ubiquitous augmented reality. *Pers. Ubiquitous Comput.* **9**(3), 169–185 (2005)
52. B. Schwerdtfeger, Pick-by-vision: bringing HMD-based augmented reality into the warehouse. Dissertation, Technische Universität München, München (July 2010)
53. R.M. Sirkin, *Statistics for the Social Sciences*, 3rd edn. (SAGE, Thousand Oaks, 2006)
54. F. Steinicke, G. Bruder, K. Hinrichs, P. Willemsen, Change blindness phenomena for stereoscopic projection systems, in *Proc. IEEE Virtual Reality Conference (VR'10)* (2010), pp. 187–194
55. J.E.I. Swan, S.R. Ellis, A.B. Adelstein, Conducting human-subject experiments with virtual and augmented reality, in *Tutorial at the IEEE Virtual Reality Conference (VR'07)* (2007). <http://www.cse.msstate.edu/~swan/teaching/tutorials/Swan-VR2007-Tutorial.pdf>
56. S. Thrun, W. Burgard, D. Fox, *Probabilistic Robotics* (MIT Press, Cambridge, 2006)
57. M. Tönnis, A. Benzina, G. Klinker, Utilizing consumer 3D TV hardware for a flexibly reconfigurable visualization system. Technical report TUM-I-11-13, Technische Universität München (2011)
58. M. Tönnis, R. Lindl, L. Walchshäusl, G. Klinker, Visualization of spatial sensor data in the context of automotive environment perception systems, in *Proceedings of the 6th International Symposium on Mixed and Augmented Reality (ISMAR)* (2007)
59. E.R. Tufte, *The Visual Display of Quantitative Information*, 2nd edn. (Graphics Press, Cheshire, 2001). ISBN 0961392142. <http://www.amazon.com/Visual-Display-Quantitative-Information-2nd/dp/0961392142%3FSubscriptionId%3D192BW6DQ43CK9FN0ZGG2%26tag%3Dws%26linkCode%3Dxm2%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0961392142>