34

Designing Multimedia Applications for User Experience

34.1 Introduction .................................................34-1
34.2 Media Properties and Design Principles ..................34-2
34.3 Media Selection and Combination ..........................34-4
    Information Architecture • Matching Information to Media
34.4 UX and Multimedia ........................................34-8
    Interaction and Flow • Presence • Engagement and Social Presence
34.5 Components of User Engagement .........................34-9
    Media for Emotional Effects • Multimedia and Aesthetic Design • Metaphors and Interaction Design
34.6 Media Design for User Engagement .......................34-12
34.7 A Brief Example ...........................................34-16
34.8 Conclusions .................................................34-16
References .........................................................34-17

Alistair Sutcliffe
University of Manchester

34.1 Introduction

Multimedia is rapidly becoming the default in most applications apart from transaction processing systems where numbers and text dominate, although multimedia has also been used extensively in task-based applications for process control and safety-critical systems (Hollan et al., 1984; Alty, 1991). With the advent of the Web 2.0 and beyond, interactive multimedia is a continuing design challenge. Note the “interactive” in multimedia; this is important since the objective is not only to deliver information through a variety of media (text, speech, image, video, music, etc.), but also to engage the user in simulating and exciting interaction. The view that the user interface (UI) design goes beyond functional and usability into user experience (UX) is well established in the Human–Computer Interaction (HCI) community (Hassenzahl and Tractinsky, 2006). In this chapter, the author takes a UX-oriented view of multimedia, so interaction will be treated as one medium in the multimedia design space.

This chapter has four aims:

1. Describe the properties of media resources and how they are used in design.
2. Propose design guidelines for conveying information in multimedia.
3. Explain the concept of UX and its relationship to interactive multimedia.
4. Propose multimedia design guidelines for UX.
The first part of the chapter will cover the design process for multimedia applications; starting with an information analysis, it then progresses to deal with issues of media selection and integration. A more detailed description of the process and information modeling techniques is given in Sutcliffe (2003, 2012). Multimedia design involves several special areas that are technical subjects in their own right. For instance, design of text is the science (or art) of calligraphy that has developed new fonts over many years; visualization design encompasses the creation of images, either drawn or captured as photographs. Design of moving images, cartoons, video, and film are further specializations, as are musical composition and design of sound effects. Multimedia design lies on an interesting cultural boundary between the creative artistic community and science-based engineering. One implication of this cultural collision is that space precludes “within-media” design being dealt with in depth in this chapter, i.e., guidelines for design of one particular medium.

The second part of the chapter broadens multimedia design to cover UX and design of applications where the goal is to persuade, entertain, or educate the user. Design in this section considers interaction design, aesthetics, and emotional effects of multimedia. UX design is a complex subject, so the background to this topic will only be reviewed briefly. For a more in-depth treatment of UX design, the reader is referred to Sutcliffe (2009).

### 34.2 Media Properties and Design Principles

We perceive multimedia using either our visual sense (text, photographs, drawings, animations), audio (speech, music, natural and artificial sounds), or both (film with a sound track). Haptic (touch), olfactory (smell), and gustatory (taste) senses may be more important in the near future; however, for the present, visual and audio senses dominate. How we perceive and understand multimedia is limited by cognitive constraints inherent in the way our brain processes information. These are summarized as follows; more in-depth explanation of these issues is given in Sutcliffe (2012):

- **Selective attention:** We can only attend to one input on one sensory channel at a time. Visually, this is a consequence of the way our eyes focus on images, on only one small area at a time. For audio, we can only pick out one message from the background noise; two or more separate sounds interfere, and we have to selectively attend to one or the other. We can overcome this limitation by time-slicing our attention between different images or sounds, but at the penalty of forgetting.

- **Working memory:** Information from our senses is perceived rapidly, but the understanding phase encounters a very limited bottleneck, working memory. Working memory is a very limited scratch pad or cache memory containing about five facts/ideas or concepts (chunks in the psychological jargon) at any one time. As information is received, it overwrites the current content of working memory, so we can only actively process a small quantity of information at once.

- **Knowledge integration:** We make sense of the world by seeing or hearing patterns, and those patterns are easier to understand when we can associate them with things we already know, our memory. This allows us to overcome some of the selective attention and working memory limitations. For example, when moving images and an audio sound track are integrated in a film, we can process information on two channels without difficulty.

- **Arousal and affect:** Any change in images or audio automatically demands our attention (see selective attention); however, change also alerts us, and this increases our arousal. Change in multimedia is therefore more interesting and possibly exciting. Content of multimedia can connect with our emotions as we react instinctively to scary images or loud noises.

The links between different types of media and these psychological implications are summarized in Table 34.1.

The psychological implications also motivate multimedia design principles (ISO, 1997, 1998). The principles are high-level concepts, which are useful for general guidance, but they have to be interpreted
Designing Multimedia Applications for User Experience

The following design principles are cross-referenced to the corresponding psychological constraints:

- **Thematic congruence.** Messages presented in different media should be linked together to form a coherent whole. This helps comprehension as the different parts of the message make sense by fitting together. Congruence is partly a matter of designing the content so that it follows a logical theme (e.g., the script or story line makes sense and does not assume too much about the user’s domain knowledge), and partly a matter of attentional design to help the user follow the message thread across different media (knowledge integration, selective attention).

- **Manageable information loading.** Messages presented in multimedia should be delivered at a pace that is either under the user’s control or at a rate that allows for effective assimilation of information without causing fatigue. The rate of information delivery depends on the quantity and complexity of information, the effectiveness of the design in helping the user extract the message from the media, and the user’s domain knowledge and motivation. Some ways of reducing information overload are to avoid excessive use of concurrent dynamic media and give the user time to assimilate complex messages. This guards against overloading user’s working memory. Since our working memory has a small capacity and gets continually overwritten by new input, controlling the pace and quantity of input can reduce overloading.

- **Compatibility with the user’s understanding.** Media should be selected that convey the content in a manner compatible with the user’s existing knowledge, for example, the radiation symbol and road sign icons are used to convey hazards and dangers to users who have the appropriate knowledge and cultural background. The user’s ability to understand the message is important for designed image media (diagrams, graphs) when interpretation is dependent on the user’s knowledge and background (knowledge integration).

- **Complementary viewpoints.** Similar aspects of the same subject matter should be presented on different media to create an integrated whole. Showing different aspects of the same object, for example, a picture and a design diagram of a ship, can help memorization by developing richer schema and better memory cues (knowledge integration).

### TABLE 34.1 Psychological Constraints and Implications for Multimedia Design

<table>
<thead>
<tr>
<th>Medium</th>
<th>Selective Attention</th>
<th>Working Memory</th>
<th>Arousal and Affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image: Photograph</td>
<td>Eye scanning depends on salient features on images</td>
<td>No detail of image persists; only a few facts can be extracted at once</td>
<td>Highlighting components increase arousal; other effects depend on content</td>
</tr>
<tr>
<td>Text</td>
<td>Eye scanning is set by text format and reading order</td>
<td>Only a few (four to five) facts can be extracted at once</td>
<td>Depends on content</td>
</tr>
<tr>
<td>Moving image: Film, cartoons</td>
<td>Attention is automatic; it is difficult to ignore dynamic media</td>
<td>Rapid overwriting; only high-level “gist” facts can be extracted</td>
<td>Dynamic media are more arousing than static media; content can be very effective in emotional reactions</td>
</tr>
<tr>
<td>Audio: Speech</td>
<td>Attention is automatic. Speech dominates over other sound</td>
<td>Rapid overwriting; only high-level “gist” facts can be extracted</td>
<td>Dynamic media are more arousing; content and delivery (voice tone) can be used for emotional effects</td>
</tr>
<tr>
<td>Audio: Other sounds</td>
<td>Attention is automatic</td>
<td>Facts depend on recognition, i.e., matching sound to memory</td>
<td>Audio has a range of arousal and emotional effects</td>
</tr>
</tbody>
</table>
• Consistency. It helps users learn an interface by making the controls, command names, and layout follow a familiar pattern. People recognize patterns automatically, so operating the interface becomes an automatic skill. Consistent use of media to deliver messages of a specific type can help by cueing users with what to expect (selective attention, knowledge integration).
• Interaction and engagement. These help understanding and learning by encouraging the user to solve problems. Memory is an active process. Interaction increases arousal, and this makes the user’s experience more vivid, exciting, and memorable (arousal and affect).
• Reinforce messages. Redundant communication of the same message on different media can help learning. Presentation of the same or similar aspects of a message helps memorization by the frequency effect. Exposing users to the same thing in a different modality also promotes rich memory cues (knowledge integration; see also complementary viewpoints).

Many of the principles are motivated by knowledge integration, which is one of the reasons for the “multi” in multimedia.

The following section provides some guidelines to address these issues.

### 34.3 Media Selection and Combination

Designers have to solve two problems:

1. Which media resources to select (or create) to convey information?
2. How to integrate media resources to convey information more effectively and improve UX?

Media resources may be classified as either static (do not change, e.g., text, images) or dynamic in the sense that they change during presentation (e.g., speech, all audio media, moving image media such as films, cartoons). Another distinction is also useful: realistic/not realistic, reflecting the designer’s involvement in creating the medium. For instance, a photograph or video may capture an image of the real world; alternatively, the hand of the designer may be more overt in drawings, sketches, and diagrams. Note that it is the content that determines the media resource rather than the presentation medium; for example, a diagram is still a diagram whether it has been photographed or drawn on paper and scanned. The design of media resources is important, because abstract concepts can only be conveyed by language (text/speech), designed audio (music), or by designed images (diagrams, cartoons).

More fine-grained taxonomic distinctions can be made, for instance, between different signs and symbolic languages (see Bernsen, 1994), but richer taxonomies increase specification effort. Media resources can be classified using the three criteria in the decision tree illustrated in Figure 34.1.

The approach to classifying media uses a decision tree “walkthrough” with the following questions that reflect the facets of the classification:

• Is the medium perceived to be realistic or not? Media resources captured directly from the real world will usually be realistic, for example, photographs of landscapes, sound recordings of bird song, etc. In contrast, nonrealistic media are created by human action. However, the boundary case category that illustrates the dimension is a realistic painting of a landscape.

• Does the medium change over time or not? The boundary case here is the rate of change, particularly in animations where some people might judge 10 frames/s to be a video but 5 slides in a minute shown by a PowerPoint presentation to be a sequence of static images.

• Which modality does the resource belong to? Most media are visual (image/text) or audio (music/speech), although one resource may exhibit two modalities. For example, a film with a sound track communicates in both visual and audio modalities.

The walkthrough process informs selecting appropriate media resources to convey information content. Information content may be apparent from the resources provided, such as diagrams, text, etc; however, in some cases, only high-level user requirements may be given, such as “present
information to influence people to adopt a healthy diet, give up smoking, or purchase a product from an e-commerce website.” The designers will have to source the content if it is not provided, and then further media resources may have to be designed to amplify the message. For example, education materials to explain the working of a car engine may be simple diagrams and text. To really convey how an engine works requires an interactive animation so the student can see the sequence of movements and relate these to a causal model of the internal combustion engine. Multimedia design, therefore, often involves transforming and augmenting information content for effective communication.

Recommendations for selecting media have to be interpreted according to the users’ task and design goals. If information provision is the main design goal, for example, in a tourist kiosk information system, then persistence of information and drawing attention to specific items is not necessarily as critical as in tutorial applications. Task and user characteristics influence media choice; for instance, verbal media are more appropriate to language-based and logical reasoning tasks; visual media are suitable for spatial tasks involving moving, positioning, and orienting objects. Some users may prefer visual media, while image is of little use for blind users.

### 34.3.1 Information Architecture

This activity consists of several activities, which will differ according to the type of application. Some applications might have a strong task model; for instance, a multimedia process control application where the tasks are monitoring a chemical plant, diagnosing problems, and supporting the operator in controlling plant operation. In goal-driven applications, information requirements are derived from the task model. In information provision applications such as websites with an informative role, information analysis involves categorization, and the architecture generally follows a hierarchical model. In the third class of explanatory or thematic applications, analysis is concerned with the story or argument, that is, how the information should be explained or delivered. Educational multimedia and websites with persuasive missions fall into the last category.

In information-provision applications, classification of the content according to one or more user views defines the information architecture; for example, most university departments have an information structure with upper-level categories for research, undergraduate courses, postgraduate courses, staff interests, departmental organization, mission, and objectives. For explanatory applications,
a theme or story line needs to be developed. This will depend on the application’s objectives and the message the owner wishes to deliver. A thematic map from a health awareness application is illustrated in Figure 34.2.

In the case described in Figure 34.2, the requirement is to convince people of the dangers of heart disease. The theme is a persuasive argument that first tries to convince people of the dangers from smoking, poor diet, stressful lifestyles, etc. It then explains how to improve their lifestyle to prevent heart disease, followed by reinforcing the message with the benefits of a healthy lifestyle such as lower health insurance, saving money, longer life, etc. Subthemes are embedded at different points, so users can explore the facts behind heart disease, the statistics and their exposure, how to get help, etc. Information is then gathered for each node in the thematic map. How this architecture will be delivered depends on interaction design decisions: it could, for example, become an interactive story to explore different lifestyle choices, combined with a quiz. The outcome of information architecture analysis will be an information-enhanced task model, a thematic map, or a hierarchy/network to show the structure and relationships of information categories. The next step is to analyze the information content by classifying it by types.

Information types are amodal, conceptual descriptions of information components that elaborate the content definition. Information components are classified into one or more of the following:

- Physical items relating to tangible observable aspects of the world
- Spatial items relating to geography and location in the world
- Conceptual abstract information, facts, and concepts related to language
- Static information that does not change: objects, entities, relationships, states, and attributes
- Dynamic, or time-varying information: events, actions, activities, procedures, and movements
- Descriptive information, attributes of objects, and entities
- Values and numbers
- Causal explanations

Information is often complex or composite, so one component may be classified with more than one type; for instance, instructions on how to get to the railway station may contain procedural information (the instructions “turn left, straight ahead,” etc.), and spatial or descriptive information (the station is in the corner of the square, painted blue). The information types are “tools for thought,” which can be used either to classify specifications of content or to consider what content may be necessary. To illustrate, for the task “navigate to the railway station,” the content may be minimally specified as “instructions how to get there,” in which case the information types prompt questions in the form “what sort of information does the user need to fulfill the task/user goal?” Alternatively, the content may be specified as a scenario narrative of directions, waymarks to recognize, and description of the target. In this case, the types
classify components in the narrative to elucidate the deeper structure of the content. The granularity of components is a matter of the designer’s choice and will depend on the level of detail demanded by the application. To illustrate the analysis:

Communication goal: Explain how to assemble a bookshelf from ready-made parts.

Information component 1:

Parts of the bookshelf, sides, back, shelves, connecting screws

Mapping to information types:
Physical-Static-Descriptive: parts of the bookshelf are tangible, do not change, and need to be described
Physical-Static-Spatial: dimensions of the parts, the way they are organized
Physical-Static-Relationship: this type could also be added to describe which parts fit together

Information component 2:

How to assemble parts instructions?

Mapping to information types:
Physical-Dynamic-Discrete action (each step in the assembly)
Physical-Dynamic-Procedure (all the steps so the overall sequence is clear)
Physical-Static-State (to show final assembled bookshelf)

The mapping physical information at the action and then the procedure level is to improve integration of information and hence understanding, following the “reinforce messages” principle. First, each step is shown, then the steps in a complete sequence (the procedure), and finally the end state of the completed assembly.

34.3.2 Matching Information to Media

The following heuristics are supplemented by more detailed examples in Sutcliffe (2012) and ISO (1998).

- To convey detail, use static media, for example, text for language-based content, diagrams for models, or still image for physical detail of objects (Booher, 1975; Faraday and Sutcliffe, 1998).
- To engage the user and draw attention, use dynamic media, for example, video for physical information, animation, or speech.
- For spatial information, use diagrams, maps, with photographic images to illustrate detail and animations to indicate pathways (Bieger and Glock, 1984; May and Barnard, 1995).
- For values and quantitative information, use charts and graphs for overviews and trends, supplemented by tables for detail (Bertin, 1983; Tufte, 1997).
- Abstract concepts, relationships, and models should be illustrated with diagrams explained by text captions and speech to give supplementary information.
- Complex actions and procedures should be illustrated as a slideshow of images for each step, followed by a video of the whole sequence to integrate the steps. Text captions on the still images and speech commentary provide supplementary information (Hegarty and Just, 1993). Text and bullet points summarize the steps at the end, so choice trade-offs may be constrained by cost and quality considerations.
- To explain causality, still and moving image media need to be combined with text (Narayanan and Hegarty, 1998), for example, the cause of a flood is explained by text describing excessive rainfall with an animation of the river level rising and overflowing its banks. Causal explanations of physical phenomena may be given by introducing the topic using linguistic media, showing cause and effect by a combination of still image and text with speech captions for commentary; integrate the message by moving image with voice commentary and provide a bullet point text summary.
34.4 UX and Multimedia

The term User EXperience (UX) grew out of concerns that traditional concepts of usability (ISO, 1997; Shneiderman and Plaisant, 2004) did not cover the more aesthetic aspects of design. Traditionally, usability has emphasized ease of use, ease of learning, and effective operation, in other words, the “drivability of an interface,” and how well it fits the user’s tasks and goals. Norman (2004) questioned the traditional view of usability in his book on emotion in design and pointed to the importance of aesthetic aspects in UIs and users’ emotional responses to well-designed products. Researchers in HCI began to question how aesthetic design might be related to usability, led by the pioneering studies of Tractinsky (1997) and Lavie and Tractinsky (2004) who experimentally manipulated usability and aesthetic qualities of a design to coin the now well-known aphorism “what is beautiful is usable.”

UX also evolved from marketing concepts such as “consumer experience” (Thomas and Macredie, 2002), which refers to the totality of product experience including sales, setup, use, support during use, and maintenance. Other influences came from research into enjoyment and fun (Blythe et al., 2004) and application areas such as games and entertainment where traditional usability appeared to be less appropriate. When excitement and amusement are the major design goals, interaction, metaphor, and aesthetics become important concerns. Hence, UX generally refers to a wider concept of design beyond functional products, which encompasses interaction, flow (Csikszentmihalyi, 2002), and aesthetic design. The content of media as well as the design of interaction in multimedia systems contribute to UX.

The criteria which may influence UX, and hence user satisfaction and use of IT products, are summarized in Figure 34.3.

Content and functionality are related to users’ goals or requirements, so this aspect is dealt with in processes for requirement analysis (Sutcliffe et al., 2006) and user-centered design. Multimedia represents content to improve UX compared to monomedia (e.g., text or image alone) and is used to implement interaction using graphical metaphors, characters, speech, and audio. Content and services are closely related to task goals, so these criteria will be more important in work-related applications. In contrast, engagement and aesthetics may gain importance in entertainment and leisure applications.

Customization interacts with content and services since users are rarely a homogenous group, and different subgroups often require subsets of the overall content and functions. Customization may be at the subgroup or individual level (personalization), to increase the fit between the application and the user’s needs, which, in turn, increases the user’s commitment to a design. UX is enhanced not only through more stimulating interaction and better fit to the user’s task and abilities (Sutcliffe et al., 2005), but also through the sense of ownership. Personalizing the choice of media and content within images, videos, etc. are examples of customization.

FIGURE 34.3 Components of UX.
Brand and identity interact with content and the users’ background. If users are familiar with product brands, they are naturally going to be well disposed toward an application. Similarly, familiar content can be motivating as it is easier to assimilate. However, the users’ reaction to content is complex; in some domains, for example, education and entertainment, challenging and unfamiliar content may be at a premium for stimulating learning and excitement, while in traditional information processing, familiar patterns of content make task performance more effective. At a high level, content and the tone of its expression in language may be familiar or not in a cultural sense, for example, the expression of material in English, American English, or the user’s native language in international UIs.

Usability is always important, but the degree of ease of use may depend on the domain, varying between high-quality usability when content and functionality are priorities, to sufficient usability when user engagement and fun are the main criteria. Aesthetics or the “look and feel” of a multimedia application may be important in domains where design quality and brand image are at a premium (especially on websites). The debate about the relative influence of aesthetics is complex, where aesthetics is often restricted to the perception or “look” of a UI, separate from interaction or the “feel” of users’ interfaces, which focuses on making interaction exciting and motivating for the user while also delivering effective content, and functionality that meets the users’ goals.

### 34.5 Components of User Engagement

User engagement synthesizes several influences to promote a sense of flow and fluid interaction leading to satisfying arousal and pleasurable emotions of curiosity, surprise, and joy. The contributing influences are summarized in Figure 34.4.

The three main components of user engagement are interaction, multimedia, and presence. Interaction describes how the user controls the computer. The default is standard WIMP (windows, icons, mouse, pointer) interaction on graphical UIs. Multimedia not only deals with the representation of content, but also the interactive environment and how the user and the means of interaction are represented, ranging from simple cursors to icons and interactive avatars with fly-through interactive paradigms. Presence is determined by the representation of the user, which ranges from a cursor in a 2D interactive surface to an avatar in a more elaborate 3D interactive world. Presence is also related to, and augmented by, interaction. Acting in a 3D world and interaction with objects all increase the sense of “being there” and arousal. Flow is the key concept for understanding interaction in terms of the pace of action, complexity of actions, and the rate of change.

Flow, presence, and immersion are concepts which influence the choice of media for interaction; so in this case it is UI rather than content which is the subject of multimedia design. Text dominates
traditional form-filling and menu interfaces, although graphics in the form of icons and structures (windows, sliders, menu bars) are now standard UI components. In the emerging and future UIs, new media will appear, as the user interacts not just via a cursor but also through an avatar, situated in a graphical virtual world, occupied by animated objects, with speech and audio. Multimedia interaction improves UX through flow, presence, and immersion.

34.5.1 Interaction and Flow

Flow is a finely tuned balancing act between the user’s abilities and skills, and the challenges provided by learning new interactions, and then responding to events. It is the sense of engagement and being absorbed in an interactive experience. The concept involves optimal arousal produced by a “sweet spot” trade-off between challenge and difficulty on one hand and ease of operation and achievement on the other. If operating a UI is too difficult, users will get frustrated and discouraged and may give up, leaving them with negative emotions and adverse memory of the experience. In contrast, if operating the UI is too easy, they become bored, excitement (or arousal) decreases, and they turn their attention to more interesting things. The trick is to keep interaction in the flow zone (see Figure 34.5), an intuition appreciated by game designers. Games need to maintain the pace of change with unpredictable events, while not overwhelming the user with too much change that exceeds their capabilities. They rapidly become used to patterns of events leading to decreased arousal as the unfamiliar becomes familiar.

34.5.2 Presence

The origins of presence are in virtual reality (VR), in which the user is represented by an avatar or virtual character. The avatar places the user inside a 3D graphic world, so interaction becomes very different from operating an interface through a cursor. Presence is the sense of “being there” inside a virtual world as a representation of oneself (embodied or immersed interaction). Embodied interaction is more complex than standard 2D interfaces, since the user can control movement directly, manipulate objects in an almost natural manner, and even feel objects if haptic feedback is provided. The sense of immersion and control is enhanced via natural movement. Virtual worlds become engaging because they invoke curiosity and arousal. Interaction becomes transparent, i.e., the user is not aware of the computer, instead feeling immersed and becoming absorbed in the virtual graphical world (see Figure 34.6).

FIGURE 34.5  Change in arousal as events become familiar. The concept of flow.
34.5.3 Engagement and Social Presence

Social presence theory (Short et al., 1976) argues that different communication channels and representations promote more or less sense of social presence, i.e., awareness of the identity, location, and personalities of other people. The theory does not give a formal classification or model of social presence; although it can be reinterpreted in terms of a model of communication which describes degradations from the ideal of face-to-face, colocated communication. E-mail is the least engaging since the time gap between turns destroys the sense of a dialog, whereas instant messaging (IM) approaches synchronous exchange and becomes more engaging. In IM, the pace of exchange approaches a conversation, and the sense of presence increases. Adding video or even still images improves presence by providing more information about the other person (see Figure 34.7). Contextual information gives us the background to an ongoing conversation. The context may be provided by a user profile, a room or whatever has been constructed in the virtual world for interpreting the conversation.

FIGURE 34.6 Virtual world illustrating the user’s presence represented as two virtual hands.

FIGURE 34.7 SecondLife, illustrating avatars that represent the characters people adopt and the location with others, both of which promote social presence.
Strange as it may seem, it takes very little to create an illusion of presence. Our powerful imaginations just need a few hints (priming or framing effects in psychology) to conjure up a perceived reality, as Pinter* and other proponents of minimalist theater have demonstrated. In a series of experiments, Reeves and Nass (1996) showed how we treat computer-based media as if they corresponded to real people. Their computer as social actor (CSA) paradigm explains how we treat computers as virtual people even when we are presented with limited cues, such as a photograph of a person or human voice. Indeed, the image can be artificial, cartoon-like, with little correspondence to reality; the same applies to the voice. Chatterbots, avatars on the web equipped with simple semi-intelligent scripts for responding to human conversations, are treated like real characters, and some people actually form relationships with these virtual characters (De Angeli and Brahnam, 2008).

### 34.6 Media Design for User Engagement

Multimedia design is frequently motivated by the need to attract users’ attention and to make the UX interesting and engaging. These considerations may contradict some of the earlier guidelines, because the design objective is to please users and capture their attention rather than deliver information effectively. First, a health warning should be noted: the old saying “beauty is in the eye of the beholder” has a good foundation. Judgments of aesthetic quality suffer from considerable individual differences. A person’s reaction to a design is a function of their motivation (Brave and Nass, 2008), individual preferences, knowledge of the domain, and exposure to similar examples, to say nothing of peer opinion and fashion.

In human–human conversation, we modify our reactions according to our knowledge (or assumptions) about the other person’s role, group identification, culture, and intention (Clark, 1996). For example, reactions to a military mannequin will be very different from those to the representation of a parson. Male voices tend to be treated as more authoritative than female voices. Simple photographs or more complex interactive animations (talking heads or full body mannequins) have an attractive effect; however, the effectiveness of media representing people depends on the characters’ appearance and voice (see Figure 34.8).

![Sapient AD](https://example.com/sapient.png)

**FIGURE 34.8** Effective use of human image for attraction. The picture attracts by the direction of gaze to the user as well as by the appearance of the individual.

* http://en.wikipedia.org/wiki/Harold_Pinter
Use of human-like forms is feasible with prerecorded videos and photographs; however, the need depends on the application. Video representation of the lecturer can augment presentations, and video communication helps interactive dialog. A good speaker holds our attention by a variety of tricks, such as maintaining eye contact, varying the voice tone, using simple and concise language, as well as delivering an interesting message. These general effects can be reinforced by projected personality. Friendly people are preferred over colder, more hostile individuals. TV announcers who tend to be middle-aged, confident, but avuncular characters have the attention-drawing power of a dominant yet friendly personality. Both sexes pay attention to extrovert, young personalities, while the male preference for beautiful young women is a particularly strong effect. These traits have long been exploited by advertisers. There are lessons here for multimedia designers as the web and interactive TV converge, and when we want media to convey a persuasive message (Reeves and Nass, 1996; Fogg et al., 2008). Media selection guidelines for motivation and persuasion, adapted from Reeves and Nass (1996), can be summarized as follows:

- Human image and speech invokes the CSA effect to facilitate motivation and persuasion.
- Photographs of people attract attention especially when the person is looking directly at the user.
- Faces that represent the norm in a population (Mr./Ms. Average) and young children are more attractive. We are very susceptible to the large-eyes effect in young animals, as exploited by Disney cartoons.
- Polite praise: Use of please, thank you, and simple compliments such as “that was an excellent choice” increase people’s tendency to judge the computer as pleasant and enjoyable.
- Short compelling argument such as the well-known British World War I recruiting poster featuring General Kitchener gazing directly at the viewer with the caption “your country needs you.”

34.6.1 Media for Emotional Effects

Media design for affect (emotional response and arousal) involves both choice of content and interaction. Arousal is increased by interactive applications, surprising events during interaction, use of dynamic media, and challenging images. In contrast, if the objective is to calm the users, arousal can be decreased by choice of natural images and sounds, and soothing music. The most common emotional responses that designers may want to invoke are pleasure, anxiety and fear, and surprise. Pleasure, anxiety, and fear usually depend on our memory of agents, objects, and events (Ortony et al., 1988), so content selection is the important determinant. Anxiety can be evoked by uncertainty in interaction and cues to hidden effects, while emotional response of fear or pleasure will depend on matching content to the user’s previous experience. Some guidelines to consider are as follows:

- **Dynamic media**, especially video, have an arousing effect and attract attention; hence, video and animation are useful in improving the attractiveness of presentations. However, animation must be used with care as gratuitous video that cannot be turned off quickly offends (Spool et al., 1999).
- **Speech** engages attention because we naturally listen to conversation. Choice of voice depends on the application: female voices for more restful and information effects, male voices to suggest authority and respect (Reeves and Nass, 1996).
- **Images** may be selected for mood setting, for example, to provide a restful setting for more important foreground information (Mullet and Sano, 1995). Backgrounds in half shades and low saturation color provide more depth and interest in an image.
- **Music** has an important emotive appeal, but it needs to be used with care. Classical music may be counterproductive for a younger audience, while older listeners will not find heavy metal pop attractive. Music can set the appropriate mood, for example, loud strident pieces arouse and excite, whereas romantic music calms and invokes pleasure, etc.
- **Natural sounds** such as running water, wind in trees, bird song, and waves on a seashore have restful properties and hence decrease arousal.
• **Dangerous and threatening episodes**, for example, being chased by a tiger, gory images (mutilated body), and erotic content all increase arousal and invoke emotions ranging from fear to anger, whereas pleasant images (e.g., flowers, sunset) tend to decrease it, having calming effects and producing pleasurable emotional responses.

• **Characters** can appear threatening or benevolent depending on their appearance or dress. For example, disfigured people appear threatening and evoke emotions ranging from fear to disgust. Characters familiar from popular culture can be used for emotional effect.

• **Spoken dialog** is probably the most powerful tool for creating emotional responses, from threats to empathy. Emotional effects are additive; so choice of character with a threatening appearance, complemented by a menacing voice tone and an aggressive dialog, all reinforce the emotions of anxiety and fear.

Media integration rules may be broken for emotive effects. For example, use of two concurrent video streams might be arousing for a younger audience, as Music TV (MTV) and pop videos indicate. Multiple audio and speech tracks can give the impression of complex, busy, and interesting environments.

### 34.6.2 Multimedia and Aesthetic Design

Judging when aesthetics may be important will be set by the design goals of the owner of the application. For example, in e-commerce applications with high-value, designer-label products, aesthetic presentation is advisable; similarly, when selling to a design-oriented audience. Aesthetic design primarily concerns graphics and visual media. Evaluation questionnaires can assess aesthetics and more creative aspects of visual design (Lavie and Tractinsky, 2004); however, these measure user reaction to general qualities, such as “original,” “fascinating,” “clear,” and “pleasant.” The following heuristics provide more design-directed guidance, but they may also be employed for evaluation (Sutcliffe, 2002; Sutcliffe and De Angeli, 2005).

- **Judicious use of color**: Color use should be balanced, and low saturation pastel colors should be used for backgrounds. Designs should not use more than two to three fully saturated intense colors. Yellow is salient for alerting, red/green have danger/safety positive/negative associations, and blue is more effective for background. Low-saturated colors (pale shades with white) have a calming effect and are also useful for backgrounds. Color is a complex subject in its own right; for more guidance, refer to Travis (1991).

- **Depth of field**: Use of layers in an image stimulates interest and can attract by promoting curiosity. Use of background image with low-saturated colors provides depth for foreground components. Use of layers in an image and washed-out background images stimulate curiosity and can be attractive by promoting a peaceful effect.

- **Use of shape**: Curved shapes convey an attractive visual style, in contrast to blocks and rectangles that portray structure, categories, and order in a layout.

- **Symmetry**: Symmetrical layouts, for example, bilateral, radial organization, can be folded over to show the symmetrical match.

- **Simplicity and space**: Uncluttered, simple layouts that use space to separate and emphasize key components.

- **Design of unusual or challenging images** that stimulate the users’ imagination and increase attraction; unusual images often disobey normal laws of form and perspective.

- **Visual structure and organization**: Dividing an image into thirds (Right, Center, Left; or Top, Middle, Bottom) provides an attractive visual organization, while rectangular shapes following the golden ratio (height/width = 1.618) are aesthetically pleasing. Use of grids to structure image components promotes consistency between pages.

Although guidelines provide ideas that can improve aesthetic design and attractiveness of interfaces, there is no guarantee that these effects will be achieved. Design is often a trade-off between ease of use...
and aesthetic design; for instance, use of progressive disclosure to promote flow may well be perceived by others as being difficult to learn. Visual effects often show considerable individual differences and learning effects; so, a well-intentioned design might not be successful. The advice, as with most design, is to test ideas and preliminary designs with users to check interpretations, critique ideas, and evaluate their acceptability. There are several sources of more detailed advice on aesthetics and visual design (Mullet and Sano, 1993; Kristof and Satran, 1995; Lidwell et al., 2003).

### 34.6.3 Metaphors and Interaction Design

While task and domain analysis can provide ideas for interaction design, this is also a creative process. Interaction design is essentially a set of choices along a dimension from simple controls, such as menus and buttons where the user is aware of the interface, to embodiment in which the user becomes involved as part of the action by controlling an avatar or other representation of their presence. At this end of the dimension, multimedia interaction converges with virtual reality. Interactive metaphors occupy the middle ground.

Some interactive metaphors are generally applicable, such as timelines to move through historical information, the use of a compass to control the direction of movement in an interactive space, controls based on automobiles (steering wheels) or ships (rudders). Others will be more specific, for example, selecting and interacting with different characters (young, old, male, female, overweight, fit, etc.) in a health-promotion application. Design of interaction also involves creating the microworld within which the user moves and interactive objects that can be selected and manipulated.

Interaction via characters and avatars can increase the user’s sense of engagement first by selecting or even constructing the character. In character-based interaction, the user can either see the world from an egocentric viewpoint, i.e., from their character’s position, or exocentric when they see their character in the graphical world. Engagement is also promoted by surprise and unexpected effects; so, as the user moves into a particular area, a new subworld opens up, or system-controlled avatars appear. These techniques are well known to games programmers; however, they are also applicable to other genres of multimedia applications. The design concepts for engagement can be summarized as follows (see Sutcliffe, 2009, for more detail):

- **Character-driven interaction:** It provides the user with a choice of avatars or personae they can adopt as representations of themselves within the interactive virtual world (see Figure 34.6). Avatar development tools enable virtual characters to be designed and scripted with actions and simple speech dialogs. Most sophisticated, semi-intelligent “chatterbots” (e.g., Alice [http://alice.pandarobots.com] & Jabberwacky [http://www.jabberwacky.com]) use response-planning rules to analyze user input and generate naturally sounding output; however, it is easy to fool these systems with complex natural language input.

- **Tool-based interaction:** This places the tools in the world which users can pick up; the tool becomes the interface, for example, a virtual mirror magnifies, a virtual helicopter flies (Tan et al., 2001).

- **Collaborative characters:** In computer-mediated communication, these characters may represent other users; in other applications, system-controlled avatars appear to explain, guide, or warn the user.

- **Surprise effects:** Although conventional HCI guidelines should encourage making the affordances and presence of interactive objects explicit, when designing for engagement, hiding and surprise are important.

Interaction design for an explanatory/tutorial application is illustrated in Figure 34.9. This is an interactive microworld in which the user plays the role of a dinosaur character, illustrating the use of the engagement concepts. A compass navigation metaphor allows the user to act as a dinosaur moving around the landscape, as illustrated in the figure. The user is given feedback on the characteristics of other predators and prey in the vicinity, and has to decide whether to attack or avoid them.
Other controls that might be added to such interactive microworlds could be the settings to change the environment, for example, add more predators, change the weather, etc.

### 34.7 A Brief Example

A healthcare example illustrates the use of some of the UX-multimedia guidelines. The objective is to persuade people through a web application to adopt healthy diets to tackle the obesity problems. Current designs (e.g., [http://www.nhs.uk/livewell/](http://www.nhs.uk/livewell/); [http://www.nhs.uk/Change4Life](http://www.nhs.uk/Change4Life)) use image and text with limited video to provide advice on the types of food to eat, what to avoid, portion size, etc. However, interactive functions are limited to simple quizzes on diet choices, which are analyzed to show the calories consumed. More adventurous designs might employ an interactive “day in your life” metaphor, where the user controls an avatar while progressing through a typical day making lifestyle choices about exercise (e.g., walk to work or drive), what to eat for each meal, choice of portion sizes, etc. The interactive game-style application provides feedback on choices, approving sensible ones or disapproving poor ones. The CSA guidelines would suggest feedback with praise for good choices given by a sympathetic character. Long-term tracking of diet and weight loss or gain could employ emotional effects to encourage (empathy, praise, empowerment), and to admonish poor progress with warnings about the dangers of obesity. The potential for multimedia design can be seen in the increasing number of mobile applications which use interactive games–style interaction with monitoring to encourage fitness.

### 34.8 Conclusions

Multimedia still poses many issues for further research. The design method described in this chapter coupled with user-centered design can improve quality; however, there is still a need for experts to create specific media resources, for example, film/video, audio. Furthermore, considerable research is
still necessary before we fully understand the psychology of multimedia interaction. Design for UX in multimedia is still poorly understood, for example, some people prefer avatars and rich graphical metaphors, while others find graphical interfaces annoying and prefer simpler designs. Although guidelines can improve design with advice that is based on psychological knowledge, the reaction to designs often depends on the individual; so, one-size-fits-all solutions will not work for everyone. As with most design, testing with users is always advisable, and where possible personalizing and customizing designs for the user or giving them the choice can improve the UX.

Current research frontiers are unpacking the role of interactivity in UX, as multimedia converges with VR and 3D interactive worlds, such as Google street view technology, become commonplace. Immersion and presence increase user engagement with games (Jennet et al., 2009); however, application of virtual worlds to other domains, such as education, does not always improve UX (Sutcliffe and Alrayes, 2012). Although interactive features do appear to have a positive appeal in many websites (Cry et al., 2009), the debate continues about the relative importance of interactive and aesthetic design on overall UX (Bargas-Avila and Hornbæk, 2011; Lindgaard et al., 2011). On balance, it appears that “pragmatic” design considerations such as functionality and usability outweigh “hedonic” factors such as aesthetics (Diefenbach and Hassenhalz, 2009), but the contribution of interaction design remains unclear.

In the future, language and multimodal communication will change our conception of multimedia from its current CD-ROM or web-based form into interfaces that are conversational and multisensory. Multimedia will become part of wearable and ubiquitous UIs where the media is part of our everyday environment. Design for multisensory communication will treat media and artifacts (e.g., our desks, clothes, walls in our homes) as a continuum, while managing the diverse inputs to multimedia from creative design, technology, and usability engineering will be one of the many interesting future challenges.

References


