Aesthetic Interaction: A Framework
Paul Locher,¹ Kees Overbeeke,² and Stephan Wensveen³

The rise in the development of interactive electronic products has been accompanied by growing interest in the aesthetic aspects not only of the artifacts themselves but in the aesthetics of interactive systems. Petersen, Iversen, Krogh, and Ludvigsen⁵ point out that these two approaches to the aesthetics of design reflect Shusterman’s⁶ distinction between analytical aesthetics and pragmatic aesthetics, respectively. From an analytic perspective, aesthetics arise as a product property, as “added value” to an artifact. The focus of the design process here is on the aesthetics of appearance, on the creation of artifacts that are attractive and pleasurable. The pragmatic approach, on the other hand, is concerned with the aesthetics of use. According to this view, the aesthetics of an artifact emerge out of a dynamic interaction between a user and this artifact and is an integral part of what has been labeled an aesthetic interaction by some researchers⁷ in design and as a resonant interaction by others.⁸

At the same time the scope of design is changing from human/artifact interaction, mainly focused on opening up the functionality of a product, toward a broader approach that seeks to enhance interpersonal and societal values, including personal, aesthetic, and socio-cultural ones, through the application of intelligence (i.e., smart electronics) in artifacts.

Much has been written concerning the factors that contribute to the aesthetics of human-artifact interaction. However, to our knowledge, no framework or conceptual model of the structure of the interactive aesthetic experience that incorporates these factors has appeared in the literature. In this paper we integrate an information-processing model of the nature of an aesthetic experience with visual art proposed by Locher and his colleagues⁹, ¹⁰ with a framework proposed by Wensveen¹¹ that describes the coupling of a user’s actions (i.e., handling an artifact) and a product’s function; the result is the formation of a general theoretical framework for understanding the nature of a user’s aesthetic interaction with design products. Our hope is that the proposed conceptual framework will serve as a valuable basis for the development of experimental studies into the nature of aesthetic interaction to complement the experimental tradition of usability studies among designers.

Before presenting the framework, it is important to note, as have Petersen and her colleagues,¹² that the notion of aesthetic is used in ambiguous ways by theoreticians when it comes to answering the key question: What is the nature of the resulting
emotion arising out of an aesthetic interactive experience? In other words, what is the aesthetic of interactive systems? This ambiguity is evidenced by the many terms found in the literature used to describe the affect generated—terms such as fun, surprise, delight, engagement, and rewarding. Furthermore, the failure to provide technical distinctions among the concepts used to describe the aesthetic outcome of an aesthetic interaction remains a central problem in this field, as well as in the arts. The purpose of this paper is to outline the structure of the aesthetic experience; it is beyond the paper’s scope to provide an empirically based explanation of the nature of the affect (either positive or negative) that results from this experience. We note, however, that the pragmatic view of the nature of an aesthetic interaction with artifacts presented herein closely mirrors Csikszentmihalyi and Robinson’s” empirically based interpretation of a viewer’s aesthetic experience with art, which is indistinguishable from what they call a flow experience. Briefly stated, their explanation of a flow experience asserts that individuals engage art objects “not because they expect a result or reward after the activity is concluded, but because they enjoy what they are doing to the extent that experiencing the activity becomes its own reward.”

A flow experience is one that contains its goal in itself; it involves deep involvement in and effortless progression of the activity with an artwork. In our view, this heightened state of awareness when one experiences great art, and during other types of activity, such as sports, hobbies, and challenging work, is the same type of involvement that occurs between a user and an artifact during an aesthetic experience with interactive systems (qualitatively so, if not quantitatively).

We turn now to the framework of the interaction of artifact-driven and cognitively driven processes (referred to as bottom-up and top-down processes, respectively, in Information Processing Theory) underlying user-product interaction and the resulting aesthetic experience described in this paper (see Figure 1). The directions of the arrows in the figure indicate that in the experience of a product there is a continuous, dynamic bottom-up/top-down interaction between the properties (form) and functionality of the artifact, the user’s sensory-motor-perceptual (i.e., visual, handling or active touch, auditory) processes involved, and the user’s cognitive structure. Thus, as an aesthetic experience progresses, the artifact presents continually changing, “action driven” affordances. These affordances in turn influence the timing, rhythm, flow, and feel of the interaction—factors seen as playing important roles in aesthetics of interaction by Djadadiningrat, Wensveen, Frens, and Overbeeke.
between attention and action. According to Baddeley,\textsuperscript{16} the central executive, which is one of the three components of working memory (the other two being subsidiary storage systems: the phonological loop and the visuospatial sketchpad), performs four important executive processes: “the capacity to focus attention, to divide attention, to switch attention, and to provide a link between working memory and long-term memory.”

Together the top-down and bottom-up component processes underlying thought and action create both meaning and aesthetic quality of the artifact from which the aesthetic experience with the artifact and the resulting affect emerge. For example, Wensveen\textsuperscript{17} designed an alarm clock in such a way that it offers freedom of interaction: The user can set the alarm time in a myriad of ways by moving one of twelve sliders. This allows the alarm clock to measure the user’s mood when the user sets the alarm time in an expressive way. The intertwining of perceptual-motor, cognitive, and emotional elements thus leads to an aesthetic experience. In experiments it was striking to see that when the users were in a good mood, they made symmetric and balanced patterns, and when they were in a bad mood, they made more irregular ones.

As shown in Figure 1, the two driving forces of the system are the artifact itself and a person context that reflects the user’s cognitive structures. The aesthetic experience is a product of the dynamic, ongoing interaction between these two components of the system. With respect to the artifact context, it has been shown

\textsuperscript{16} Alan Baddeley, Working Memory, Thought, and Action (Oxford: Oxford University Press, 2007), 119–120.

\textsuperscript{17} Wensveen, A Tangibility Approach to Affective Interaction, 117.
that features of an artifact provide a user with different types of information. Specifically, research has identified at least six ways in which the appearance of a product influences consumer product evaluation and choice, typically in an artifact-driven or a bottom-up fashion. An artifact’s appearance can convey its aesthetic and symbolic value and provide a quality impression; it can communicate functional characteristics and ease of use; it can draw attention by visual novelty and communicate ease of product categorization. In addition to presenting product properties, interactive artifacts can be designed so that their use contributes to a dynamic aesthetic interaction between their form and functionality and the user. Although the primary focus of this paper is the aesthetics of interaction, the aesthetics of appearance of an artifact must always be taken into consideration as contributing factors to a user’s interaction with it.

The second major contributing component to an aesthetic interaction is the user’s cognitive structure, which contains several types of information (semantic, episodic, and strategic) acquired throughout his or her life. It is also the repository of one’s personality, motivations, and emotional state. All these components are brought to bear in a top-down fashion on a user’s interaction with a product, and they determine how he or she invites, perceives, and evaluates it. These components simultaneously contribute to and create what we call the “person context” in which the aesthetic experience takes place.

As mentioned, the aesthetics of use emerges out of the dynamic interaction between a user and the product’s form and functionality. When using mechanical products, such as a pair of scissors, there is a natural or unmediated coupling between a product’s appearance, the action possibilities for its use, the action, and the function, which supports intuitive interaction with the product. Interactive electronic products, on the other hand, require an interface for individuals to interact with them. Users need information from the product, both in the form of feedforward and feedback, to guide their actions toward the couplings between actions and functions. Wensveen has presented a framework to conceptualize the person-product interaction that focuses on three types of information the user can receive from an interactive system: inherent, augmented, and functional.

Inherent information is the information provided by the natural consequences of taking an action—that is, by touching an object while simultaneously observing it visually. This type of information ties together the action possibilities of the product and the perceptual/motor abilities of a user. Inherent feedforward information from the product communicates the kind of actions possible when using it, such as pushing, sliding, or rolling its components, and how the action can be carried out (e.g., the amount of force required).
of force required to bring about an action). Inherent feedback is the information returned from acting on a product’s action possibilities (e.g., the feel and sound of a button on a product when it is pushed). Both feedforward information and feedback information are acquired in a bottom-up fashion by the user as indicated by the arrow drawn from the artifact to the sensory-motor system in Figure 1.

Augmented information comes not from an action on the product itself but from an additional source about either the action possibilities of the artifact or the purpose of the action possibilities. This source informs a user about an internal state of the system through the use of such artifact features as LCDs, light-emitting diodes, and sounds. This type of information draws on the user’s knowledge about such artifacts and is added to the product by the central executive (as indicated by the arrow from the artifact to the central executive in Figure 1). Feedforward augmented information provides information about the action possibilities of the product in the form of, for example, on-screen messages (e.g., words, pictograms, or graphical labels) indicating what to do. The information that a user receives when these sources are activated and inform the user of the internal state of the system (indicating, e.g., “processing,” “stand by,” “log off”) is called augmented feedback.

The third type of information, functional information, relates directly to the function of the product; it is the goal of the interaction, the actual purpose of the product. Functional feedforward information is provided by the visible functional parts or components of a product, which inform the user about the functionalities of the product (e.g., the speakers and the screen of a television). When users receive information (feedback) from the functional parts of a product, it is clear to them that their actions were successful. Thus, functional information is generated by the combined output of both bottom-up and top-down processes (i.e., by artifact and central executive processes), as indicated by the arrows and their directions in Figure 1.

Because interaction with a product involves the simultaneous use of visual and haptic (exploratory touch) perception following an initial glance at it, a brief description of the physical relationship between the two sense modalities is in order to understand how handling an object can add to the perception and aesthetic evaluation of an artifact beyond vision’s contribution, the basic premise of this framework. Research suggests that vision and haptics are differentially suited to extract and encode information about objects (e.g., haptics for texture and vision for spatial location) and that the two modalities interact in various ways at the encoding stage of processing. The nature of the interaction is mediated by differential attention to an object’s features, based on the demand characteristics of the perceptual performance required by the task being performed. However, there is reasonably good behavioral and neuroscience evidence that visual and haptic object representation

---

is shared between these two perceptual modalities; thus, stimulus information from an artifact obtained separately by the two senses becomes combined at the cortical level into a common bimodal, cognitive representation of the object. Supporting this assertion are findings by a number of neuroimaging studies, which have demonstrated, for example, that haptic exploration of novel objects produces activation not only in the somatosensory cortex when the same objects were later viewed, but also in areas of the occipital cortex associated with visual perception. Findings such as these provide support for the view that the haptic component of a dynamic interaction with an artifact not only makes its own modality-specific contribution to the aesthetic experience with it but, more importantly for the present discussion, also combines with vision’s contribution to the aesthetics generated by the user-product interaction at “higher levels” of processing, as described. (It should be noted that the auditory qualities of the artifact likely contribute to the aesthetics of a multi-sensory interaction; however, this modality is not part of the focus of the present discussion.)

Behavioral evidence that tactile information can affect the aesthetic evaluation of artifacts is provided by the findings of a recent study by Jansson-Boyd and Marlow. They asked undergraduate students to rate the aesthetic quality of DVD containers that varied in three types of plastic textures (a smooth surface, a ribbed plastic surface, or a thick matte plastic material) under three viewing conditions (visual only, blind haptic evaluation, or simultaneous visual and tactile sensing of the DVD). The front cover of the DVD (E.T.: The Extra-Terrestrial special edition) was the same across conditions. It was found that the aesthetic quality of the DVD containers was influenced by both visual and tactile evaluation and that the extent to which the modalities influenced evaluations, individually and collectively, was a function of which of the three surface textures was being evaluated.

We return now to a description of an aesthetic interaction with an artifact as depicted in Figure 1. As mentioned, the components of the framework and the processes indicated by the arrows are adapted from the model describing the nature of an aesthetic experience with visual art, proposed by Locher and his colleagues. According to this two-stage model, exploration of a painting by a viewer starts with rapid encoding of the content of its pictorial field to acquire an initial impression (or gist) of the structural arrangement and semantic meaning of the composition. The gist information detected with the initial glance at a composition drives the second stage of an aesthetic experience, which consists of visual scrutiny or focal analysis of presumably interesting pictorial features detected initially to satisfy cognitive curiosity and to develop aesthetic appreciation of the painting. We propose that a user’s experience with a product follows these same two stages.

---


There is ample evidence that many physical properties of an art work (e.g., its structural complexity, symmetry, and organizational balance) are detected by the visual system automatically or pre-attentively by genetically determined, hard-wired perceptual mechanisms. Research also shows that the sense of touch is capable of rapidly recognizing stimulus properties of objects simply from sensory information (e.g., shape, texture, “sensuous feelings”). Such innate processes are indicated in Figure 1 by the arrow between the sensory-motor system and the aesthetic experience, which reflects a rapid initial impression of the object by haptics as well as by vision. The initial stage of processing by the visual and haptic sensory systems just described is similar to the first of three levels of processing of artifacts proposed by Norman, called the visceral level, which involves the rapid generation of a first impression of the artifact based on hard-wired, automatic processes. Such reactions have been referred to as “natural perceptive responses to products” by Overbeeke and Forlizzi, who, like Norman, assert that they are evoked in the absence of significant interaction with products. The aesthetics of artifacts must, therefore, be concerned with the immediate impressions of products, obtained first by visual perception and then by initial handling of the product. That is, designers must create “effective visceral designs,” in Norman’s words, that are attractive at first glance (both visually and then haptically) and that appear pleasurable to use when they present themselves to us. Thus, an analytical approach to aesthetics is, in a sense, an important “first step” of a pragmatic approach to design. A positive first impression of a product is essential if there is to be any further interaction with it. It is most likely the case that a user’s initial reaction to an artifact also influences how the artifact is “processed” during the aesthetic experience, as is the case for artworks, although to our knowledge this influence has not been demonstrated empirically.

In addition to the automatic detection of physical properties of artworks and artifacts, it has also been shown that individuals are capable of rapidly detecting and categorizing learned properties of a stimulus. For example, Locher and others have demonstrated that characteristics of the artistic style of a painting (e.g., abstract, representational) and a composition’s pleasantness and interestingness can be detected with a single rapid (100 ms) glance at it. In addition, Creusen and Schoormans report that almost all members of a consumer household panel were able to perceive the overall form and appearance of three product alternatives of two artifacts (viz., a clock radio and hairdryer) within 800 ms of presentation onset. These responses occur by a rapid and direct match in activated memory between the structural features of an art object or artifact generated by the sensory-motor system and a viewer’s knowledge about the stimulus stored in his/her cognitive system (person context). The resulting rapid automatic reaction to the stimulus, represented in

26 See Heller and Schiff, The Psychology of Touch.
the framework by the arrow drawn from activated memory directly to the aesthetic experience in Figure 1, also contributes to one’s first impression of it.

Once an initial impression of an artifact is formed based on information obtained from seeing and handling it, the second stage of processing—focused attention to its form and functionality—follows, directed by the central executive. For the visual modality, users gather information about an artifact by moving their eyes over it in a sequence of rapid jumps, or saccades, followed by pauses or fixations. The number, location, and duration of fixations used to visually scrutinize the artifact constitute the spatial-temporal aspects of encoding, in Figure 1. For touch, information about an artifact is similarly obtained by users actively moving one or both hands about the product to select and manipulate its features, usually in concert with vision in sighted individuals. The encoding activity of both modalities is indicated by the action arrow in Figure 1 drawn between the sensory-motor system and the artifact. Once again, it is important to note that the perception and aesthetic evaluation of an artifact emerges out of the dynamic interaction of input obtained by both looking at and handling an artifact. Product information in activated memory, acquired by visual and haptic experience with the artifact during the second phase of processing, spontaneously activates subsets of featural and semantic information in the user’s knowledge base. The information drawn into active memory across the time course of the interaction is determined by effortful processing on the part of the central executive as the user/product interaction unfolds within the ongoing interaction space, as shown in Figure 1. This ongoing process is influenced by the factors of the person context shown in the figure, including the user’s level of aesthetic sophistication (i.e., experience in the arts and design), personal tastes, level of education, cultural background, personality, and his or her emotional and cognitive state during the aesthetic experience, to name but a few of the factors most relevant to an interaction with a product. In this respect, the central executive corresponds in function to Norman’s31 reflective level of processing, which, along with the behavioral level of processing in his model, are very sensitive to experience, training, culture, and education. However, Norman asserts that the reflective level does not have access or control over sensory input or behavior, whereas these functions of the central executive are critical in the present model.

Research from our laboratory has demonstrated how some of the factors shown in the person context of Figure 1 contribute in an interactive way to a user’s aesthetic experience with a product. For example, we32 observed that positive affect, induced by the gift of a small bag of candy, enhanced ratings of the appeal of digital cameras by participants untrained and trained in principles of design, compared to control groups who did not receive candy.

Analysis of cognitive process measures obtained from participants’ verbal protocols collected as they completed the task revealed that individuals in a positive mood state differentially influenced the way the groups of participants thought about the cameras as they made their rating decisions. For those untrained in design, positive affect cued and facilitated access to more positive material in memory, which enhanced their perception of the cameras’ appeal. In contrast, design-trained individuals in whom positive affect had been induced showed greater access and use of design-related information in memory than design-trained students who did not receive candy. This enabled them to identify more aspects of good design in the cameras, and correspondingly, made the cameras more appealing from a design perspective. These findings illustrate how aesthetic expertise (or lack thereof) and motivation (two factors shown in Figure 1) combine in an interactive way through the central executive to influence in a top-down fashion the spatial-temporal aspects of interaction with the cameras (the artifact).

There is yet another set of factors that contributes to a user’s interaction with an artifact, and these factors constitute the artifact context in Figure 1. They include product characteristics and situational characteristics. As mentioned previously, the appearance of an artifact communicates at least six different roles of a product, of which the symbolic role was one of the most frequently mentioned by participants in a study by Creusen and Schoormans. The social-cultural and socio-economic factors related to an artifact, its historical significance, the quality of the materials out of which the artifact is constructed (e.g., wood vs. plastic), and the marketing programs used to sell the artifact (e.g., brand names) all influence the perceived symbolic associations and social value of products. These factors, in turn, contribute to a user’s self-perception of his or her cultural taste (i.e., the values and standards to which he or she aspires, either as an individual or as an expression of group membership). The positive or negative values assigned to products are based on pre-existing knowledge in the user’s knowledge base (i.e., the person context) and, as such, function to influence an interaction in a cognitively driven or a top-down fashion. Situational characteristics, or the environment in which one experiences an artifact, also provide an artifact context that influences the nature and outcome of an interaction with an artifact. Using a product in either a store, at home, or in a product test situation likely influences in a differential fashion the experience one has with it. Observation time available to process product information is another factor that determines how superficially or intensely one can pay attention to the product. In addition, the salience and functional dimensions of an artifact can be “primed” in a user through subtle factors created by previous advertised exposure.

Hummels, Ross, and Overbeeke also see context mode as an important aspect to consider when designing for resonant interaction.
between a user and a product. In addition to the influence of where and when an interaction takes place, already discussed, the context factors include “how often,” “how long,” and “how frequently” a product is used. The “use factor” relates to a key question not yet addressed here: At what point will interactive consumer products designed to provide rich interactions become just “things” in users’ lives? If, as is the case with all stimuli, the brain naturally adapts to repeated exposure to a stimulus, how then can repeated experience with an artifact maintain a user’s excitement, interest, and pleasure through an interaction, even after long acquaintance with it? Norman’s37 answer is that, if an artifact is to give continued pleasure, two components are required: the skill of the designer in providing a powerful, rich, and compelling experience, and the skill of the user to detect this richness. He notes that works in the fields of art, music, and literature that have stood the test of time are rich and deep so that there is something new to be encountered on each experience with such “classics.” These two factors fall within the artifact and person context factors, respectively, of the present framework.

In conclusion, we believe the framework presented in this paper provides a comprehensive foundation upon which the nature of an aesthetic experience, in an interaction with a design product, may be better understood. The important point conveyed by the framework presented herein is that there are many moderating factors contributing in complex, dynamic ways to influence a user’s aesthetic experience with a design product. We hope that it suggests promising future research directions and offers the design community the potential for developing concrete guidelines for designing interactive products.38

37 Norman, Emotional Design, 111.
38 We wish to thank Lilian Admiraal, who designed Figure 1 as part of her M1.2 research project at the TU/e Industrial Design Department.