

A Media Taxonomy

Rachelle S. Heller and C. Dianne Martin
George Washington University

This media taxonomy serves both research and development of multimedia applications. Two graduate courses used it successfully, one on multimedia design and another on the research and evaluation of interactive multimedia. It correlates well with previous categorizations of multimedia and helps researchers better understand the impact and value added by an individual medium in a multimedia presentation.

In just a few years the use of interactive multimedia in various computer applications has moved from an interesting novelty to a standard. Examples abound of manufacturers extending existing productivity tools to handle various media. Spreadsheets such as Lotus Notes and word processors such as Microsoft Word have moved from text only to include images and even sound. Tools for creating presentations, such as Microsoft Power Point, have moved from text and still graphics to include Apple QuickTime movies and animated presentations. Development tools such as Asymmetric Multimedia ToolBook and Macromedia Authorware Professional have been extended for easy inclusion of images, sound, and movies in new applications. New tools for presentation and application development, such as the MediaText project,¹ provide access to sources of media including, but not limited to, CD-ROM (audio and video), videodisc, PhotoCD, stored animations, graphics, and sound. Researchers have written much about the need to create easy-to-use multimedia "toolkits" so that designers untrained in sound composition or image creation can put together multimedia presentations.

Despite an emphasis on developing tools to easily include media in presentations and applications, the literature contains little discussion or consideration of determining the impact of using particular media or of developing guidelines or structures to describe the various representations of each medium. The material published in the open literature considers the technical aspects of producing media within an application. Articles in trade publications, textbooks, and journals address algorithms for efficient digital storage, compression, and presentation of media. Other articles consider issues related to bandwidth and enhancements of color or sound to improve image clarity.

Some researchers have focused on how people

Learning and Media

Some researchers have addressed how people learn to use media. Hay et al. described working with students using a multimedia word processor. The student work seemed to evolve and become more complex as students became more familiar with the tool and, more importantly, as they developed a mental model for using different media. The researchers concluded that users who have only a short time to work with a new tool will persist in using already existing mental models. In this case, the familiar mental model would be text only.

K. Hay et al., "Students as Multimedia Composers," *Computers & Education*, Vol. 23, No. 3, 1994, pp. 301-318.

Others have researched how people learn with media. Nugent reported a series of studies on the impact various media had on learning, noting that the differences are most acute in combined media comparisons. She stated that separate stimuli components are intrinsic to each medium, and it is the interaction of these components that characterizes each medium. She looked at the impact media had on factual learning across single and dual combinations of visual, audio, and print media, but did not deconstruct the specific attributes of each medium. However, she did point out that the differences of impact on learning within a medium can be as great as the differences of impact between media. She cited research that classifies each medium by whether it supports the main or ancillary information flow.

G. Nugent, "Pictures, Audio, and Print: Symbolic Representation and Effect on Learning," *Ed. Comm. and Tech. J.*, Vol. 30, No. 3, 1982, pp. 163-174.

Salomon asserted that, when looking at the media impact on learning, the way different media structure or represent information is their most important attribute. She felt these attribute characteristics need further study.

G. Salomon, *Interaction of Media, Cognition and Learning*, Jossey-Bass, San Francisco, 1979.

learn to use media, while others considered how people learn using media. A number of researchers have tried to classify media, but with little attention paid to the value added by the media. For research

Media Classification Schemes

A number of researchers have tried to classify various aspects of media, but the literature provides little information on the quantitative value added by various media sources. For example, in his much-cited paper, "Learning with Media," Kozma focused on the technical characteristics of media presentation platforms, from books to video to computers. He was particularly interested in how media affect cognition. He noted that text, especially on the printed page, has a certain stability that allows learners to establish a relationship with the information and use that in their learning. He addressed the impact of combining media in an example describing how novices and experts use text and images to create an understanding of the material. However, he did not address how to create or analyze a media presentation to direct this understanding, nor did he offer guidance on how to select, assemble, and combine media content to promote understanding.

R. Kozma, "Learning with Media," *Review of Ed. Research*, Vol. 61, No. 2, Summer 1991, pp. 179-211.

When describing the impact of graphics in a presentation, the adage "A picture is worth a thousand words" is often quoted, but a methodology or classification scheme for determining the content of the words replaced by a picture is rarely addressed. One notable exception is a description by Tufte, in which he calculated that a graph on the monthly mail workload generated by Congress equals about 700 words of text by comparing the information in the graph to a published news story on the same topic.

E.R. Tufte, *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, Conn., 1983, p. 37.

Molitor et al. suggested that the purpose and context of the message should guide the use of pictures. For example, if the message concerns procedural tasks, pictures are the preferred delivery medium over text.

S. Molitor, S-P. Ballstaedt, and H. Mandl, "Problems in Knowledge Acquisition from Text and Pictures," in *Knowledge Acquisition from Text and Pictures*, H. Mandl and J.R. Levin, eds., North Holland, New York, 1989, pp. 3-35.

Arens et al. considered how to determine which medium should be used to express a specific content or topic. However, they did not assign specific media choices to a data item based on the attributes or characteristics of a particular medium, instead focusing on the nature of the information being presented.

Y. Arens, E. Hovy, and M. Vossers, "On the Knowledge Underlying Multimedia Presentations," in *Intelligent Multimedia Interfaces*, M. Maybury, ed., MIT Press, Cambridge, Mass., 1993, pp. 470-306.

Seidman reflected on the functional categorization of music as it supports total media productions. His categories relate to mood setting or background, and he made the case that including music in a production should be purposeful, not "tacked on." He described how to use various aspects of music, such as the tempo, rhythm, or timbre, to contribute to a presentation's mood.

S. Seidman, "On the Contributions of Music to Media Productions," *Ed. Comm. and Tech. J.*, Vol 29, No. 1, 1981, pp. 49-61.

on how people learn to use media and how they learn with media, see the sidebar "Learning and Media." For information on other researchers' efforts to develop classification schemes for media, see the sidebar "Media Classification Schemes."

In working with graduate computer science students whose area of concentration is the design and implementation of interactive multimedia, we created a media taxonomy to help students better understand the possible media forms they can use in a presentation.² This article presents that taxonomy, its elements, and its possible uses in both the research and development of multimedia applications.

A media taxonomy

The media taxonomy shown in Table 1 presents three categories of expression for each medium. The taxonomy appears as a two-dimensional space, with the increasingly complex media types (text, graphics, sound, and motion) as the vertical dimension. The horizontal dimension encompasses the expres-

Table 1. Media taxonomy.

Media Type	Media Expression		
	Elaboration	Representation	Abstraction
Text	Free text, sentences, paragraphs	Bold, italics, bullets, underlines, headlines, subheads	Shapes, icons
Graphics	Photographs, renderings, scanned images	Blueprints, schematics	Icons
Sound	Speech, audio transcripts	Intensity, tone, inflection	Sound effects
Motion	Raw film footage	Animation, time-lapsed photography	Animated models, highly edited video

sion of each medium in three categories, ranging from a fully delineated Elaboration category to the metaphorical Abstraction category.

The Elaboration category of media expression encompasses any form of each medium in which

Relationship to Other Taxonomies

In describing a taxonomy for music, Pope made a strong case for the usefulness of taxonomies. He suggested that formal specifications help people mark or weigh various aspects of a rapidly changing or evolving field, such as multimedia. Additionally, he noted that the mere activity of developing a structuring framework helps clarify the terms and their usage.

Researchers in the field of computer music struggling with the notion of a taxonomy identified seven hierarchical areas related to composition. Our taxonomy fits well within their organization.

S.J. Pope, Website http://www.mitpress.mit.edu/Computer_music_journal, 1994.

Both Messaris and Weidenmann categorized images based upon their interpretative function: the recognition of salient features (which Messaris called *representational*) or the aesthetic appreciation or understanding of the image's manipulative content (which Messaris called *interpretation*). These two categories are consistent with the media taxonomy categories of Elaboration and Abstraction. Furthermore, in their view images lie along a continuum from concrete to abstract to analogy in much the same way that our media taxonomy provides the classifications of Elaboration, Representation, and Abstraction. Weidenmann provided a set of questions that an image maker should ask: What is the purpose of the illustration? What is its message? Could you summarize the points being made? For what other purpose could the image be used? These questions provide a useful analysis protocol for the media designer for all categories of media.

P. Messaris, *Visual Literacy: Image, Mind and Reality*, Westview Press, San Francisco, 1994.

B. Weidenmann, "Codes of Instructional Pictures," in *Comprehension of Graphics*, W. Schnotz and R. Kulhavy, eds., North Holland,

Amsterdam, 1994, pp. 29-42.

Messaris based his work on Arnheim's concern about viewer interpretation of an image. He contended that no image can perfectly reflect the real world. This analysis extends our understanding of the Elaboration category of media. He listed 10 discrepancies between concrete representational images (which include stills and moving images) and the appearance of things as they are. Knowing about such discrepancies should cause the media designer to be more critical of images, especially elaborative ones.

B. Arnheim, *Visual Thinking*, Univ. of Calif. Press, Berkeley, 1969.

Our media taxonomy could be combined with the list of attributes described by Hunter et al. to categorize graphics: embellish, reinforce, elaborate, summarize, and compare. In either case the taxonomy provides a framework for a design space that accommodates any of these cognitive areas.

R. Hunter, R. Crismore, and P.D. Pearson, "Visual Displays in Basal Readers and Social Studies Textbooks," in *The Psychology of Illustration, Basic Research*, Vol. 1, D.M. Willows and H.A. Houghton, eds., Springer, New York, 1987, pp. 116-135.

Considine and Haley suggested that an expanded definition of literacy is needed to include a new media literacy. The definition, which they said both designers and users must understand, should stress that media do more than merely record reality and reflect it, similar to the taxonomy's Elaboration category; media also create multiple representations of reality, as reflected in the Representation and Abstraction categories.

D. Considine and G. Haley, *Visual Messages: Integrating Imagery into Instruction*, Teachers Ideas Press, Englewood, Colorado, 1992.

no information has been edited out (although information can be added). For example, in a text-based presentation, this category would include a full narrative or transcripts of proceedings. The Elaboration form of graphics would include photographs or scanned realistic images, with no additional manipulation of the images. The Elaboration form of sound might include an audio narrative or transcript of an event, with no removal of problem spots. The Elaboration form of motion would be raw film footage, uncut.

Moving from the fully elaborated form of information to Representation, the second category, the same content is expressed in a more abbreviated, stylized mode within each medium. In this case, the developer has more control over the expression of the information, and the reader/user must adjust to the author's ellipsis. For example, an outline or a series of bulleted items can serve for text. A blue-

print, schematic, or map might represent graphics. A different tone, intensity, or inflection can convey a mood for sound, and an animation or time-altered video might be used for motion.

In the last category, Abstraction, the developer relies on metaphor and culturally common understandings to convey information. Thus, text can work as a graphic, as in a logo or registered trademark, or by using a word to look like its meaning, such as "hill" written as

hill

Graphics could be represented by an icon or well-known symbol. In the sound category the use of sound effects requires the user to make mental connections based on familiarity with those sounds. Animated models or highly edited film clips for motion again require the user to fill in the gaps.

The Abstraction category echoes some of the considerations Bolter³ put forward when he suggested that electronic text is actually a photographic representation of text in contrast to nonelectronic text, which can be described as typographic. While all media are culturally related, perhaps visual media rely more on the culture than does text. Some of this might stem from the fact that text is a standard commodity, while images are seen as more rare and hence more important—even in today's world of image explosion.

Arnheim⁴ delineated these categories similarly when he considered visual thinking. He suggested that straight text—Elaboration in our taxonomy—engages interpretation by inferences, while the other categories, Representation and Abstraction, engage interpretation by perception or metaphor.

To better understand our taxonomy, consider an example in which the presentation explains replacement of a car muffler. Applying Elaboration to the text, the author could include the entire text, in full sentences, from a document on muffler replacement. In the Representation category, the author might use a numbered list of the steps necessary to replace the muffler. Further, the author could use an Abstraction-based presentation of the material by using words and symbols to express the same replacement information.

Clearly, some categories of media presentation work better than others in a given context. In the car muffler example, the Elaboration and Representation categories seem the most straightforward in their application to this process. In fact, one purpose of the taxonomy is to encourage designers to experiment with various forms of each medium, such as abstract pictorial directions, in an attempt to realize the best presentation technique for a given content.

Thus, the taxonomy can be understood across the horizontal dimension as a movement from the concrete to the increasingly abstract expression of a media type. Also included is the notion that this taxonomy represents a movement from little editorial license to increasing author control of the information expression. Finally, the taxonomy can be understood as an increase in design effort from left to right. The Abstraction expression of a medium often takes considerably more effort to design than does the Elaboration expression.

Comparing our taxonomy to others

If we review the media taxonomy in light of the media characteristics defined by Arens et al.,⁵ we can see that the Elaboration category relates well

Media Characteristics⁵

Temporal endurance (permanent or transient)

An indication of whether the created exhibit varies during the lifetime of the presentation.

Granularity (continuous or discrete)

An indication of whether arbitrarily small variations along any dimension of the presentation have meaning in the denotation or not.

Baggage (low or high)

A gross measure of the amount of extra information a consumer must process to become familiar enough with the substrate to correctly interpret it.

Detectability (low, medlow, medhigh, or high)

A measure of how intrusive to the consumer the exhibit created by the medium will be.

Media type (aural or visual)

The type of medium necessary for presenting the created exhibit.

Table 2. Media taxonomy with media characteristics.⁵

Media Type	Media Expression		
	Elaboration	Representation	Abstraction
Text (detectability = low)			
Temporal*	permanent	transient	transient
Granularity	continuous	discrete	discrete
Baggage	low	high	high
Graphics (detectability = medlow)			
Temporal*	permanent	transient	transient
Granularity	continuous	discrete	discrete
Baggage	low	high	high
Sound (detectability = medhigh)			
Temporal*	permanent	transient	transient
Granularity	continuous	discrete	discrete
Baggage	low	high	high
Motion (detectability = high)			
Temporal*	permanent	transient	transient
Granularity	continuous	discrete	discrete
Baggage	low	high	high

* Our definition of temporal includes whether meaning changes over time.

to their descriptor "real world." As we move to the right across the dimension of "media expression," the real worldliness of a particular representation diminishes. The "Media Characteristics" sidebar shows the definition of their media characteristics, and Table 2 shows our media taxonomy augmented with the applicable media characteristics: temporal, granularity, and baggage.

In the temporal category, we extend the meaning to include how the understanding of content changes with time. For example, Elaboration text is permanent, but as we move to the Abstraction

expression of text, the temporal aspect shifts to transient. The Elaboration expression of media is more "permanent" because it is raw transcripts or footage. For instance, a transcript or videotape of a trial proceedings is a permanent record of it. A news story or news clip is transient and can change with each representation.

The characteristics of granularity and baggage relate to the increasing degrees of abstraction represented by the dimension "media expression." Note that for each medium, the overlay of the media characteristics is the same, suggesting an internal consistency within the taxonomy.

The issue of detectability relates more to the medium than to format and is exactly correlated with the four media shown in the taxonomy. Nonaural media will have a lower detectability, with text being the lowest and graphics rated as medium low. The detectability of sound is medium high, and motion has the greatest detectability.

Levin et al.⁶ proposed a functional classification for images—drawn or photographed—in light of an image's effect on content recall. They identified decoration, representation, organization, interpretation, and transformation as categories for pictures accompanying text.

We can apply our media taxonomy to the Levin classification as a way of enhancing both taxonomies. For example, Levin classified as decorative an image designed to make the page more attractive without relating to the content expressed in the other media. The media taxonomy would allow someone who wanted to use an image as a decoration to consider a border (Representation) or icon (Abstraction) as possible choices.

The connection is even stronger when considering Levin's representation category, in which images serve to tell exactly the same story as the other media. In this case the developer could consult the media taxonomy to insure that all media were at the same level of abstraction. Levin's organizational function, usually relying on maps or other diagrams, is analogous to the Representation category of the media taxonomy.

It is also interesting to note that Levin et al. identified a "form" category, called transformation. In this category, information is recorded in a memorable form, not unlike an icon, and in a form that differs from the other media presentations.

One purpose for delineating each medium across three categories of expression is to increase our ability to determine which representation of a particular medium will enable us to transmit information in a way that will increase understanding

and retention by the user of that information. Some readers may question our selection of only four media types. For example, many authors categorize maps differently from graphics. They could extend the taxonomy to include refinements to the media type. In the case of maps, a detailed contour map could be classified as Elaboration, while a scaled map showing stores in a mall might be a Representational map. A cognitive map could be considered Abstraction. In fact, it is the extensibility of this taxonomy that underscores its power.

Using the media taxonomy

The taxonomy in its current form has been used successfully for the past three years in our multimedia graduate program in computer science. Students in the introductory multimedia course we developed² use the taxonomy in designing their media portfolios. They experiment with the various presentation forms of each medium before attempting to combine media into a multimedia application. Students in the advanced research and evaluation course use the taxonomy as part of their evaluation protocols to determine whether each medium in a multimedia presentation is used effectively.

Single-media class projects

In the introductory course, students must develop a portfolio of presentations that use all aspects of each medium to describe some place. The place may be real or imaginary, physical or metaphysical. They focus on one medium at a time to fully understand the technical creation and interactive presentation of information in each medium before putting together multiple media. Students use the taxonomy to develop examples for their portfolio in each of the single medium formats: text, graphics, sound, and video.

When developing the first media element in the portfolio of text, the students' first inclination is to write fully elaborated descriptions of the place, using paragraph form and narrative style. The taxonomy provides guidance about how to go beyond simple narrative to try outlines, headlines, different fonts and formats, and iconic use of text. One student, writing about Korea, chose to use an illuminated manuscript style to describe his homeland. He experimented with color for the anchoring letters of the paragraphs and used a font indicative of Korea. Another student, describing his apartment, followed the style aspect of the taxonomy for text and used the coded descriptions usually found in newspapers for rent listings. Another student,

describing the university food court in text, applied the idea of abstraction in using letters to show the shape of the area (see the sidebar "Using the Taxonomy to Learn MultiMedia Design" pp. 42-43). She also used different fonts, colors, and styles to represent the different kinds of food available, such as Chinese, Mexican, deli, and grille.

In the sound exercise, one student described the whale sanctuary in Antarctica by using narrative, mood music, sound effects of whales singing, and sound as an "earcon" when moving from one part of the presentation to another.

In the graphics exercise, another student described Jamaica by using a map of Jamaica, a schematic drawing of the roads with major tourist stops, and scanned pictures of the island. He also developed graphical icons as navigation buttons.

During the video exercise, a student described a popular local area by shooting video sequences that took on the aspect of someone walking down the main street, with close-up shots of the stores, restaurants, and theaters. During his animation sequence, an animated map showed viewers where they were and allowed them to go in any direction to any location.

Another student used video and animation in his ski school project to show actual footage of different skiing techniques, with an animated overlay of how the skis moved in each lesson. He used the dimensions of the taxonomy to develop his ideas of starting with spectacular long distance skiing shots, then zooming into the ski school where the user could view the lessons in actual time, slow motion, stop action, and with the animated ski overlay.

Using multiple media

The taxonomy also offers a structure in which to understand how to put two media together more effectively. For example, in the two-media exercise in the portfolio, students are asked to combine animation with one other medium to convey the idea of their place. One student describing the Washington, DC Metro system used an animated map of one of the subway lines with the sounds of the subway in the background.

Admittedly, design is more an art form than a science, but guidelines such as the taxonomy can help students as they find their own artistic voice. The taxonomy has helped students experiment more fully with the many aspects of each medium by using the three categories in each of their single-medium explorations. It provides a framework for both students and professional designers to com-

bine different media more intentionally, using the cognitive impact from one aspect of a particular medium to offset or enhance the cognitive impact of a different aspect of another medium.

One challenge that designers face is how to insure that what is said in two media, text and images for example, represent unique, reinforcing expressions of each medium and not mere duplicates of each other. For example, when describing a Smuckers fruit preserves commercial, Messaris⁷ dissected the different roles played by the sound, images, and language to explain how to combine concrete text with abstract images and sound. In his example, the words say "fresh air," while the images and sound convey a farmyard scene. Even though both the sound and the image might be characterized as Elaboration, in this case we can use them as support or metaphor for the text.

In the earlier example on the training presentation about muffler replacement, a designer might select a representational text of a numbered list with a fully elaborated video showing the process from start to finish. Conversely, the designer might present the material with a complete, elaborated textual presentation and use an animation or abbreviated video to focus only on the troublesome or unusual aspects of the muffler replacement. These two design presentations build on the use of two or more media and insure that they do not merely duplicate each other.

If a designer selects two or more media using the same type of expression, the possibility of diminishing the overall impact due to duplication of expression increases greatly. This often occurs in presentations where text appears on the screen at the same time it is read by a narrator. If, on the other hand, design choices create a seamless integration of multiple media using a different type of expression for each, multiple media will strengthen the presentation. A simple example of the power of mixing levels of expression is the rebus, an expression that combines text with appropriate icons.

Using the taxonomy to evaluate media

The media taxonomy can also be used to evaluate multimedia applications. In an advanced course following the multimedia design course, our graduate students focus on the development and implementation of evaluation protocols for multimedia applications. In this course the media taxonomy, added to the technical and cognitive guidelines addressed by other published evaluation criteria, has proven very useful as a framework

Using the Media Taxonomy to Learn Multimedia Design

One of Martin's students, Nadia Abdou, crafted a class project that nicely demonstrates use of the media taxonomy. The process she followed for her project is described here in first person.

Our assignment was to describe "a place" using each medium individually and then put several of the media together with a nice interface. When thinking about the media taxonomy as a guide, I wanted to choose a place that would let me experiment with each medium at all levels. I chose J Street, the food court located in the student center at the University. I decided to use Macromind Director on the Macintosh as my development platform, as opposed to Asymmetric Multimedia ToolBook on the IBM platform.

First we had to use text to describe our place. As my declarative description I wrote a narrative describing J Street. However, to

make it more interesting, I placed the text around the screen so that it actually represented a map of the area. Then I thought about other ways that text would be used, such as for restaurant menus. I designed text menus for each restaurant using different fonts and colors to reflect the different food types, such as Mexican, Italian, and

Chinese. Finally, I designed a logo with a large "J" to represent J Street.

The second assignment required using graphical images only. With the digital camera I took pictures of each restaurant and the type of food it offered. I created a map of J Street as an organizer for the pictures. The map had buttons that caused the pictures of each restaurant to appear on the screen. I also developed graphical icons to represent each restaurant, plus the main menu and the help screen. In this part of my portfolio, all navigation was done using the graphical icons.

In the third assignment we had to use sound only. With the microphone I created an audio narrative about J Street. Then I took a tape recorder and went to J Street to record activity in the area. I recorded small musical selections to represent the flavor of each restaurant. Finally I created "earcons" for the buttons to initiate the sounds. Audio instructions told users how to access the different sounds.

Next, we were to use video to describe our place. I took the lab camcorder to J Street to capture the live sights and sounds. I started with a large pan of the area, used the subjective camera technique so that it appeared as though you were walking through the area, and then created small video clips of each restaurant. I put all this together in a video presentation of J Street. At one point I had several video clips going at once in a collage effect to reflect the bustle of activity always present at J Street. Despite technical difficulties with the video capture, I was pleased with the results.

The next assignment required us to use animation with one other medium. I had the most fun creating this part of my portfolio. I decided to combine the text food menus created previously with animated cartoon food items. For example, I had a tray filling up with food, a pizza rolling across the screen, and fruits jumping into a basket. The effect was amusing and interesting.

Last, I had to decide which of the previously developed pieces to include in the final project. This proved more difficult than I expected. At first, I thought I would just put all of them in, but I soon realized that this would create a confusing, overly busy presentation. So I had to decide

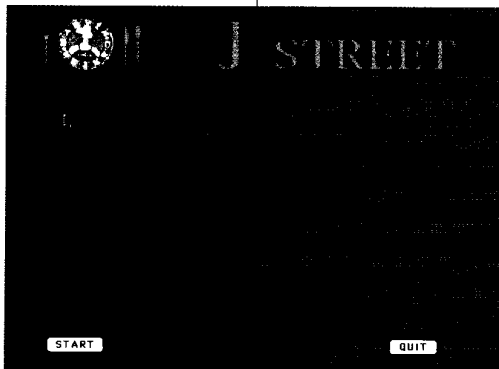


Figure A. The opening screen of "J Street Presentation."

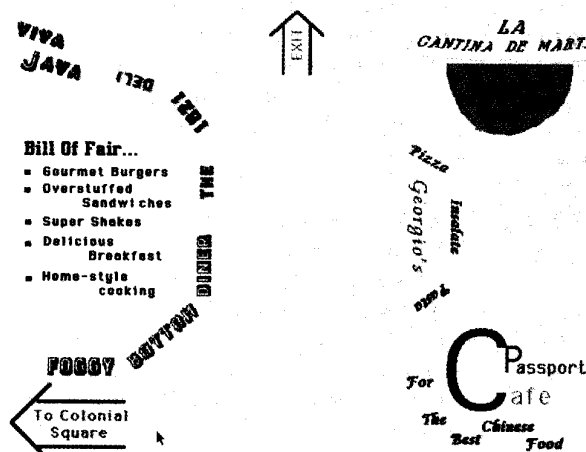


Figure B. The organizing metaphor for "J Street Presentation."

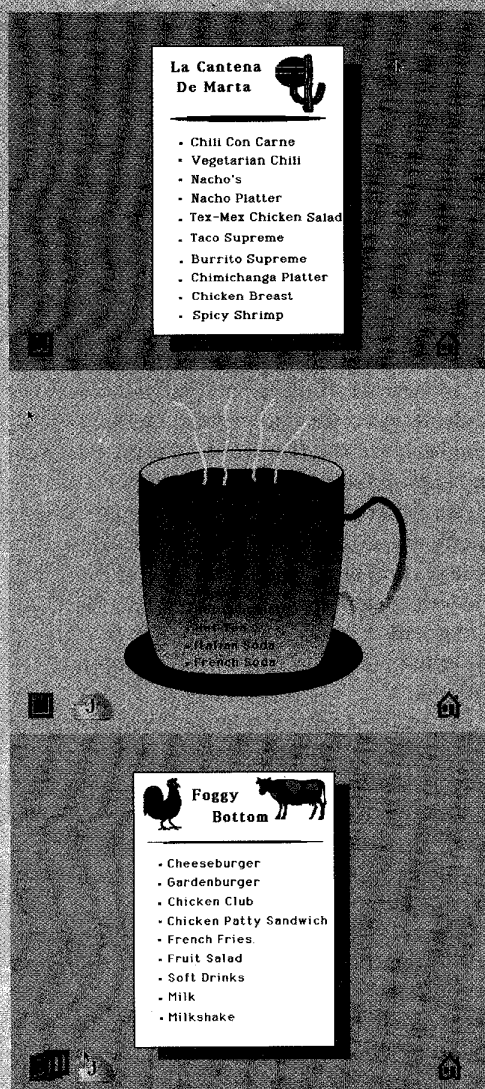


Figure C. Sample menus linked to pictures.

which parts fit together best.

To solve the problem of creating a nice interface to pull it all together, I used the organizing metaphor of a map of the area. The opening screen showed a corrugated metal door with the J Street logo on it (Figure A). When the user clicked "open," the metal door rolled up with appropriate sound effects, revealing the map of J Street with a graphical icon for each of the restaurants. An audio introduction welcomed the user to the area and provided instructions on how to proceed (Figure B).

Clicking on each restaurant icon called up a brief video clip with sights and sounds of that restaurant.

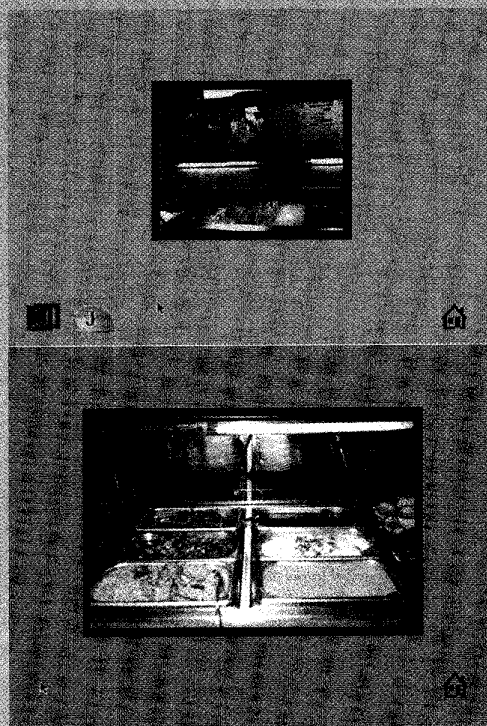


Figure D. Sample pictures of food.

A button let users view the menu (Figure C) and also a picture of the food served (Figure D). The musical clip representing each restaurant played in the background to set the mood. At each stop, the user could return to the map and exit or continue exploring J Street. As it turned out, I did not use my cute animations because they didn't add to the overall presentation and might actually have distracted from it.

In reflecting on how I used the taxonomy to develop my portfolio, I would say that it made me think about using each medium in at least three different ways. Also, for my final project, I tried to think about the different levels of the taxonomy in using the different media. In the opening sequence, I used an elaborated audio narrative as well as representational sound effects and an abstract icon. I used graphics at the representational level in my map and text icons at the abstract level. At each restaurant area I used video, pictures, and text menus to elaborate the area, but also music at the representational level to set a mood and abstract icons and earcons for navigation aids. Having the taxonomy helped me think carefully about how I was going to use each medium and forced me to express my ideas at multiple levels.

Text	1	2	3	4	5
Information	Unclear				Clear
Presentation	Syntactically incorrect				Correct
Sequence of screens	Confusing				Clear
Content	Uninteresting				Interesting
Reading characters on screen	Hard				Very easy
Screen spacing	Crowded				Well spaced

Figure 1. A Likert scale developed for the evaluation of text after considering the Elaboration category of the media taxonomy.

for categorizing the various media aspects of an application. The taxonomy allows the reviewer to identify the levels of design at which the application operates. From that information, the reviewer can measure how well the level is achieved. For example, in our muffler example, if the presentation uses text at the Elaboration level, it is then appropriate to ask how well this is accomplished. Is the text complete and fully represented? If the presentation is iconic, questions can be developed to address the clarity or accessibility of the abstraction or metaphor to the intended user.

To understand the process, we follow two students developing a complete evaluation protocol. The students teamed up to prepare and implement an expert evaluation of the Medical Tutorial program prepared by the university's medical school staff. Other students in the class prepared user evaluation protocols. They wanted to be sure that their evaluation protocol accounted for the impact of each medium as well as the synchronization between and among various media. They wanted to make the survey questionnaire as simple to fill out as possible while making it as complete as possible. Thus they decided to use a Likert scale instrument. As a start, they asked questions to elicit a ranking from 1 to 5, with 1 being ineffective and 5 effective, of the following six items: text, icons, sound, video, pictures, and media synchronization.

During the class peer review session, they were asked how the six questions related to the taxonomy and whether or not someone could really make such a high-level judgment. As a result they immediately expanded the synchronization question to include synchronization between text and sound, between video and sound, and between video and text.

However, the taxonomy can provide more direction than just insuring that each medium used in conjunction with another was validated for synchronization. In that context, the two students were asked to review each medium listed above and decide how better to discuss or categorize the spread between ineffective and effective. When considering text more carefully, one of the

students suggested that if we looked at the Elaboration category—which implied a fully declarative presentation of the material using text—they should include the questions about text shown in Figure 1.

Thinking about the Abstraction category more carefully led them to add items concerning the navigation and instructional messages, the input prompts, and the ellipses of information using outlines and other textual expressions. The team wanted to add items about readability under this category. However, when they put the protocol together for testing, they realized that asking people to indicate whether fully elaborated text was readable was comparable to having them specify whether abstract text was printed using clear fonts and appropriate spacing. On the other hand, the Abstraction category did lead the students to add questions concerning the understandability of the outlines and other modified textual forms. They tried adding a two-part question under "Text" and "Information": on a scale of 1 to 5, was it difficult to follow or easy to follow? Incomplete or complete?

When their peers challenged them to be sure that they had included issues related to the Representation category, they used a question about esthetics to deal with that concern: Under "Text" they added the category "esthetically pleasing," ranging from not attractive (1 on the scale) to attractive (5 on the scale).

The analysis of individual medium by category allowed the students to address the different expressions of a medium and to insure they prepared questions about them for evaluation of the application by multimedia experts. Each medium received the same deep analysis. The result was a robust evaluation protocol that took into account the impact made by each individual medium to the presentation as well as an evaluation of the interaction between the media and the overall integration of media into an application.

Future directions

What this taxonomy does not address is the design process and the order of use of the different media. For example, it does not suggest that an author begin with the visuals of graphics or motion and then add text, nor does it suggest that the author start with text and build from there. These considerations relate to the application's targeted audience. For example, some readers/users prefer to problem solve using visuals first. For them, seeing the intended outcome allows them to better

understand the text. Others like to "read the manual" before attempting to solve any problems.

Further, this taxonomy does not attempt to discern features across media or draw judgments as to whether one medium is better than another. Similarly, the synergy frequently created when several media are combined into a whole larger than the sum of its media parts is not addressed by the taxonomy. Issues related to whether certain kinds of information should be presented electronically are also outside the purview of this article.

Further, this taxonomy provides only minimal guidance against the improper coordination of media, such as presenting text on one topic and an image on another at the same time. Typical questions of media coordination can be found in the paper by Mayer and Anderson,⁸ which identifies the strength of pictures and text in contiguous space and time as an instructional support. They did not address the result of types of images or types of text in this contiguous space. The media taxonomy can supply organizing principles to supplement their research design.

As an organizing principle this taxonomy can be used to understand both design and content messages. The next step is to review the taxonomy in light of various studies in the psychological and cognitive aspects of understanding of multimedia applications to determine whether it can shed light on these areas. Questions such as how we come to understand an image and how the route to that understanding differs from how we come to understand text can be answered in part by using the taxonomy. Finally, the media taxonomy represents an attempt at the formalism needed as the field of automatically generated interactive multimedia presentations continues to grow. **MM**

References

1. K. Hay et al., "Students as Multimedia Composers," *Computers & Education*, Vol. 23, No. 3, 1994, pp. 301-318.
2. R.S. Heller, "Creating an Advanced Degree Program in Multimedia," *Proc. ASEE*, ASEE, Washington, D.C., 1995, pp. 353-356.
3. J. Bolter, *Writing Space: The Computer, Hypertext and the History of Writing*, Lawrence Erlbaum Assoc., Hillsdale, N.J., 1991.
4. B. Arnheim, *Visual Thinking*, Univ. of Calif. Press, Berkeley, 1969.
5. Y. Arens, E. Hovy, and M. Vossers, "On the Knowledge Underlying Multimedia Presentations," in *Intelligent Multimedia Interfaces*, M. Maybury, ed., MIT Press, Cambridge, Mass., 1993, pp. 470-306.

6. J.R. Levin, G.J. Auglin, and R.N. Carney, "On Empirically Validated Functions of Pictures in Prose," in *The Psychology of Illustration, Basic Research*, Vol. 1, D.M. Willows and H.A. Houghton, eds., Springer, New York, 1987, pp. 51-85.
7. P. Messaris, *Visual Literacy: Image, Mind and Reality*, Westview Press, San Francisco, 1994.
8. R.E. Mayer and R.B. Anderson, "The Instructive Animation: Helping Students Build Connections Between Words and Pictures in Multimedia Learning," *J. of Ed. Psych.*, Vol. 84, No. 4, Dec. 1992, pp. 444-452.



Rachelle S. Heller is an associate professor in the Electrical Engineering and Computer Science Department at George Washington University. She earned a BS in chemistry from the State University of New York, Stony Brook, and an

MS in computer science and a PhD in education from the University of Maryland. Her research interests focus on computers in educational settings, including the impact of interactive multimedia on learning, the development of courseware for new application areas, and the integration of computers into existing educational settings and public spaces.

Heller is co-editor of *Computers & Education: An International Journal* and served as the design consultant for the Interactive Multimedia Learning Center of the US Holocaust Memorial Museum in Washington, D.C.



C. Dianne Martin is an associate professor in the Electrical Engineering and Computer Science Department at George Washington University. She earned a BA in math education from Western Maryland College, an MS in computer science

from the University of Maryland, and an EdD in education from George Washington University. Her research interests focus on the use of hypermedia in qualitative research, the social and ethical impact of technology, and applications of multimedia in education and training.

Martin is chair of the ACM Special Interest Group for Computers and Society (SIGCAS) and is an ACM National Lecturer.

Readers may contact the authors at EECS Dept. Academic Center, 6th Floor, George Washington University, Washington, DC 20052, e-mail {sheller, diannem}@seas.gwu.edu.