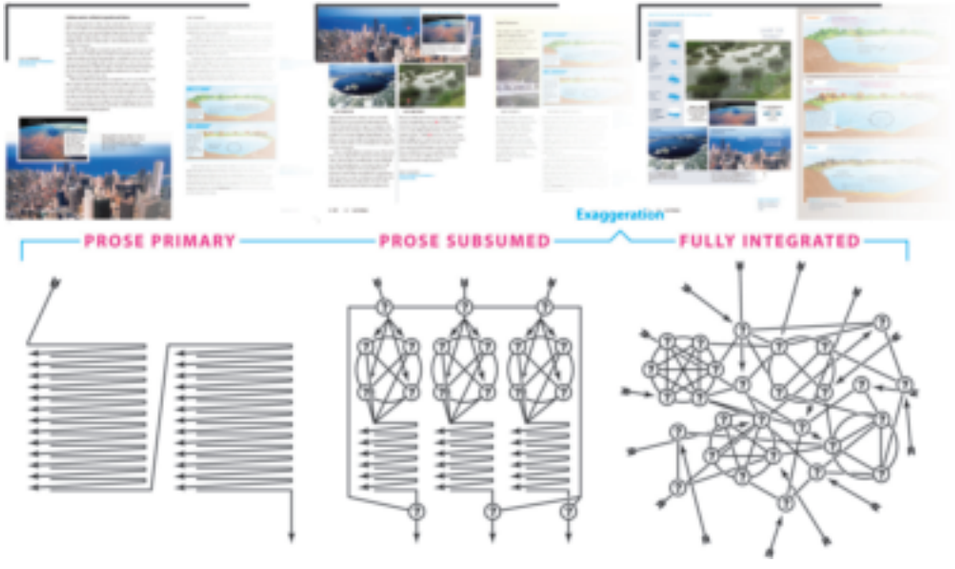
Visible Language

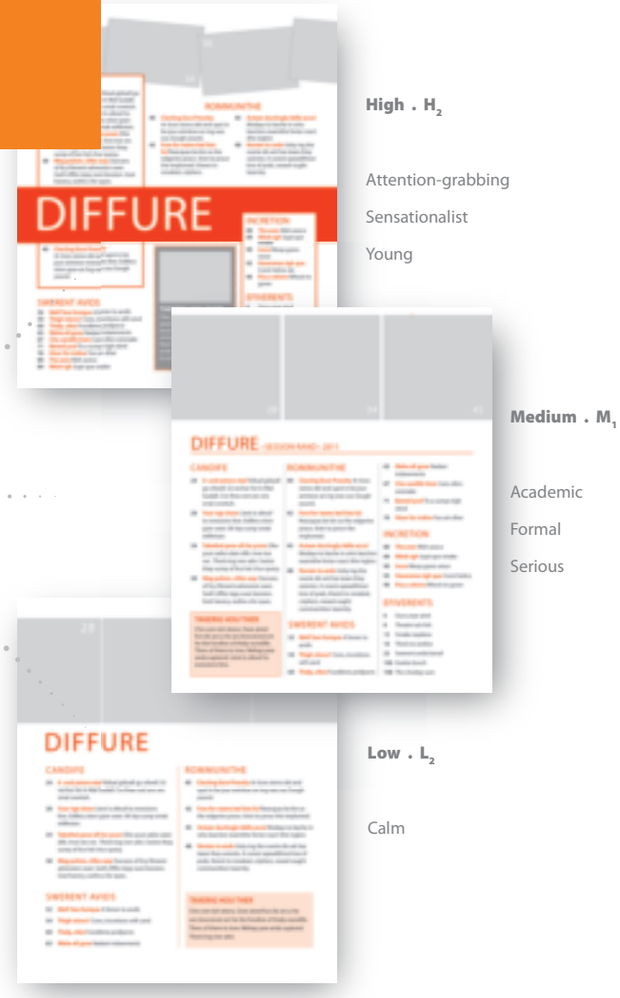
the journal of visual communication research

48 I



THE INTEGRATION OF TEXT AND IMAGE IN MEDIA AND ITS IMPACT ON READER INTEREST

Matthew O. Peterson, Ph.D.



TYPOGRAPHIC LAYOUT AND FIRST IMPRESSIONS

Jeanne-Louise Moys

ISSN 0022-2224

May 2014

Before there was reading there was seeing. *Visible Language* has been concerned with ideas that help define the unique role and properties of visual communication. A basic premise of the journal has been that created visual form is an autonomous system of expression that must be defined and explored on its own terms. Today more than ever people navigate the world and probe life's meaning through visual language. This journal is devoted to enhancing people's experience through the advancement of research and practice of visual communication.

If you are involved in creating or understanding visual communication in any field, we invite your participation in *Visible Language*. While our scope is broad, our disciplinary application is primarily design. Because sensory experience is foundational in design, research in design is often research in the experience of visual form: how it is made, why it is beautiful, how it functions to help people form meaning. Research from many disciplines sheds light on this experience: neuroscience, cognition, perception, psychology, education, communication, informatics, computer science, library science, linguistics. We welcome articles from these disciplines and more.

Published continuously since 1967, *Visible Language* maintains its policy of having no formal editorial affiliation with any professional organization — this requires the continuing, active cooperation of key investigators and practitioners in all of the disciplines that impinge on the journal's mission as stated above.

Website
http://visiblelanguagejournal.com

Postmaster:
send address changes to:
circulation manager name
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016

Published tri-annually in January, May and October

SUBSCRIPTION RATES

United States	Individual	Institutional
1 year	$35.00	$65.00
2 year	$65.00	$124.00
3 year	$90.00	$183.00
Canadian*	Individual	Institutional
1 year	$44.00	$ 74.00
2 year	$83.00	$142.00
3 year	$117.00	$210.00
Foreign**	Individual	Institutional
1 year	$ 56.00	$ 86.00
2 year	$107.00	$166.00
3 year	$153.00	$246.00

Prepayment is required. Make checks payable to University of Cincinnati *Visible Language* in U.S. currency only, foreign banks need a U.S. correspondent bank.

* Canadian subscriptions include additional postage ($9.00 per year).
**Foreign subscriptions include additional postage ($21.00 per year).

ISSN 0022-2224
Published continuously since 1967.
Index included in last issue of volume year.

BACK COPIES

A limited number of nearly all back numbers is available. The journal website at http://visiblelanguagejournal.com is searchable and lists all issues, contents and abstracts.

COPYRIGHT INFORMATION

Visible Language

the journal of visual communication research

48 I

May 2014

48_1

Visible Language

the journal of visual communication research

Contents

Slide Presentations, Seriously

Per Mollerup

ABSTRACT

This article addresses the informative quality of slide presentations in university lectures. The arguments also apply to slide presentations in other situations. The article presents a number of principles to improve the graphic quality and use of slide presentations. These principles build on a review of relevant literature and on the author's experience and reflection. Research in this area is limited in quantity and depth.

KEYWORDS

learning style, bimodal teaching format, multimedia learning, slide presentation, PowerPoint, progressive disclosure

INTRODUCTION

Slide presentations play a major role in teaching and learning at universities. Two conflicting facts characterize the widespread use of slide presentations. They are preferred by lecturers and criticized by students. In principle, two factors can account for this problem. One is the nature of slide presentations. The other is the skill of the presenter. This article begins by considering slide presentations as a format for teaching and learning. Then, it focuses on the quality of slide presentations and how to make them effective.

This article includes references to previous authors on slide presentations. Richard E. Mayer, Professor of Psychology at University of California, Santa Barbara, describes nine theory-based multimedia effects, some of which are cited. Most of the other cited authors reflect on their own experiences as slide presenters. Some have a distinguished academic background and address slide presentations as an interesting subsidiary aspect of academic life. Among them are Stephen M. Kosslyn, Professor of Psychology at Harvard, Robert A. H. Anholt, Professor of Zoology and Genetics at North Carolina State University, and Edward R. Tufte, Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University. I also draw on my own prior publication, a manual written for the International Institute for Information Design (Mollerup, 2011). This article also offers prescription based on my extensive professional experience before taking on my professorial position: four decades as editor of design magazines and managing director of a professional design practice, as well as frequent lecturing.

This article does not argue for slide presentations in higher education or elsewhere. Slide presentations are already an overwhelming aspect of contemporary life. The sheer ubiquity of slide presentations means that their well-known weaknesses warrant a closer look. According to UNESCO (2013, np) there are now more than 17,000 universities in the world. It is a modest guess that each of these universities offers 1,000 slide presentations to an average of 25 students every day. At this rate, at least 425 million students attend 17 million slide presentations every day. To this estimate we can add slide presentations outside academic life. These numbers are modest compared with Web use, but they are still huge.

To position slide presentations among other teaching methods, this article starts by discussing learning from lectures compared with learning from books. It continues by discussing two hybrid forms of teaching that address both hearing and seeing: slide presentations and seminars with handouts. The discussion of slide presentations vis-à-vis other teaching methods delineates the arena where slide presentations should prove their worth. The article continues by describing the basics of slide presentations before dealing with lists, details, and handouts. Some of

the cited statements are contradictory, possibly because they have their origins in different didactic situations. One example of this concerns reading aloud from text on the screen. Authors are divided whether this is a good idea. Personally, I do and I don't read text aloud from the screen. It depends on the situation. After all, there is a subjective element in any personal presentation.

Skilled slide presenters sometimes draw on their expertise to deviate from the principles in this article. They can make idiosyncratic presentations in the same way that skilled writers can sometimes present their messages in particular ways with great effect.

UNIVERSITY LECTURES OR BOOKS?

Seventy-five years ago, Virginia Woolf wrote that university lectures are "an obsolete practice dating from the Middle Ages when books were scarce" (Woolf, 1938, chapter 1, note 30). While many individuals agree, university teachers around the world think otherwise. They base most education on oral lectures where one lecturer talks face-to-face to a multitude of students. In spite of the widespread opinion that lectures rank low on effective learning, old-fashioned lectures do have some value. Several possible benefits relate to lecturer-student contact and to the contents:

The lecturer gets contact with the students.
Students meet a person enthusiastic about their subject.
Students see the lecturer as a role model.
Students can ask questions and get immediate answers.
Some spoken content is not found in written material.
The lecturer emphasizes content that will prove useful at exams.

Two of these benefits, the lecturer's contact with the students and the students' opportunity to ask questions, probably come in inverse proportion to the number of the students attending the lecture. What the university enjoys as an advantage of scale works to the disadvantage of the students.

Some drawbacks of lectures relate to their timing. Students must follow the pace of the lecturer. They cannot speed up or slow down, they cannot stop to digest, and they cannot go back. There is no fast-forward, stop, pause, or rewind. These issues are especially apposite to lectures dealing with difficult subjects. What is only heard is easily forgotten; therefore students take notes during lectures. This means that the students much of the time think about *what has been said* while trying to listen to *what is being said.* This problem can be overcome or reduced if the students are told that they will get comprehensive handouts after the lecture.

Compared with lectures, books have some obvious benefits:

Students can read and digest the material at their own speed.
Students can stop reading and resume reading as they prefer.

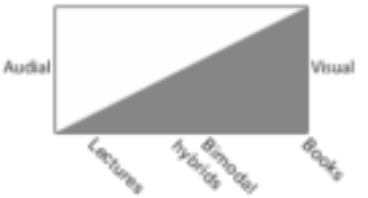
Students can go back to repeat reading when needed.
Students can make notes, which increase the value of the book.
Students are not disturbed while making notes.
Students have a greater chance of understanding difficult subjects.

The students' preferred learning style influences the relative importance of specific benefits. The benefits of oral lectures do not depend exclusively on listening. Attending oral lectures also involves seeing, seeing the lecturer, and seeing the visuals the lecturer might present. Visuals have improved greatly. 25 years ago, the lecturer would talk and use chalk, whiteboards, or overhead transparencies; today the lecturer will typically present PowerPoint slides while talking. While PowerPoint presentations are used everywhere in university teaching, they have some notable weaknesses (Edward Tufte, 2003).

Slide presentations are bimodal hybrids. So are seminars discussing handouts. These didactic formats have been adapted in university teaching to reinforce the spoken word.

Speech-only lectures, bimodal hybrids, and books address a continuum of learning styles that moves from pure listening to pure reading, from hearing to seeing.

FIGURE 1.



Teaching methods using audial and visual learning styles

On-line teaching involves the full continuum of audial and visual learning styles. We will not discuss on-line teaching here, but many of the arguments that follow apply to on-line teaching as well as to slides.

SLIDE PRESENTATIONS OR SEMINARS?

Lectures and books each have their advantages and disadvantages. To combine advantages and exclude or reduce disadvantages is the purpose of hybrid presentations. Slide presentations and seminars with handouts are two cases in point.

Slide presentations are oral lectures accompanied by PowerPoint or similar kinds of slides projected on a large screen by a computer. They can have any number of participants. In universities there are sometimes up to 600 students. Seminars are defined here as meetings with a lecturer and a relatively small number of students, often less than 25. Seminars often discuss handouts. In many cases, these handouts

reproduce slides. In other cases, handouts are custom-designed. Proponents of seminars take delight in the etymology of the word: *seminarium*: seed bank, seedbed, a place where seeds are planted. The sheer size of the audiences gives slide presentations a cost advantage while seminars have a quality advantage in terms of intimacy and students' propensity to ask questions, and possibility to get in-depth answers.

Due to the problems of distance viewing, slides have a limited capacity to show detailed content with required readability. Printed handouts that accompany slide presentations or seminars overcome these limitations. Handouts can be printed in an appropriate format and be as detailed as human vision allows.

In slide presentations, the lecturer, in principle, has full control of viewer attention, no matter how many slides. Students cannot look at the wrong slide if it is not shown. In seminars, lecturers don't enjoy this level of control. Nevertheless, a smaller audience and a smaller number of handouts make it easier for the lecturer to see that the students are on the same page – (literally).

In a polemic critique of PowerPoint, the most common computer software for slide presentations, Tufte (2003) strongly advocates seminars with detailed handouts instead of PowerPoint presentations: "For serious presentations, it will be useful to replace PowerPoint slides with paper handouts showing words, numbers, data graphics, images together. High-resolution handouts allow viewers to contextualize, compare, narrate, and recast evidence" (Tufte, 2003, p.8). Many factors may influence the choice between slide presentations and seminars. Seriousness should not be among them. Slide presentations and seminars can both be serious didactic presentation forms.

Andrew Abela (2008) distinguishes between two presentation idioms: ballroom style and conference room style. The purpose of ballroom presentations is to inform, impress, and entertain a large audience. Ballroom presentations are colourful, vibrant, attention grabbing, and noisy. Ballroom presentations are a one-way communication format and should use projected slides. The purpose of conference room presentations is to engage, persuade, and drive action in a smaller audience. They are black and white, with lots of details. Conference room presentations are two-way and should use printed handouts.

According to Abela the biggest single mistake presenters make is to confuse the two idioms. "The main determinant of which style to use is whether you are trying to persuade a small audience, in which case you should use conference room style, or whether you are trying to inform or entertain a larger audience, which would instead call for ballroom style" (Abela 2008, p.107). Neither of Abela's presentation idioms accommodates universities' needs for bimodal presentations.

In some situations, flipped classrooms may be an alternative to slide presentations. Flipped classrooms is a didactic format

where the students are given short video lectures to study at home before the class and then use the time in the class for discussion, exercises, or project work. (Educause, 2012). Like good slide presentations, the flipped classrooms format demands careful preparation.

BASICS

A slide presentation is a series of slides projected on a large screen while the presenter talks. The term slide presentation is used for both the performance and for the series of slides to be presented. Slide presentations are produced by a computer and projected by a computer. Before the computer era, most slide presentations consisted of 35mm slides or overhead transparencies. PowerPoint from Microsoft Corporation is the most widespread software used for slide presentations, so much that PowerPoint is used as a synonym for slide presentation. Other presentation software products exist, notably Apple iWork Keynote. Some authors are PowerPoint specific and explain in detail how certain effects and whole presentations are made using PowerPoint (Rotondo & Rotondo, 2002; Bunzel, 2007; Atkinson, 2008).

The intended function of PowerPoint and similar software products is to assist a lecturer with projected slides while the lecturer talks. However, some authors suggest other uses of PowerPoint such as printed reports and material for websites, and sometimes criticize the software for not being good at these. Tufte (2003) discusses at some length how badly a deck of PowerPoint slides is suited for exchanging technical information. Tufte (2003, pp.7-11) substantiates his argument by relating a disastrous exchange of PowerPoint slides between Boeing and NASA – and internally in NASA – preceding the Colombia Space Shuttle disaster.

Tufte also addresses the role of PowerPoint in assisting speakers when he criticizes PowerPoint for distributing information sequentially in time rather than simultaneously in space. "When information is stacked in time, it is difficult to understand context and evaluate relationships. Visual reasoning usually works more effectively when the relevant information is shown adjacent in space within our eyespan" (Tufte, 2003, p.4). Nevertheless, lectures are sequential; they involve words distributed in time.

Work with slide presentations has three phases: planning, design, and delivery. While recognising the importance of planning and delivery, this article addresses primarily the design phase. Most slide presentations fail here, but they can be much improved by applying a few design principles. Planning deals with crafting the story. The lecturer should bear in mind that the audience come to hear, not to read. Slides are assistants and should be used as such. Slides should not replace lecturers, but support them by helping to make their message perceivable, understandable, and convincing. A slide presentation can include three types of slides or combinations of these: text slides, image slides, and break slides.

Among other tasks, text slides can overview the lecture, clarify main points, clarify new words, clarify important concepts, present definitions, and recap earlier topics. "Text slides should be used for text that supports what is said. The text on screen should be as short as possible, cues, rather than full sentences. The common practice of lifting complete sentences from the paper into the slides….is not helpful" (Salmond & Smith, 2011, p. 590). Text on slides is problematic when it is difficult to read, and when speech interferes with it. These two mishaps often come together. Text on slides should be readable. Most presenters neglect this fact. They show too much type in sizes too small in too short a time. Low colour contrast and badly legible type often worsen the situation. Designers not totally accustomed to doing slide presentations should test their slides projected on the lecture room screen. What is perfectly readable on a computer screen may be unreadable, even invisible, on the screen in the lecture theatre.

There are several views on reading aloud from text on slides: "Don't, however, make the classical mistake of actually reading your slides to the audience. There is nothing worse than that" (Kapterev, 2011, p.118); "If a slide contains complete sentences, it is practically impossible for even the most accomplished presenters to avoid reading the entire slide word for word" (Altman, 2007, p.9); "When you read your slides word for word, you sound like an idiot" (Altman, 2007, p. 9). Kosslyn (2007) does not agree. "I read the slide aloud, telling the audience that I'll read a set of directions, such as the ones they are about to see" (Kosslyn 2007, p.42).

Rather than reading aloud from text on slides, the lecturer should in most cases give the audience time to let them read at their own speed. There are exceptions: the lecturer can read aloud absolutely succinct cues, and confirm difficult words. Showing a wordy text slide while saying something else is anathema to good presentation. Nobody, absolutely nobody, can listen to one message and read another message exceeding a few words at the same time. The presenter should in principle give the verbal presentation orally. Text on the screen should only support the spoken presentation. There is no exact rule on how much text a slide can contain. It depends on the way it is presented. If in doubt, use as little text as possible.

Image slides are used to show images, when images are better than spoken or written words in helping the audience to understand what the lecturer presents. Image slides should preferably show only one image at a time, perhaps with a short text. "Each visual image should illustrate a single point and, like the presentation itself, have only one focus" (Anholt, 2006, p.76). Image slides together with the presenter's speech may actuate the multimedia effect as described by Mayer (2002) "learners perform better on transfer tests when they receive an explanation in words and pictures rather than in words alone." (Mayer, 2002, p.105). By extension (Mayer testes the use of animation, not slides), the modality effect may also be in play: "A modality effect (for transfer) occurs if students

perform better on subsequent transfer tests when the words are spoken rather than printed, that is when they receive animation and narration rather than animation and on-screen text." (Mayer, 2002, p.118).

Break slides are used when there is nothing to show and leaving the previous text or image slide visible distracts the students (Anholt, 2006, p.74). Break slides let the students address their full attention to the lecturer. Break slides are in principle blank, but they may contain a few words such as 'More to come' or another sign that confirms that the break is intentional.

Mayer (2002, p.113) shows that "the knowledge construction process is facilitated when extraneous information is excluded from the presentation" and calls the resulting improvement the coherence effect. As types of 'extraneous information' Mayer includes sounds and music. In line with Mayer's findings designers of slide presentations should refrain from presenting extraneous information in the form of ready-made graphic solutions, animations, or transitions offered by PowerPoint, Keynote, and other software providers. These are worse than useless. They disturb communication and distract from the real content. A good presentation doesn't need icing. "PowerPoint templates (ready-made designs) usually weaken verbal and spatial reasoning" (Tufte, 2003, p.1). Not all authors agree. "Well-chosen effects do polish a presentation", "An arrow that spirals in and points at a particularly critical data point helps focus attention and can be very effective, but such effects should be used sparingly" (Anholt, 2006, p.98).

Slide presentations are used in many sectors outside universities: business and industry, public administration, primary and secondary education, and the military. Most authors on slide presentations seem to gain their experience in the corporate world. They tend to suggest presentations that address feelings rather than understanding. A few authors, such as Anholt (2006) and Kosslyn (2007), focus on academic and scientific presentations.

Anholt (2006) deals exclusively with scientific presentations. He illustrates his arguments with numerous examples of work with students and conference presentations. The graphic design of slides is a minor consideration for Anholt. While Anholt is a great admirer of PowerPoint, he offers no systematic advice concerning the graphic parameters in play. Advice like "Lettering on slides can never be too big!" (Anholt, 2006, p. 80) is obviously misleading. It should rather be: type should be big enough to be read by the whole audience, not bigger. Type that is too large reduces the available space.

Psychologist Stephen Kosslyn (2007, pp.3-18) offer advice that stands out because of an analytical approach based on eight psychological principles that all slide presentations should respect. The eight psychological principles serve three major goals: 1) Connecting with the audience, 2) Directing and holding viewers' attention, and 3) Promoting understanding and memory.

Connecting with the audience is supported by the principles of relevance and of appropriate knowledge. The principle of relevance states that the slides should include only relevant information; the principle of appropriate knowledge suggests that the slides should neither talk down to, nor over the heads of the audience. Language, displays, and concepts should be understandable to the audience (Kosslyn, 2007, pp.4-6).

Three principles support directing and holding attention. The principle of salience suggests that important material should be shown extra clearly. The principle of discriminability suggests that difference should be shown clearly. Finally, the principle of perceptual organization means that viewers see elements that are shown together as belonging together (Kosslyn, 2007, pp. 6-9).

Three principles support understanding and memory: The principle of compatibility suggests that form and content should harmonize. The height of soldiers should be shown in a vertical, not horizontal, bar chart [this author's example]. The principle of informative changes means that change in meaning should be shown by change in form, and that change in form should illustrate change in meaning. Finally, the principle of capacity limitation tells us to respect the limits of viewer perception and cognition (Kosslyn, 2007, pp.9-12).

Kosslyn's eight principles
Connecting with the audience
relevance
appropriate knowledge
Directing and holding attention
salience
discriminability
perceptual organization
Promoting understanding and memory
compatibility
informative changes
capacity limitations
Mollerup's ninth and tenth principles
graphic restraint
semantic chunking

Kosslyn's eight principles would benefit from a ninth principle suggesting graphic restraint. Kosslyn's (2007) examples would benefit from this principle. Text in the examples tends to be unnecessarily large. The result is lines that are too short with inappropriate separation of words that belong together. Beyond a certain limit, text does not become easier to read because it is larger. Neither does the presenter become easier to hear because he shouts. Also, several of Kosslyn's slides (2008, pp.28, 29, 38, 41, 43) would benefit from semantic line change: one line one chunk of meaning, one chunk of meaning one line. A tenth principle could be the principle of semantic chunking.

While some who use slide presentations outside universities may want to entertain, and may have wide parameters to do so, academic and scientific presenters should be more interested in facts than in bells and whistles. Facts should speak for themselves. Too much design may disturb the message, giving the presentation the tone of a sales pitch. This does not mean that academic presenters have nothing to learn from authors with business background, especially

those who show some restraint. Garr Reynolds (2008) recommends slide presentations inspired by Zen principles: restraint in preparation, simplicity in design, and naturalness in delivery. Nancy Duarte (2008) specialised in advanced slide presentations for the corporate world offers healthy advice that is also applicable in higher education.

TECHNICALITIES

The typeface used in designing a slide presentation must also exist in the computer used for projection. If it does not, the computer projecting the slides will substitute a typeface with uncontrollable results. In practise this means that slide presentations should use a typeface found in the Microsoft Office software package. This problem does not exist if a slide presentation is presented as a PDF.

Slide presentations are not suitable media for showing new, experimental, or delicate typefaces. Typographic subtleties should be left to print on paper. A robust sans serif typeface such as Ariel or Calibri does well projected on a screen. These typefaces do not use the delicate details that easily become unclear. Bold type and italic type should only be used for emphasis or not at all. They are not as easy to read as ordinary type.

Capitalisation, use of upper case letters, should take place with restraint. Words written exclusively with capitals are more difficult to read than words written with lower case letters. Sentence case should be used: only the first letter of the first word in a sentence or a string of words is capitalized, except for proper nouns and other words which generally have capitalized first letters. Kosslyn (2007) uses title case in his examples: he capitalises the first letters of all important words, but does not explain why.

Readability determines the type size to be used for slide presentations. Type that is too small is unreadable or difficult to read. Nevertheless, type that is larger than necessary is not necessarily easier to read, but it occupies too much space and often separates words that belong together. Visual shouting does not drive an argument home. On the contrary, the audience may feel attacked, asking what are they going to sell? Unless special conditions suggest otherwise, 24-point text is perfectly readable on the big screen. The size of the room in which the presentation takes place should not play a role here. Larger lecture rooms have larger screens.

Leading, the vertical space between text baselines, should be ample, especially if the lines are longer than a few words. 26-point or 28-point leading works well with 24-point type.

Used with moderation type emphasis is a means for guiding the audience. If everything is emphasized, nothing is emphasized. On text slides, the designer can emphasize one or more words to guide attention and to indicate importance. Kosslyn (2007, pp.66-67) suggests that emphasis should also be used to indicate that some words

belong together. The tools for emphasizing type are colour, size, weight, italics, capitals, and underlining. Some presenters use several of these typographic tools at a time. This is bad attention economy. One form of emphasis is enough. As a rule, that form should be colour. Clear yellow or clear green type gives crisp effect on black background. On a white background, green, blue, and red type will serve for emphasis. "Should it become necessary to highlight the particular bullet you are talking about, there are right ways and wrong ways to do it. Dimming everything except that bullet is the wrong way. Showing everything and highlighting the current one is the right way" (Altman, 2007, p.48). Underlining for emphasis should be used with care. In most typefaces, including Ariel and Calibri, underlining collides with the descenders of the lower case letters g, j, p, and q.

Modern typography is basically an asymmetrical affair. Asymmetrical text on slides running from left to right complies with our habit of reading in a Z-pattern. Slides are not tombstones. They should not be symmetrical. Symmetrical arrangement of a text of more than one line means that readers start each line at a different horizontal position, reducing readability.

As a rule, text on slides should begin at the top of the screen. One reason is that beginning at the top allows text to begin at the same position on all slides, no matter how much text. This adds to consistency and clarity of the presentation. "Uniformity of style throughout the presentation accentuates and underscores the flow and coherence of the talk" (Anholt, 2006, p.80). Another reason for text to begin at the top of the screen is that the lower part of the screen is occasionally not visible from all seats. Images that are less wide than the slide can be centred horizontally. As a rule they should be positioned at the top of the screen for the same reasons that apply to text,

Slide backgrounds should be unobtrusive, remaining in the background. They should not steal the picture. Background patterns are visual noise. It is best to avoid them. "Using a single background gives all your slides a uniform look, but it also prevents you from using a range of design techniques to visually highlight the most important information" (Atkinson, 2008, p.38). White, grey, and black are good background colours for most slide presentations. They are compatible with all other colours. "Here is a good description for your background: It's white." (Kapterev, 2011, p.117); "A black slide background lets bright text stand forward with maximum colour contrast. The content catches the eye, not the background" (Mollerup, 2011, p.28); "Create black backgrounds" (Altman, 2007, p.33). Other background colours can be used but should always be tested for functionality. Anholt offers two viewpoints: "One should choose a quiet, muted background, which should be uniform throughout the presentation" (Anholt, 2006, p.96). "Enhanced backgrounds done in good taste render a presentation classy. However, the border between good taste and gaudiness may not be obvious to all times" (Anholt, 2006, p.97). Following Anholt's ad-

vice may compromise the commitment to simplicity: "A complicated design wastes not only your time but also the audience's attention" (Kapterev, 2011, p.116).

LISTS

Slide presentations have often been identified with bulleted lists, and criticism of slide presentations invariably attacks bulleted lists. There have no doubt been too many slide presentations with too badly designed bulleted lists. "Projecting slides with text bullet points and/or irrelevant graphics such as clip art during your presentation will likely have worse results than speaking with no visual aids at all" (Abela, 2008, p.90). Some information becomes clearer when represented in lists. Other information doesn't. The function of lists in slide presentations is to help the audience to see and remember important issues that belong together. If the lecturer is talking about trade among the nations in the Baltic Region, it might be helpful for the students to see the names of these nations.

In addition to the principle of belonging, lists in slide presentations can show sequence by showing what comes first and what next. A list showing the outline of the lecture as a horizontal line on the top of the slide is a simple instance of this:

Question | Details | Solution | Action

The designer should emphasize the current part by colour. This gives the audience an overview of the lecture and facilitates monitoring the progress of the presentation. Finally, lists can show hierarchy, what is more and what is less:

Kingdom
Phylum
Class
Order
Family
Genus
Species

Whatever their function, lists on slides become troublesome when they become too long and when points have too many words. Several authors recommend restrictions: do not have more than so many lines, each with so many words. Different authors offer different maximum number of lines and words: 8x8 (Rotondo & Rotondo, 2002, p.58); 7x7 (Reynolds, 2008, p.130); (6x6) Forrest, ny, np); 4x4 (Kapterev, 2011, p.119). McKinsey & Co. reportedly recommend maximum three lines, each with maximum three words (Salmond & Smith, 2011, p.589). Rather than more or less arbitrary rules, parsimony should be the ruling principle: Use as few lines as possible with as few words as possible. Much depends on the circumstances of the presentation including how the list is shown in relation to the speech.

Lists need no bullets if the points only have a few words each and do not fill more than one line. Numbers instead of bullets may be useful if the list involves a sequence. It is best to avoid hierarchical lists with several layers of indents. They may be difficult to understand on paper and they are even more difficult on slides. They often seem designed to please the presenter rather than to inform the audience. Using different varieties of bullets for different hierarchical layers do not change this.

Vertical lists, bulleted or not, are often best presented as sequential disclosed lists, lists with progressive disclosure. The speaker should reveal points on the list when talking about them, not before. This prevents visual distraction and reading ahead of the presenter. This principle respects the sequential nature of the lecture and limited audience attention. Dimming (greying) completed points on the list can reinforce this principle. Dimmed points show that they are already passed, but they can still be read. "I prefer to have text built sequentially as I'm not sure why anyone would want the audience to jump ahead. Remember, if the audience can see your bullets, they know the points you're going to make. They'll get bored or agitated waiting for you to catch up with them" (Duarte, 2008, p.145). Progressive disclosure lists, with colour emphasize of the current point and/or dimmed passed points, mean extra work to the slide designer, but pay off with improved audience attention. Kosslyn (2007, p. 11) recommends the principle of progressive disclosure applied to complex illustrations; he lets the illustrations grow while explaining orally the single parts.

If the points of a progressive disclosure list stand for progression, it may occasionally be a good idea to show the list as a staircase moving from lower left to upper right. An example could be George Pólya's *Four steps to problem solving* (2013, pp.1-4):

Looking back
Carrying out the plan
Devising a plan
Understand the problem

Not all authors recommend progressive disclosure lists. Rick Altman (2007 pp.42-43) presents three arguments against "click-by-click bullet advancement". His first argument is that the audience will lose the context if they don't see the full list in advance. This argument leaves no room for lectures that build a context sequentially. Altman's second argument is that the presenter may forget which bullet is the last. Yes, but the list will remind the presenter. Altman's third argument is that spoon-feeding information insults the audience. By this standard, all lectures are insulting because they are all delivered one word, one sentence, and one argument at a time. Altman argues that sequential disclosure lists combined with dimming past points represents "a nadir of growing lists" (2007, p.43). Tufte warns against sequential disclosure lists read aloud: "Worse is the method of line-by-line slow reveal. Beginning with a title slide, the presenter unveils

and reads aloud the single line on the slide, then reveals the next line, reads that aloud, on and on, as stupefied audience members impatiently await the end of the talk" (Tufte, 2003, p. 23).

Sometimes it is preferable to show a full list at once. Kosslyn (2007, p.32) suggests providing a road map, the conceptual structure of the presentation, immediately on one slide and highlighting the single points as the lecture progresses by dimming the other entries. Yates & Orliowski agree: "The typical introduction includes a slide containing a preview of the talk's structure" (2007, p.15).

Slide presentations consisting exclusively of long bulleted lists, with progressive disclosure or not, may be soporific. Consistency is good; but so is variation. Sometimes, series of bulleted lists can be broken with other types of slide content, perhaps a relevant illustration. If such an illustration prevents the audience from dozing off, it is functional. "[A] variety of visuals will keep your audience's interest and attention" (Kosslyn, 2007, p.21).

DEALING WITH DETAILS

A crucial factor in choosing lecture form is the need for discussing detailed visual material. A seminar format may be better when the subject matter and the kinds of material considered are served better by documentation on paper handouts – not prints of slides. This format can possibly be combined with an introductory slide presentation.

What if a slide presentation is the only available option? Practically all authors recognize the limitation of slide presentations for showing such detailed information as tables, formulas, and technical drawings. "it is better to be clear than technical. If presenters are clear, then they may induce audience members to read their written work, which is where technical competence and flair really shine" (Salmond & Smith, 2011, p.592). Presenters can deal with detailed information in at least three ways. Sometimes, a table with many figures for use in a slide presentation can be represented by a simple graph with few details. Sometimes, emphasizing important parts or omitting less important parts can simplify a complicated technical drawing. Sometimes, a complicated technical drawing can be divided into two or more less detailed drawings. "Figures composed of multiple panels should be avoided. Instead, the individual panels should be presented sequentially as separate images" (Anholt, 2006, p.76). "It is important to avoid showing tables. Tables containing rows and columns of numbers are an excellent way to document data in written form, but nobody in the audience can read, compare, and analyse tabulated data points during an oral presentation. Instead, the data should be converted into a bar graph or, if possible, a line drawing…" (Anholt, 2006, p.76). "An unreadable formula is waste of time, even if it is substantively appropriate. This may require showing only the most important or novel part of a formula rather than the whole thing" (Salmond & Smith, 2011, p.592). "One way to

simplify intricate diagrams is to start by showing the basic components and gradually increasing its complexity" (Anholt, 2006, p.89). "The less busy a figure appears, the more justice it does to the information it attempts to communicate" (Anholt, 2006, p.79).

HANDOUTS

If the subject of a slide presentation is complex and complicated, the lecturer should carefully decide what to present on the slides and what to show on a handout. "It is helpful to provide audience members with at least one mode of information that allows *them* to control the order and pace of learning – unlike slides and unlike talk. Paper handouts for talks will help provide a permanent record for reviews – again unlike projected images and talks" (Tufte, 2003, p.23).

A handout can considerably enhance the benefit of a lecture, which otherwise may be a short-lived experience. A handout can be a complete manuscript, or selected parts. It should, as a rule, not be a deck of printed slides. Slides are produced to support the presenter while talking; they are not stand-alones. Handouts should be distributed after the presentation to avoid competing with the presenter. If the presenter wants to discuss something during the presentation, which can only be shown on paper, distribute the handout at the appropriate time for discussion. There may be exceptions to this rule: "It is often advantageous to make instructional handouts as lecture notes available before the actual talk to give attendees an opportunity to familiarize themselves somewhat with the content of the 'upcoming attraction'. They will then be able to focus all of their attention on the lecturer without being distracted by the handout." (Anholt, 2006, p.104).

MANUSCRIPT

When planning a slide presentation the lecturer should also consider the role of a manuscript. A manuscript for an academic presentation serves several purposes. It lets the presenter prepare what to say, it serves as an aide memoire, it facilitates preparation for repeated presentations, and it may serve as a handout. If strictly necessary, and only then, the lecturer can read loud from the manuscript.

Most presentations are best if the lecturer can say what should be said without reading from a manuscript. Even though the lecturer may sometimes miss a word or two when speaking without reading from a manuscript, the loss is offset by a livelier presentation. Reading from a manuscript reduces contact with the audience. Lecturers who read generally speak in a monotonous tone, too fast, or too low. Carefully prepared cue notes on cards or a mind map may be helpful. Occasional glances will not harm the presentation in the same way that reading from a manuscript does. Using the PowerPoint notes function on the lecturer's computer screen tends to focus the presenter on the computer screen rather than on

contact with the audience. Slides may serve as prompts, but they should never show a full manuscript to be read from. Prompts should be a nice side effect. Slide presentations are made for the audience, not for the lecturer.

Most of the principles presented in this article draw on the experiences and reflexions of a great number of slide presenters. However sensible, these principles would gain in usability of they were substantiated by robust research. Different kinds of subject matters, different kinds of audiences, and different didactic situations offer a great field for future evidence based research. Well executed and well communicated these research results could affect some of the more than 17 million presentations given every day to 425 million students, not to mention slide presentations outside academia.

REFERENCES

Abela, A. (2008). *Advanced Presentations by Design*. San Francisco, CA: Pfeiffer.

Altman, R. (2007). *Why most PowerPoint presentations suck*. Pleasanton, CA: Harvest Books.

Anholt, R. R. H. (2006). *Dazzle'em with Style.* Burlington, MA: Elsevier.

Atkinson, C. (2008). *Beyond bullet points*. Redmond, WA: Microsoft Press.

Bunzel, T. (2007). *Solving the PowerPoint predicament.* Indianapolis, IN: QUE.

Duarte, N. (2008). *Slide:ology*. Sebastopol, CA: O'Reilly.

Educause (2012). *Flipped classrooms.* Downloadable from http://net.educause.edu/ir/library/pdf/ELI7081.pdf

Forrest, D. (ny). *How can I use PowerPoint more effectively?* Retrieved from-www.tltc.ttu.edu/teach/.../using_powerpoint.asp. Accessed 2013 December 1.

Kapterev, A. (2011). *Presentation secrets*. Indianapolis, IN: John Wiley & Sons.

Kosslyn, S. M. (2007). *Clear and to the point.* New York, NY: Oxford University Press.

Mayer, R. E. (2002). "Multimedia learning." In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory.* Vol. 41. San Diego, CA: Academic Press.

Mollerup, P. (2011). *PowerNotes: Slide presentations reconsidered.* Downloadable from http://hdl.handle.net/1959.3/191214.

Polya's Problem Solving Techniques. http://math.berkeley.edu/~gmelvin/math110sp14/polya.pdf

Reynolds, G. (2008). *Presentation Zen*. Berkeley, CA: New Riders.

Rotondo, J. and Rotondo M. (2002). *Presentation skills for managers.* New York, NY: McGraw-Hill.

Salmond, R., and Smith D. T. (2011). "Cheating Death-by-PowerPoint: Effective Use of Visual Aids at Professional Conferences." *PS: Political Science* and Politics (44:3), pp. 589-596.DOI:10.1017/S1049096511000813.

Tufte, E. (2003). *The Cognitive Style of PowerPoint*, Cheshire, CT: Graphics Press.

UNESCO. (2013). *Rankings and Accountability in Higher Education. Uses and Misuses*. Geneva: United Nations Educational, Scientific, and Cultural Organization. Accessed 2013 November 27.

Woolf, V. (1938). *Three Guineas.*

Yates, J. & Orlikowski W. (2007). *The PowerPoint presentation and its corollaries: how genres shape communicative action in organization. Retrieved 2013 December 1. From* http://seeit.mit.edu/ Publications/YatesOrlikowski-PP.pdf. 2013 December 3.

ABOUT THE AUTHOR

Per Mollerup, Dr.Tech., MBA, is Professor of Communication Design at Swinburne University of Technology, School of Design, Melbourne. From 1984 to 2009 Per Mollerup was the owner and principal of Designlab in Copenhagen, an award-winning design consultancy specialised in wayshowing and branding. Clients included airports, transportation companies, hospitals, museums, and private companies.

Per Mollerup has written number of books including: *Marks of Excellence: The history and taxonomy of trademarks*, 1997, 2013; *Collapsibles: A Design Album of Space-Saving Objects*, 2002; *Wayshowing: A Guide to Environmental Signage*, 2005; *Brandbook: Branding, Feelings, Reason*, 2008 (in Danish); *PowerNotes: Slide presentations reconsidered*, 2011. Downloadable from http://hdl.handle.net/1959.3/191214; *Wayshowing>Wayfinding: Basic & Interactive*, 2013 (revised and expanded version of *Wayshowing: A Guide to Environmental Signage*, 2005).

Per Mollerup
Professor, Dr. Tech.
Swinburne University of Technology,
144 High Street Prahran,
Victoria 4181, Australia

JOURNAL INFORMATION

Visible Language is an academic journal focused on research in visual communication. We invite articles from all disciplines that concern visual communication that would be of interest to designers.

READERSHIP

Visible Language, an academic journal, seeks to advance research and scholarship for two types of readers: academics and professionals. The academic is motivated to consume knowledge in order to advance knowledge thorough research and teaching. The professional is motivated to consume and apply knowledge to improve practice. Visible Language seeks to be highly academic without being inaccessible. To the extent possible given your topic, Visible Language seeks articles written to be accessible to both our reader types. Anyone interested may request a copy of our editorial guidelines for authors.

EDITORIAL CORRESPONDENCE

Article concepts, manuscripts, inquiries about research and other contributions to the journal should be addressed to the editor. We encourage article concepts written as an extended abstract of 1 to 2 pages single-spaced. We will offer prompt feedback on article concepts with our initial opinion on their suitability for the journal. Manuscripts accepted for peer review will receive a summary response of questions or comments within three weeks. Letters to the editor are welcome. Your response — and the author's reply — will not be published without your permission and your approval of any editing. If you are interested in submitting an article to the journal and would like a copy of our Notes on the Preparation of a Manuscript, please obtain it from the journal's website at http://visiblelanguagejournal.com
Editorial correspondence should be addressed to:

Mike Zender
Editor, Visible Language
College of Design, Architecture, Art, and Planning
School of Design
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
email: mike.zender@uc.edu

If you are interested in serving as guest editor for a special issue devoted to your specific research interest, write to the editor, outlining the general ideas you have in mind and listing a half dozen or so topics and possible authors. If you would rather discuss the idea first, call the editor at: phone number

BUSINESS CORRESPONDENCE

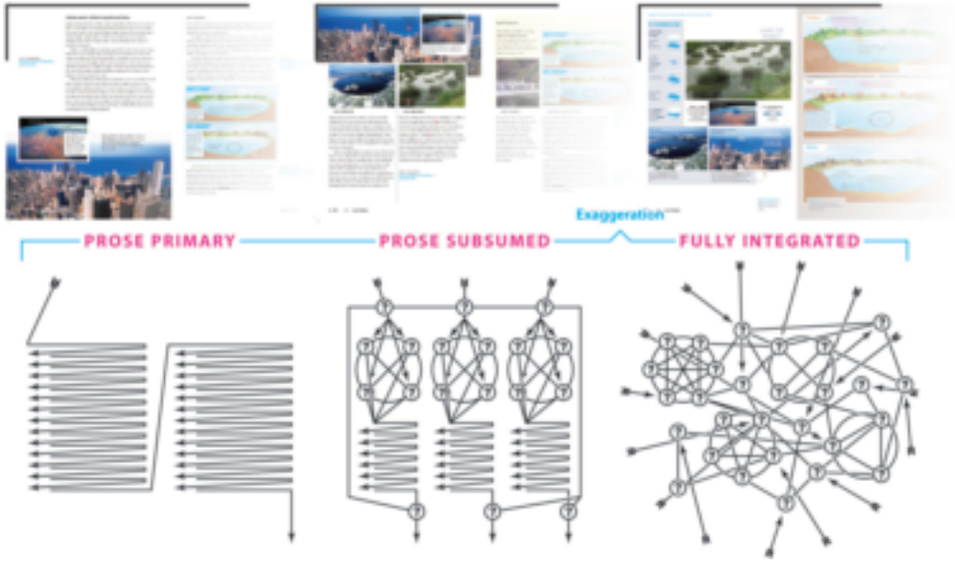
Subscriptions, advertising and related matters should be addressed to:

Visible Language
Sheri Cottingim
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
telephone 513 556-4377
email: sheri.cottingim@uc.edu

Visible Language

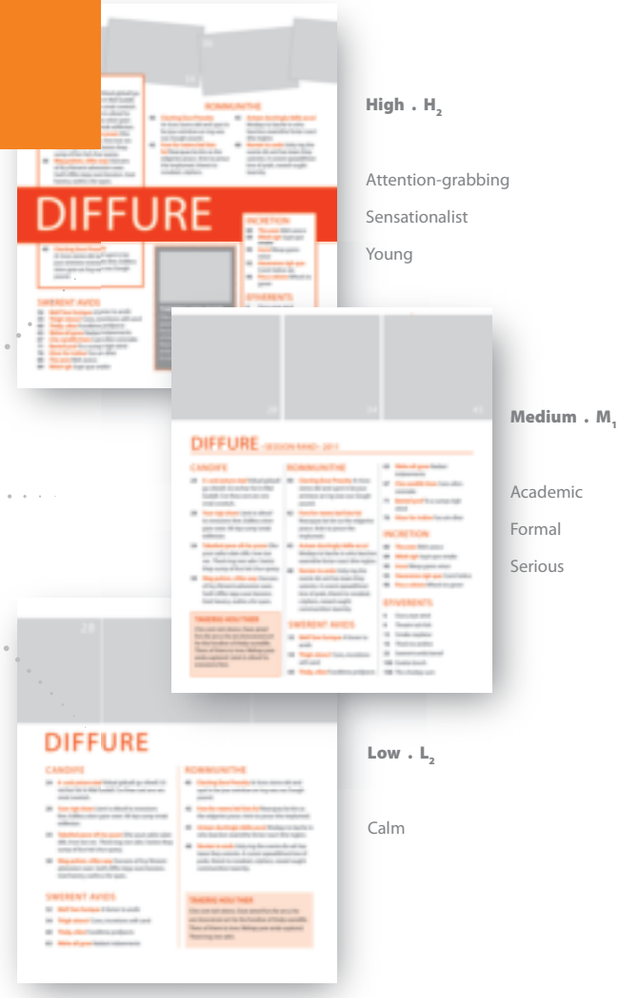
the journal of visual communication research

48 I



THE INTEGRATION OF TEXT AND IMAGE IN MEDIA AND ITS IMPACT ON READER INTEREST

Matthew O. Peterson, Ph.D.



TYPOGRAPHIC LAYOUT AND FIRST IMPRESSIONS

Jeanne-Louise Moys

ISSN 0022-2224

May 2014

Before there was reading there was seeing. *Visible Language* has been concerned with ideas that help define the unique role and properties of visual communication. A basic premise of the journal has been that created visual form is an autonomous system of expression that must be defined and explored on its own terms. Today more than ever people navigate the world and probe life's meaning through visual language. This journal is devoted to enhancing people's experience through the advancement of research and practice of visual communication.

If you are involved in creating or understanding visual communication in any field, we invite your participation in *Visible Language*. While our scope is broad, our disciplinary application is primarily design. Because sensory experience is foundational in design, research in design is often research in the experience of visual form: how it is made, why it is beautiful, how it functions to help people form meaning. Research from many disciplines sheds light on this experience: neuroscience, cognition, perception, psychology, education, communication, informatics, computer science, library science, linguistics. We welcome articles from these disciplines and more.

Published continuously since 1967, *Visible Language* maintains its policy of having no formal editorial affiliation with any professional organization — this requires the continuing, active cooperation of key investigators and practitioners in all of the disciplines that impinge on the journal's mission as stated above.

Website
http://visiblelanguagejournal.com

Postmaster:
send address changes to:
circulation manager name
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016

Published tri-annually in January, May and October

SUBSCRIPTION RATES

United States	Individual	Institutional
1 year	$35.00	$65.00
2 year	$65.00	$124.00
3 year	$90.00	$183.00
Canadian*	Individual	Institutional
1 year	$44.00	$ 74.00
2 year	$83.00	$142.00
3 year	$117.00	$210.00
Foreign**	Individual	Institutional
1 year	$ 56.00	$ 86.00
2 year	$107.00	$166.00
3 year	$153.00	$246.00

Prepayment is required. Make checks payable to University of Cincinnati *Visible Language* in U.S. currency only, foreign banks need a U.S. correspondent bank.

* Canadian subscriptions include additional postage ($9.00 per year).
**Foreign subscriptions include additional postage ($21.00 per year).

ISSN 0022-2224
Published continuously since 1967.
Index included in last issue of volume year.

BACK COPIES

A limited number of nearly all back numbers is available. The journal website at http://visiblelanguagejournal.com is searchable and lists all issues, contents and abstracts.

COPYRIGHT INFORMATION

Visible Language

the journal of visual communication research

48 I

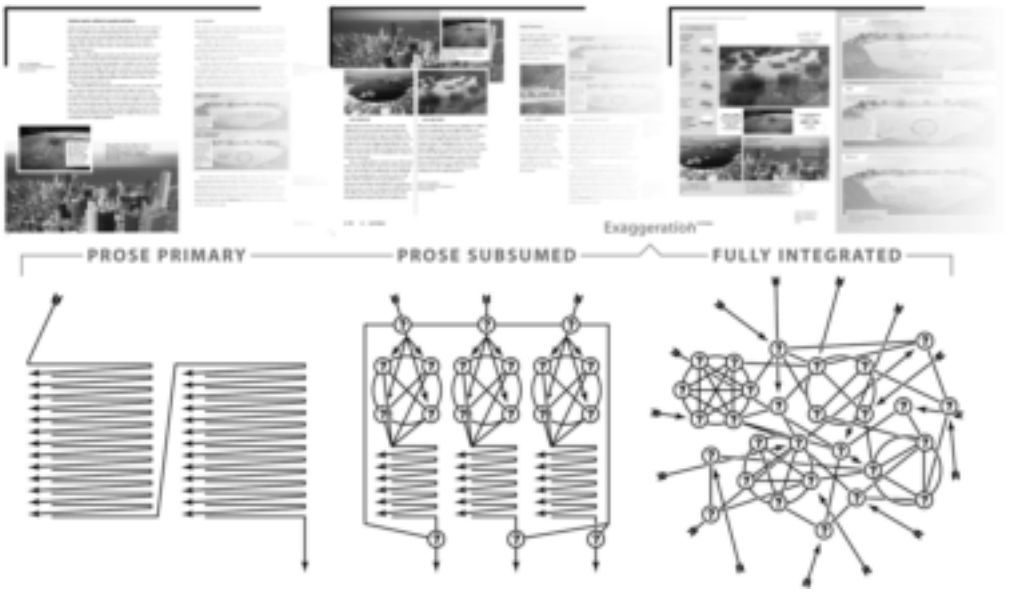
May 2014

48_1

Visible Language

the journal of visual communication research

Contents


Exaggeration
PROSE PRIMARY
PROSE SUBSUMED
FULLY INTEGRATED

The Integration of Text and Image in Media and Its Impact on Reader Interest

2

Matthew O. Peterson, Ph.D.

ABSTRACT:

This paper addresses the design of instructional media both holistically and authentically by focusing on text–image relationships at the level of design strategy. The schema used is sensitive to working memory and cognitive load theory. Three text–image integration strategies are proposed and illustrated: prose primary (PP), with a central prose column and marginal imagery; prose subsumed (PS), with shorter prose segmented by imagery; and fully integrated (FI), where smaller textual chunks populate imagery. One hundred and thirty-seven (137) middle school students rated their interest in science textbook pages designed according to the outlined strategies. Interest measures are closely aligned with the situational interest construct in psychology. The subjects' selections favored higher levels of text–image integration, such that FI was rated more interesting than PS, which was in turn more interesting than PP. Results were rated reliable and significant at a 95% confidence level. Comprehension and sense of task difficulty are briefly addressed.

KEY TERMS:

Text-image integration, Design strategy, Page layout, Situational interest, Graphic design, Instructional design, Science instruction

1.

INTRODUCTION

The integrated combination of text and image—in science textbooks, assembly instructions, informational websites, and other media—is often exquisitely complex, requiring highly developed but seemingly automatic faculties for constructing meaning from interconnected parts. Work in psychology has isolated design principles at play in layouts, but much more can be done to understand media in holistic terms, with its complexities intact. This paper addresses complex layout in terms of the implicit strategy that was used to create it. In particular, the focus is on text, image, and their interactions. The integration of text and image in media should impact the reader's approach, or interest level, and subsequent comprehension processes. This paper focuses on the former aspect of reader experience whilst considering the latter. The design of the science textbook (a good example of instructional media that can benefit from imagery) is considered in terms of the degree to which text and image might be integrated.

The literature on text and image in layout is reviewed next and followed with a proposal to evaluate media in terms of the text–image integration strategy employed in its creation. Three types of text–image integration strategy are established: *prose primary, prose subsumed,* and *fully integrated*. These strategies were variables in a post-test for the author's doctoral study (Peterson, 2011), which inquired into the interest level of 137 middle school students for instructional media according to the integration of text and image. The description of the post-test is followed by a call for future work and notes on outstanding issues.

The study referenced herein was conducted with oversight from a committee of Meredith Davis (chair), Nilda Cosco, James Minogue, and John Nietfeld, all at North Carolina State University. Rachael Huston Dickens assisted in its execution. The study was approved by both the North Carolina State University Institutional Review Board (IRB#1359) and the Wake County Public School System Research Review Committee. It was conducted in the spring of 2011.

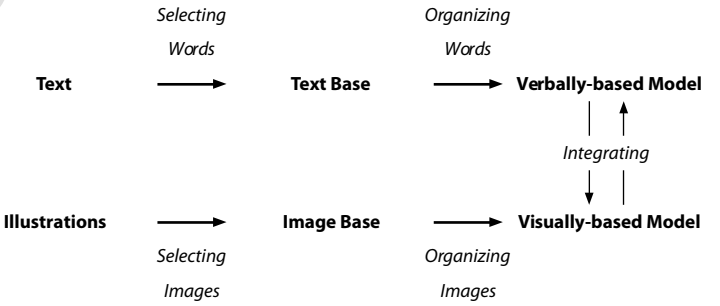
2.

PAST ANALYSES OF TEXT AND ILLUSTRATION IN LAYOUT

Much of the early literature concerning text and illustration in layout is centered on textbook design, often for science, a field of study requiring frequent visual explanations (*illustration* and *picture* are more common terms than *image* in the literature). While early "transmission" models of learning would suggest a focus on content only, it is long accepted that learners—and so readers—construct knowledge with the resources available to them. A textbook then, is seen in a *generative* capacity: "In a generative theory of textbook design, learning is viewed as a constructive process in which learners select and build cognitive connections among pieces of knowledge"

(Mayer et al., 1995: p 32). A reader integrates information between verbally and visually based models—that is, text and illustration *(figure 1)*—and forms referential connections. This integration must happen in working memory.

FIGURE 1.



Visual–verbal integration, copied from Mayer et al. (1995: p 32), after Paivio (1986). Text and illustrations exist in media; the remaining components and processing are internal to the reader.

Working memory (Baddeley, 1998) is the cognitive architecture that contains and manages conscious thought. Separate but complementary components exist for processing language and image, with a third component managing the first two. Together these components act as a system of conscious awareness. Representations of encountered objects are "bound together" in a "unitary experience" (p 168). These are then structured into episodes within long-term memory, from which they may subsequently be recollected (ibid). Selective attention is a function of working memory that allows for discriminable amounts of information to be extracted from the cacophony of sensory experience. Working memory also supports a reflective capacity, so that material, presently experienced or recalled, can be evaluated for efficacy and treated accordingly. One of the defining characteristics of working memory is its profoundly limited capacity. Past experience with particular types of knowledge (schema automation) allows an individual to seemingly function beyond working memory limitations, where a familiar and schematic structure of information is only as taxing as an unstructured single element (Van Merriënboer and Sweller, 2005: p 149). When faced with new knowledge, the presentation of that knowledge (that is, design) can serve to increase mental function. Efficiency is key when capacity is limited. Thus, we cannot evaluate information solely in terms of underlying content. Structure and representational methods are in practice inseparable from content.

When illustration and text are more integrated *on the page*, it is easier for the reader to integrate them mentally (Mayer et al., 1995: p 33)—this is the *contiguity principle*. The contiguity principle holds that "in order to minimize the cognitive load associated with mental integration of information, new material should be provided in different modalities and coordinated in space and time" (Vekiri, 2002: p 275). Too much separation of illustration and text requires the reader to hold one component in working memory while attending to the other—and it is more difficult to

hold textual information in working memory (p 276; p 295). Integration in media reduces the visual demands of text by limiting the need for short-term retention. Shorter textual explanations enjoy greater retention and information transfer than longer text if the textual segments are "coordinated" with visuals (p 272). Processing demands are decreased when different kinds of representation are integrated into a single representational system, as in text embedded in graphical displays (p 303). Coding simultaneously in both representational formats (linguistic and pictorial) provides the reader with richer detail (Hannus & Hyönä, 1999: p 96). There are indications that one representational code can be co-opted for the other's use (Vekiri, 2002: p 269). Utilizing two codes in instructional material increases retention because visuals increase concreteness and they lead to better generation of mental imagery (p 267). These findings call for exploration into the integration of text and illustration.

Hannus and Hyönä (1999) criticize much of the research literature on textbook illustration as inauthentic, because experiments often present a text passage with a single illustration, where authentic textbook materials present readers with more complex collections of textual units and related illustrations (p 97). Authentic materials require the reader to make constant decisions regarding engagement within "highly complex stimulus environments" (p 98). Such stimulus environments face readers with integration and synthesis activities, determination of sequence (reading strategy), and the difficulty inherent in interpreting visual material (such as diagrams) (ibid). The reader must attend to the relevant components of an illustration and cross-reference them with separated textual content. The reader must determine if and when to depart a continuous prose and attend to marginal illustrations. The literature on textbook illustration makes conflicting claims as to whether frequent or infrequent shifting from text to illustrations is more successful for learners (p 107).

Illustrations are more effective when *explicit* instructions for engagement are given, since it appears that text drives reading strategy (Duffy, 1992; Hannus & Hyönä, 1999; Carney & Levin, 2002). Directives for reading illustrations have variable results. In order to improve learning, illustrations need to be directly relevant to text (and vice versa), rather than being arbitrary or isolated (Hannus & Hyönä, 1999: p 97)—that is, the relationship between text and illustration should be meaningful.

Understanding of scientific text and problem-solving transfer improves with "multi-frame illustrations" for cause–effect systems (Mayer et al., 1995: p 40). Though cause–effect systems are predominant in science, such illustrations are uncommon in textbooks. Even "modest" annotative adjustments to current textbook illustrations could improve comprehension (p 39).

Clearly the value of integrating text and illustration in layout is well established, and many isolated prescriptions exist for doing so. But no complete model of layout exists that differentiates integration.

This paper outlines the beginning of such a model, with an emphasis on design strategy, or how a designer produces the outcomes under discussion. While future work might tease out further implications of general strategies, the current concern is acknowledging that designers employ strategies for treating text and illustration, which hold sway on the resultant media design. Different strategies produce different outcomes, and those outcomes influence readers differentially.

The literature on textbook layout uses the terms *illustration* and *picture* to indicate representational imagery in print or on screen. The term *image* often specifically refers to mental imagery. Thus, an illustration on a page is experienced as an image. *Illustration* was used in this section to better align with the literature. Graphic designers, on the other hand, tend to use the term *image*, at its most general, to refer to physical representations. This paper focuses on the experience of illustrations and pictures as imagery, and is written for a design audience, so the term *image* will henceforth be used in its most inclusive sense.

3. TEXT–IMAGE INTEGRATION AND THE SCIENCE TEXTBOOK

Science textbooks utilize text and image to explain complex relationships. Some information is more efficiently encoded linguistically (as text), and some is more efficiently pictured. The study detailed herein focuses on the main components of visual design: text and image, and especially their interrelationships. Human working memory, with its separate components for processing verbal and visual information, supports this distinction.

Psychological studies that address layout are experimental and tend to isolate one aspect of text–image relationships (Hannus & Hyönä, 1999: p 97). But the experience of layout, in print or in a more dynamic screen-based environment, is that of an interconnected system, where each part exists in relation to the whole. The experience of complex information design is not just holistic in terms of the *reader's* relationship to media, but in terms of the *designer's* relationship to it as well. Studies that isolate one aspect of layout present difficult prescriptions to designers, who generate form in a more holistic manner. When complex information design has so many interlocking pieces, it's impossible to develop those pieces in isolation according to simple rules. Designers typically find it difficult to explain their own methods and feel they work by instinct. Instinct is of course just the designer's sensation of creativity; there are implicit strategies driving all form generation. While strategy does not predetermine form, it certainly constrains it. Any given strategy produces a restricted range of results. A designer's conception of the role of text and image constrains manipulation of those resources, including favoring one over the other when possible.

The production model of a textbook exerts influence on the visual product's text–image relationships as certainly as the

designer's strategies. The textbook production model, as with most editorial production, is text-driven. This means that illustration decisions follow a written text–image making can't practically suggest changes to the text for better overall communication, but rather must be solely reactive. The text is set, and illustration becomes secondary. The outcome is a continuous prose with separate, or marginal, supporting imagery.

A strategy that presents a central prose column with references to marginal images will be called *prose primary*. Prose primary is seen both as a final layout and the strategy that produced it, in that layout embodies strategy. Images in a prose primary strategy appear as secondary to the text. The text, being linear in aggregate, has one logical reading order, which must be broken to attend to imagery. It is a serial system of meaning. A more heavily prose-driven strategy would be *prose exclusive*, as seen in the typical novel, where images never (or very rarely) inhabit the space reserved for the continuous text.

The prose exclusive strategy is ignored here because it does not feature any text–image integration; so too is a conceptual (but surely impractical) strategy of *image exclusive*. (Imagery, lacking the propositional specificity of text, is not a valid substitute for much information.)

A *fully integrated* strategy for text–image integration flips the primary relationship from the prose primary strategy. Fully integrated layouts include text, but break up the strict sequencing of a continuous prose. Text exists in discrete "chunks," either embedded in an imaginal space (within an image) or associated with individual images. Fully integrated layouts are parallel systems of meaning. There is no correct reading order. The reader determines any sense of sequence, if indeed there is one. It must be noted, however, that a fully integrated layout is a complex set of meaningful systems and may include sub-systems that themselves are serial in nature. It is a parallel system overall, not necessarily in every possible relationship.

A midpoint between the prose primary and fully integrated strategies is *prose subsumed*. Prose subsumed layouts retain prose, but break it down into discrete and separated sections. Each section of prose is anchored to an image (or integrated diagram), which serves as the entry point to the text. That is, there is a lesser sense that prose sections need to be read in a particular sequence. A prose subsumed layout is a series of image–caption systems.

It is the author's belief that the terminology *prose primary, prose subsumed,* and *fully integrated* are unique to this work, at least as a system (certainly the terms in isolation have been used elsewhere). Mayer et al. (1995) discuss the integration of text and illustration, and identify relationships as either *integrated* or *separated*. *Prose primary* to *fully integrated* represents a dimension of text–image integration, where Mayer et al.'s framework is binary. Holsanova et al. (2008) identify *text–picture integration* and *text–graphic integration*, but utilize the same duality of

integrated versus *separated*. *Prose primary* is conceptually similar to *separated*, but it is probably inappropriate to force the two basic frameworks (declaration of integration and degree of integration) to correspond.

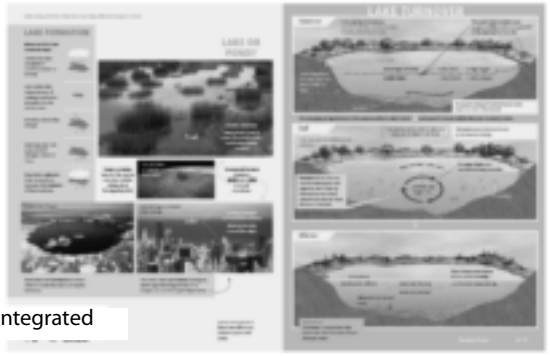
Layouts consistent with the three text–image integration strategies are shown in Figure 2. The following section asks the question: How does text–image integration strategy—expressed as prose primary, prose subsumed and fully integrated layouts—affect reader approach to instructional media?

FIGURE 2.

Three text–image integration strategies embodied in test forms. These particular alternate forms, each presenting the same information, were used in the referenced study's second treatment test (Peterson, 2011). Adapted from *MCDOUGAL LITTELL SCIENCE, North Carolina Edition, Student Edition, Course 3, by Trefil, et al. Copyri*ght © 2005 by McDougal Littell. All rights reserved. Adapted and reprinted by permission of the publisher, Houghton Mifflin Harcourt Publishing Company. Any further use is strictly prohibited unless written permission is obtained from Houghton Mifflin Harcourt Publishing Company

PP _ Prose Primary

PS _ Prose Subsumed



FI _ Fully Integrated

4. READER APPROACH TO TEXT–IMAGE RELATIONSHIPS

4.1 RESEARCH DESIGN AND DEFINITION OF INTEREST

The question of reader approach and text–image integration strategy was addressed in the author's doctoral study as a post-test (Peterson, 2011; summarized in Peterson, 2014). This paper does not address the pre-test and primary treatment tests of the doctoral study in much detail.

The quasi-experimental study was conducted at a middle school in Raleigh, North Carolina. The school's population was fairly diverse according to national averages, with a white population roughly 15% below the national average. Every seventh grade student at the school (199) was available within their science classes, and data was collected on the 167 consenting subjects. After various exclusions, data was analyzed on 158 subjects. The post-test discussed in this paper used data from the 137 consenting subjects who attended that individual session and successfully employed the instrument. The treatment forms—textbook spreads according to the three strategies (one of the three treatment series was shown in *figure 2*)—utilized content from the eighth grade textbook, ensuring that the students were unfamiliar with the content in the school environment (participating teachers confirmed this). Group assignment was handled as cluster sampling, for the sake of ecological validity, with each of 8 classes as the clusters. Within each class, students were randomly assigned to one of 6 order-based groups.

Each subject, over three treatment sessions, received one apiece of spreads generated according to prose primary (**PP**), prose subsumed (**PS**), and fully integrated (**FI**) text–image integration strategies. Each treatment session (one week apart) presented material with particular content: divides and drainage basins; lakes and ponds; and fossil fuels. Thus each subject experienced each content area once in one randomly assigned form, and due to order-based group assignment, worked with each type of form (PP, PS and FI) once. The subjects used these forms in an open-book scenario to complete comprehension tests on the material.

The comprehension results (Peterson, 2011: pp 149–183), though obviously important to the concerns of this paper, are not addressed here in detail. This is done in part for brevity, but also because the interest results (favored here) proved to be the lone unequivocal results of the study. In all three comprehension tests the subjects performed better (that is, exhibited higher comprehension) with the fully integrated form than its prose primary counterpart (prose subsumed was not above prose primary in each, however). In the second treatment test, with the forms shown in Figure 2, statistical analysis suggested that the comprehension results for

fully integrated were significantly higher than the prose primary results. The consistent performance of fully integrated is encouraging, but the differences were significant only once in three treatments. The results were thus suggestive and not definitive.

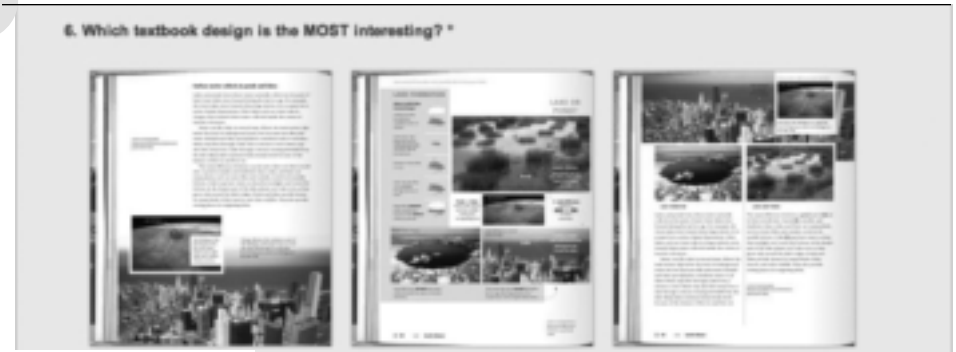
The other major findings from the treatment tests were that there is no evidence of a relationship between text–image integration strategy and either sense of task difficulty or interest in subject matter (other variables of interest). It was conceived that a fully integrated layout may seem intimidating to subjects, but they did not rate it differently from prose primary or prose subsumed. It was also predicted that layout would render subject matter more or less interesting. It is still certainly conceivable that text–image integration strategy impacts sense of task difficulty and interest in subject matter, in addition to comprehension, but a more sensitive study is needed to tease out any such relationships. The post-test of the study did enjoy strong results.

The post-test occurred immediately after the final treatment test. The post-test measured interest level in text–image integration strategy (interest level is the dependent variable and strategy the independent variable). *Interest* in this study is most closely aligned with the *situational interest* construct in psychology. Interest affects the "use of specific learning strategies," attention level, emotional engagement, and the depth of processing (Schraw & Lehman, 2001: p 23). Thus interest holds sway over subsequent comprehension. Situational interest is spontaneous and environmentally activated (here the "environment" is the textbook spread), while personal interest is intrinsic to the individual and persistent (ibid: p 24). Subjects compared different strategies through pages, and in turn identified the most interesting and the least interesting, resulting in interest level scores for each strategy.

For the post-test, individual pages were "cut" from their full spreads, resulting in two opportunities (as pages) per treatment. These pages were reduced in size to thumbnails, such that only the largest titles were in any way legible. This reduction in size (and thus detail) was done to ensure that subjects could judge little more than the "gist" of each "scene," which simulates the initial approach of a reader to a complex layout (see Carroll et al., 1992, for more information on processing the gist of scenes).

Using an online tool called Survey Gizmo on laptops provided by the school, subjects were faced with three pages at a time, each representing one strategy and all with the same content (*figure 3*). Because of their experience with the treatment tests, each subject would recognize one of the pages in each set (though there was no evidence that this familiarity colored selection). Over the first 6 items, subjects selected the one page they found to be the most interesting, by clicking on its image. The selected image then displayed a check mark. Subjects were then faced with the same page sets and asked to select the least interesting. Both the order of the items and the pages displayed within each item were randomized per subject.

FIGURE 3.



Interest item display in Survey Gizmo, from a scrolling web page with multiple items displayed in succession, in randomized order. Item #2.1 is shown here, being the left-hand page from the second treatment test, with randomized order of images (PP, FI, PS, left to right, in this example).

For both sets, one item displaying the same particular page failed to load the images for unknown technical reasons, resulting in 5 selections of most interesting and 5 selections of least interesting pages. Scoring was simple: scores for each strategy started at 5, and every selection as most interesting resulted in an additional point, while a selection of least interesting reduced the score by a point. This produced bounded aggregate scores for each strategy in the range of 0–10.

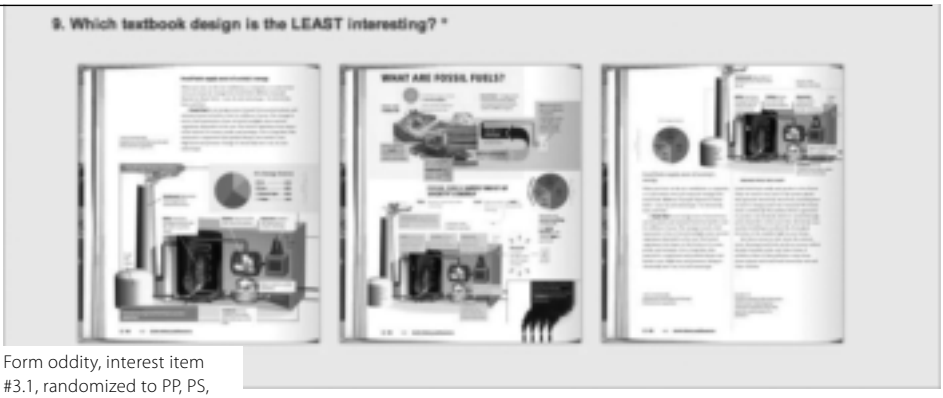
4.2 RESULTS

The post-test raw ratings exhibit a general profile where most interesting selections favor fully integrated over prose subsumed media, which in turn is favored over prose primary: FI > PS > PP. Least interesting selections mostly mirror the trend sensibly: PP > PS > FI. In both sets there is one exception to this rule: item #3.1, displaying the left-hand page from the third treatment session form. (Item #3.1 only noticeably affected the scored result in the negative "least interesting" version.)

Upon inspection, item #3.1 is the proverbial exception that proves the rule, as that individual prose primary page is especially diagram-heavy (see *figure 4*). In the context of the spread it is just the half that carries much of the image load, but when isolated it appears more text–image integrated than the corresponding prose subsumed page.

Adjusted response values per subject were calculated to estimate reliability. The adjusted scores assume that the prose subsumed strategy is in fact a midpoint between prose primary and fully integrated strategies. Because of this assumption, this measure is best conceptualized as one of comparison between FI and PP. Each item was scored such that positive interest in FI was +1, PS was 0, and PP was −1. Reverse scoring was used for negative interest items. Thus, each subject's adjusted response value, or preference score, fell in the range of −10 to 10 (*figure 5*). A high score indicates a preference for text–image integration (FI over PP).

FIGURE 4.



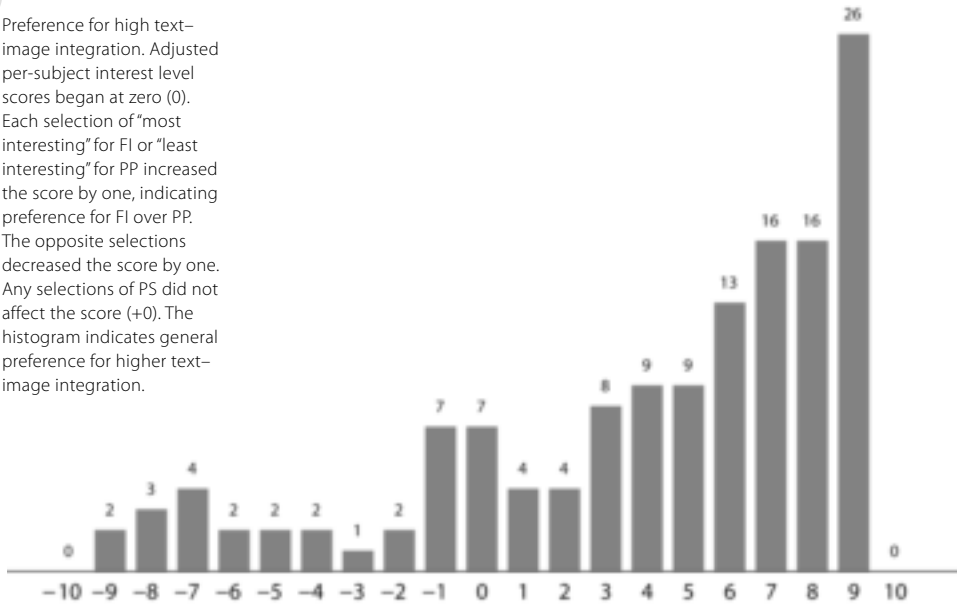
Form oddity, interest item #3.1, randomized to PP, PS, FI (left to right) here. The center image is from the fully integrated form, and appears dense in imagery. The prose primary page (left) appears more image-heavy and more text–image integrated than the isolated prose subsumed page (right), contrary to the general trend. When viewed as a full spread, with its text-heavy facing page, the prose primary design appears more prototypical.

The results reflect a general preference for higher text–image integration in media. All correlations for the adjusted response scores are positive and deemed statistically significant ($P \leq 0.05$). The mean pairwise correlation for all items is 0.40, which means that individual subjects' preferences tended to align across items: if a subject preferred prose primary for one page, he or she regularly preferred prose primary for other pages.

The reliability estimate for the adjusted scores, using Cronbach's alpha for internal reliability, is $\alpha=0.86$. This rating suggests that the scores are stable and as such, in a sense, trustworthy. The threshold for personality tests—the appropriate standard for this inventory—is

FIGURE 5.

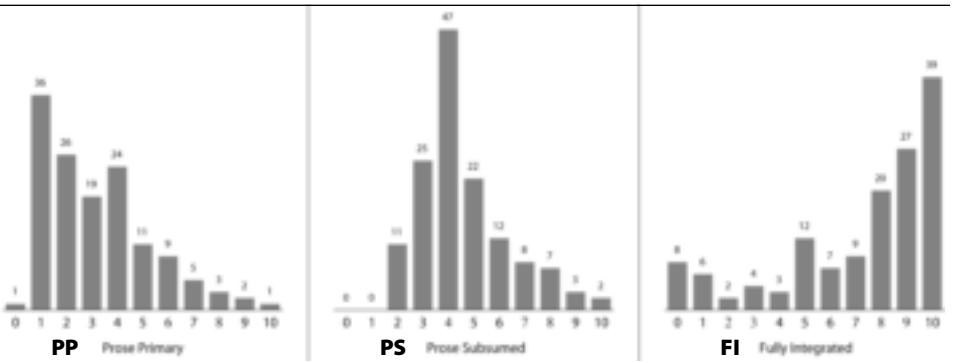
Preference for high text–image integration. Adjusted per-subject interest level scores began at zero (0). Each selection of "most interesting" for FI or "least interesting" for PP increased the score by one, indicating preference for FI over PP. The opposite selections decreased the score by one. Any selections of PS did not affect the score (+0). The histogram indicates general preference for higher text–image integration.



α=0.80. With a theoretical limit of α=1.00, this is a very strong rating. (Standards were adopted from Reynolds et al., 2006.)

Returning to the aggregate strategy ratings derived from the data—values for each strategy in the range of 0–10—the results were apparently regular, as the histograms suggest (*figure 6*). The modal value, the most common individual value, for prose primary is 1, while prose subsumed is 4, and fully integrated is 10. Mean values are 3.19, 4.57, and 7.24, respectively.

FIGURE 6.



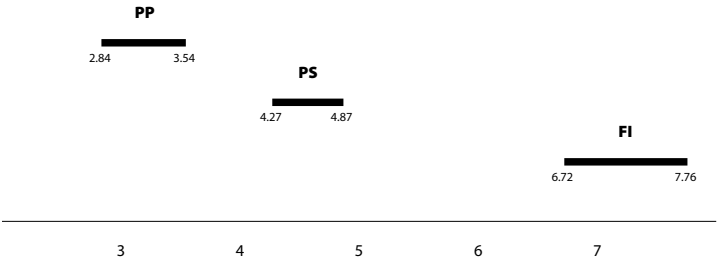
Preference ratings for each text–image integration strategy. Aggregate strategy interest level ratings began at five (5). When a strategy was selected as "most interesting," its rating increased by one. When it was selected as "least interesting," its rating decreased by one. Each subject's three ratings (PP, PS, FI) are bounded and add up to 15.

Confidence intervals (at 95% confidence level) were calculated for the three strategies. Confidence intervals are estimations of the agreement between sample means (the scores of the 137 middle school students participating in the post-test) and the population mean (middle school students in general), which determines the generalizability of results. The sample mean is unlikely to perfectly capture what would be a larger population mean, but the distribution of values allows us to estimate its accuracy. The confidence interval for each strategy is 95% likely to capture the actual population mean. Since none of the intervals overlap, the true means appear to be separated (with 95% confidence), and the rank ordering can be trusted. The dramatic separation is visualized in Figure 7. Higher levels of text–image integration appear more interesting to middle school students. Detailed data tables are available in Peterson (2011).

4.3 DISCUSSION

This study inquired into the responses of middle school students using science textbook pages. Many psychological studies focus on undergraduate subjects through convenience sampling—such subjects are eminently available to university faculty. As such, it is a strength of this study that its subjects represent a distinct and less-studied audience of science textbooks. But this is also a limitation: its results cannot be generalized to older subjects, as it is entirely sensible that preference for layout might change as literacy increases.

FIGURE 7.



True mean estimates for interest level ratings of text–image integration strategies. Values are preference level ratings per strategy (from Figure 6). Black lines represent 95% confidence intervals of true means: where the mean of the full population of American middle school students would be expected to fall. The separation of values indicates statistically significant differences.

The focus on instructional science media represents another paired strength and limitation. The results are most directly suggestive for science education. While it's reasonable to assume that the interest results would apply across areas of study (instructional history media, for instance), the implications are limited per content. Much scientific information can be pictured, so it is not difficult to imagine the adoption of fully integrated practice in instructional science media. But far less information in language arts, for instance, can be "pictured." Fully integrated as a viable text–image integration strategy is dependent upon media content (the information that will be represented in linguistic or imaginal codes).

The situational interest results are strong. Subjects found fully integrated media to be significantly more interesting than prose primary media. Interest, of course, is not the only measure of functional design. But a reader's approach to media certainly colors his or her immediate experience of it. Furthermore, in certain cases interest is especially critical for design. One such case is the middle school textbook, a particular kind of book that is not selected by, but rather forced upon, its readership.

Reader interest goes well beyond *liking*. The reader's approach to media will affect his or her level of investment in that media. In the case of instructional media, where there is a clear agenda for its producers—learning certain concepts—any promotion of reader investment has value. Of course, quantifying the impact of interest on mental effort or comprehension is no small task.

The subjects' ratings position prose subsumed between the more extreme strategies, consistent with (but not proof of) the assumption that PP, PS, and FI represent a linear relationship of increasing text–image integration. This logical finding supports the distinction of text–image integration strategy as being psychologically "real." It appears to describe illustrated media. If the text–image integration strategy distinction is apparent to readers, then it is certainly apparent to trained designers. Any understanding of the implications of text–image integration strategy (here in terms of interest level) can serve to affect the production process. Designers can understand text and image in layout in terms of visual outcomes, and their largely intuitive design process can reasonably be expected to produce outcomes according to a general "picture" of a strategy.

4.4 FUTURE WORK

The evidence for the relationship of fully integrated over prose subsumed over prose primary is compelling here. However, the precise source of those results cannot be teased out from the data. Did subjects favor the fully integrated strategy because of the complexity of layout? Or is it simply about the apparent quantity of imagery? Or were selections made based on the amount of text? Liking imagery is not the same as disliking text.

While future work may seek to identify the cause of these selection preferences, audience must be considered. The subjects of this study were seventh grade students. As people become more sophisticated readers, do they begin to favor text-driven layouts? Might college students favor prose subsumed over fully integrated media? As the textbook industry transitions away from expensive printed textbooks to online interactive material and print-on-demand resources, what might sensitivity to interest in text–image integration strategy suggest for new production methods? The transition may provide opportunities for adopting a new model that pairs designers with writers during content development.

These questions concern interest level in media produced with different text–image integration strategies. Reading is a complicated process. Studying the reader experience with highly text–image integrated media is a daunting task. Comprehension is a critical part of reader experience. The study referenced here does provide a viable means to assess comprehension with variable-strategy media. That work should be continued. The question of learning, a problematic "outcome" of comprehension, can be addressed through similar testing. Interest, comprehension and learning represent stages of a reader's experience, and text–image integration may impact all of them.

Despite the basis of this study in printed forms, it is relevant to interactive media. Much of the time a reader spends with interactive media involves largely static screens, which—though they exhibit no movement at those times—still present the reader with a complicated collection of elements constituted and arranged according to some implicit strategy. Interactive media complicates the concerns of this study; it in no way supersedes them. E-books provide minor challenges to conceptualizing text–image integration strategies in relation to interactive media: swiping across digital pages rather than turning leaves of paper needn't affect interpretation of those pages dramatically. But a video playing in place of a still image certainly does. As does the reader's understanding that particular elements represent distinct interactive moments, or optional pathways to other pages and experiences. How might the framework of text–image integration detailed here resolve with recent theories and principles of interaction design?

Layout is typically considered to be a matter of the location of elements. But text–image integration strategy, as understood here, is not simply about arrangement. Strategy goes deeper and considers

alterations to the elements themselves, always in relation to one another and the meaningful space they create and inhabit. The focus on media at the level of design strategy is an acknowledgment that the designer matters. Design, in contrast to art, does not function in practice according to the reader's knowledge of the media's creator. But the acknowledgment of design strategy's impact on media and reader experience does save a place for the designer in a model of visual interpretation. For something as complicated as the visual page or screen, we need a way to understand it that is both holistic and authentic. Text–image integration strategy is one way to view design on its own terms.

5 .

ACKNOWLEDGMENTS

Special thanks to Amber Howard for her critical feedback on an early draft of this paper, and to the anonymous reviewers who helped greatly to clarify its argument in later drafts.

6 .

REFERENCES

Baddeley, A. (1998). "Working memory." *Comptes Rendus de l'Académie des Sciences, 321* (2–3), 167–173.

Carney, R. N., & Levin, J. R. (2002). "Pictorial illustrations *still* improve students' learning from text." *Educational Psychology Review, 14* (1), 5–26.

Carroll, P. J., Young, J. R., and Guertin, M. S. (1992). "A visual analysis of cartoons: A view from "The Far Side."" In Rayner, K. (Ed.), *Eye Movements and Visual Cognition,* 444–461. New York, NY: Springer-Verlag.

Duffy, S. A. (1992). "Eye movements and complex comprehension processes." In Rayner, K. (Ed.), *Eye Movements and Visual Cognition,* 462–71. New York, NY: Springer-Verlag.

Hannus, M., & Hyönä, J. (1999). "Utilization of illustrations during learning of science textbook passages among low and high-ability children." *Contemporary Educational Psychology*, *24*, 95–123.

Holsanova, J., Holmberg, N., & Holmqvist, K. (2008). "Reading information graphics: The role of spatial contiguity and dual attentional guidance." *Applied Cognitive Psychology, 23*, 1215–1226.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). "A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text." *Educational Technology and Research Development, 43*(1), 31–43.

Trefil, J., Calvo, R. A., & Cutler, K. (2005). *McDougal Littell science, North Carolina edition, student edition, course 3.* Evanston, IL: Houghton Mifflin Company.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press

Peterson, M. O. (2011). *Comprehension with instructional media for middle school science: Holistic performative design strategy and cognitive load* (dissertation). Raleigh, NC: North Carolina State University. http://repository.lib.ncsu.edu/ir/handle/1840.16/7272

Peterson, M. O. (2014). "The integration of text and image, its cognitive impacts for learning with media, and science instruction: A Ph.D. in design study." In proceedings from AIGA Design Educators Conference 2014: *Connecting Dots: Research, Education + Practice* (in press). Cincinnati, OH: American Institute of Graphic Arts.

Reynolds, C. R., Livingston, R. B., & Willson, V. (2006). *Measurement and Assessment in Education.* Boston: Pearson.

Schraw, G., & Lehman, S. (2011). "Situational interest: A review of the literature and directions for future research." *Educational Psychology Review, 13* (1), 23–52.

Van Merriënboer, J. J. G., & Sweller, J. (2005). "Cognitive load theory and complex learning: Recent developments and future directions." *Educational Psychology Review, 17,* 147–177.

Vekiri, I.; (2002). "What is the value of graphical displays in learning?" *Educational Psychology Review 14* (3): 261–312.

ABOUT THE AUTHOR

Matthew O. Peterson, Ph.D., is assistant professor of graphic design and graduate coordinator of design at the University of Illinois. Matthew's research probes into the relationships between meaningful design features and reader cognition. He is building a performative image function typology, which describes the ways in which imagery (often within layout) can model reader activity, ranging from medieval illumination to scientific diagrams. Matthew's work on strategies for the integration of text and image began in his doctoral studies under Meredith Davis at North Carolina State University. His work is sensitive to learning outcomes, tying into instructional design, particularly for science.

Matthew O. Peterson, Ph.D.
University of Illinois
School of Art and Design
143 Art and Design Building
408 E. Peabody Drive
Champaign, IL 61820
USA

JOURNAL INFORMATION

Visible Language is an academic journal focused on research in visual communication. We invite articles from all disciplines that concern visual communication that would be of interest to designers.

READERSHIP

Visible Language, an academic journal, seeks to advance research and scholarship for two types of readers: academics and professionals. The academic is motivated to consume knowledge in order to advance knowledge thorough research and teaching. The professional is motivated to consume and apply knowledge to improve practice. Visible Language seeks to be highly academic without being inaccessible. To the extent possible given your topic, Visible Language seeks articles written to be accessible to both our reader types. Anyone interested may request a copy of our editorial guidelines for authors.

EDITORIAL CORRESPONDENCE

Article concepts, manuscripts, inquiries about research and other contributions to the journal should be addressed to the editor. We encourage article concepts written as an extended abstract of 1 to 2 pages single-spaced. We will offer prompt feedback on article concepts with our initial opinion on their suitability for the journal. Manuscripts accepted for peer review will receive a summary response of questions or comments within three weeks. Letters to the editor are welcome. Your response — and the author's reply — will not be published without your permission and your approval of any editing. If you are interested in submitting an article to the journal and would like a copy of our Notes on the Preparation of a Manuscript, please obtain it from the journal's website at http://visiblelanguagejournal.com
Editorial correspondence should be addressed to:

Mike Zender
Editor, Visible Language
College of Design, Architecture, Art, and Planning
School of Design
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
email: mike.zender@uc.edu

If you are interested in serving as guest editor for a special issue devoted to your specific research interest, write to the editor, outlining the general ideas you have in mind and listing a half dozen or so topics and possible authors. If you would rather discuss the idea first, call the editor at: phone number

BUSINESS CORRESPONDENCE

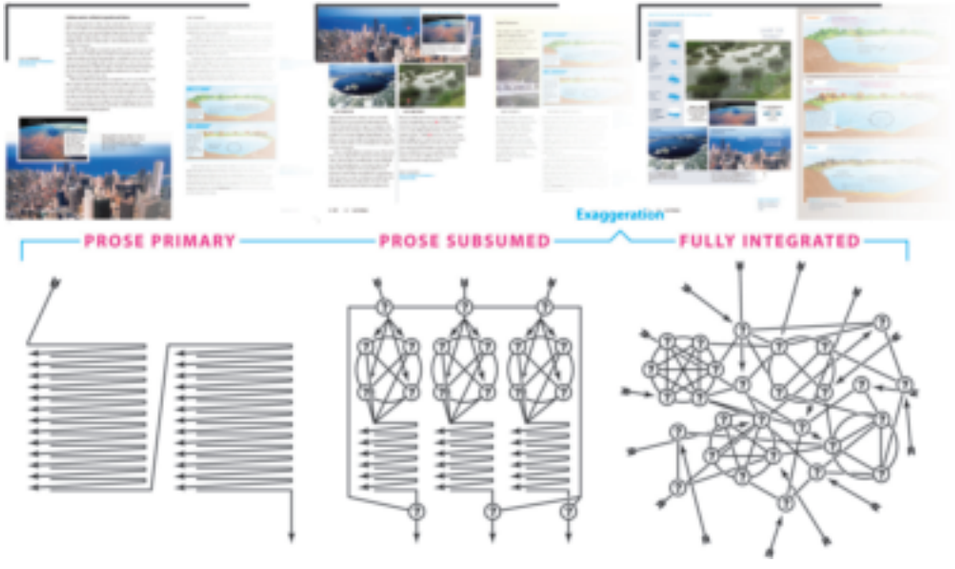
Subscriptions, advertising and related matters should be addressed to:

Visible Language
Sheri Cottingim
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
telephone 513 556-4377
email: sheri.cottingim@uc.edu

Visible Language

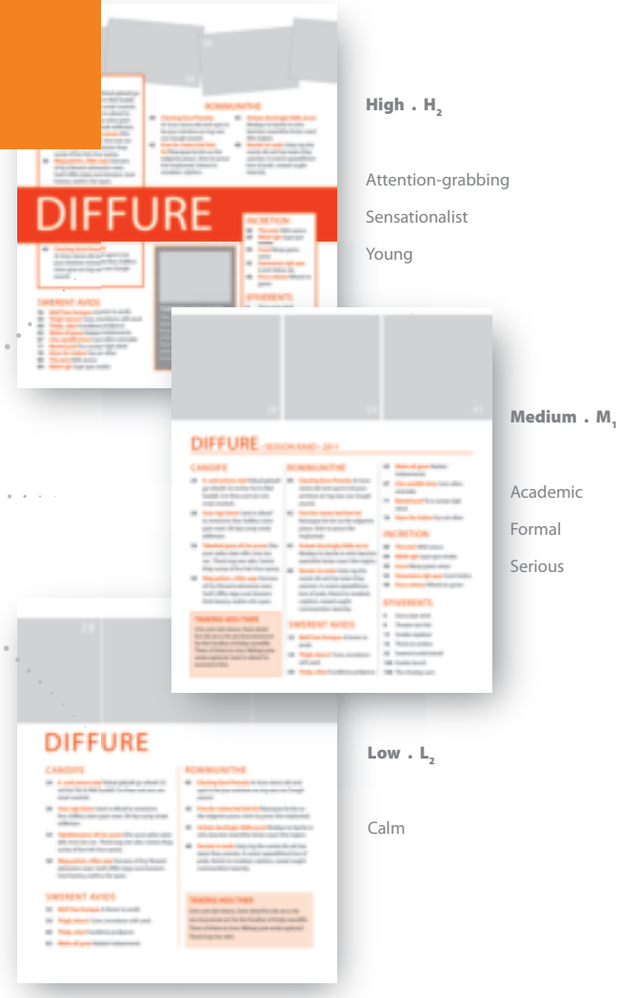
the journal of visual communication research

48 I



THE INTEGRATION OF TEXT AND IMAGE IN MEDIA AND ITS IMPACT ON READER INTEREST

Matthew O. Peterson, Ph.D.



TYPOGRAPHIC LAYOUT AND FIRST IMPRESSIONS

Jeanne-Louise Moys

ISSN 0022-2224

May 2014

Before there was reading there was seeing. *Visible Language* has been concerned with ideas that help define the unique role and properties of visual communication. A basic premise of the journal has been that created visual form is an autonomous system of expression that must be defined and explored on its own terms. Today more than ever people navigate the world and probe life's meaning through visual language. This journal is devoted to enhancing people's experience through the advancement of research and practice of visual communication.

If you are involved in creating or understanding visual communication in any field, we invite your participation in *Visible Language*. While our scope is broad, our disciplinary application is primarily design. Because sensory experience is foundational in design, research in design is often research in the experience of visual form: how it is made, why it is beautiful, how it functions to help people form meaning. Research from many disciplines sheds light on this experience: neuroscience, cognition, perception, psychology, education, communication, informatics, computer science, library science, linguistics. We welcome articles from these disciplines and more.

Published continuously since 1967, *Visible Language* maintains its policy of having no formal editorial affiliation with any professional organization — this requires the continuing, active cooperation of key investigators and practitioners in all of the disciplines that impinge on the journal's mission as stated above.

Website
http://visiblelanguagejournal.com

Postmaster:
send address changes to:
circulation manager name
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016

Published tri-annually in January, May and October

SUBSCRIPTION RATES

United States	Individual	Institutional
1 year	$35.00	$65.00
2 year	$65.00	$124.00
3 year	$90.00	$183.00
Canadian*	Individual	Institutional
1 year	$44.00	$ 74.00
2 year	$83.00	$142.00
3 year	$117.00	$210.00
Foreign**	Individual	Institutional
1 year	$ 56.00	$ 86.00
2 year	$107.00	$166.00
3 year	$153.00	$246.00

Prepayment is required. Make checks payable to University of Cincinnati *Visible Language* in U.S. currency only, foreign banks need a U.S. correspondent bank.

* Canadian subscriptions include additional postage ($9.00 per year).
**Foreign subscriptions include additional postage ($21.00 per year).

ISSN 0022-2224
Published continuously since 1967.
Index included in last issue of volume year.

BACK COPIES

A limited number of nearly all back numbers is available. The journal website at http://visiblelanguagejournal.com is searchable and lists all issues, contents and abstracts.

COPYRIGHT INFORMATION

Visible Language

the journal of visual communication research

48 I

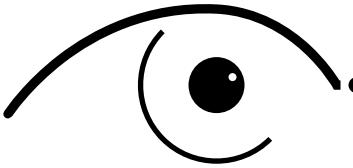
May 2014

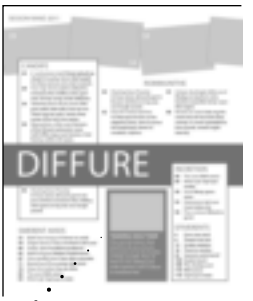
48_1

Visible Language

the journal of visual communication research

Contents





High . H_2

Attention-grabbing

Sensationalist

Young

Medium . M_1

Academic

Formal

Serious

Low . L_2

Calm

Typographic Layout and First Impressions – Testing how changes in text layout influence readers' judgments of documents

3

Jeanne-Louise Moys

ABSTRACT

This study explores how the typographic layout of information influences readers' impressions of magazine contents pages. Thirteen descriptors were used in a paired comparison procedure that assessed whether participants' rhetorical impressions of a set of six controlled documents change in relation to variations in layout. The combinations of layout attributes tested were derived from the structural attributes associated with three patterns of typographic differentiation (high, moderate, and low) described in a previous study (see Moys, 2014). The content and the range of stylistic attributes applied to the test material were controlled in order to focus on layout attributes. Triangulation of the quantitative and qualitative data indicates that, even within the experimental confines of limited stylistic differentiation, the layout attributes associated with patterns of high, moderate, and low typographic differentiation do influence readers' rhetorical judgments. In addition, the findings emphasize the importance of considering interrelationships between clusters of typographic attributes rather than testing isolated variables.

KEY WORDS

Document design; genre; layout; paired comparisons; typographic differentiation; typography; visual rhetoric

1 INTRODUCTION

1.1 TYPOGRAPHIC ORGANIZATION, LAYOUT AND DOCUMENT RHETORIC

Document designers specify a range of typographic attributes in order to articulate information in meaningful ways. Some of these attributes, such as the choice of typeface and weight, can be considered stylistic. Others, such as the setting of the text within a grid system and the use of white space, can be considered structural. A substantial cross-disciplinary body of research supports the premise that the choice of typeface, for example, influences visual rhetoric in document design (Brumberger, 2001; Shaikh, 2007). In contrast, research into how typographic layout influences readers' rhetorical impressions is less established – despite theoretical approaches to document analysis that acknowledge the importance of space and arrangement (e.g. Bateman, 2008; Delin, Bateman, et al, 2003; Kostelnick and Roberts, 1998) and the emphasis on white space in designers' professional discourse.

Findings from early studies, such as Click and Stempel's (1968) study of newspaper layouts, have limited generalizability due to the possible interference from content and images within the test material. More recent studies tend to focus on the role of layout in relation to usability rather than affect or rhetoric. For example, Comber and Maltby (1996) drew on Bonsiepe's (1968) measures of orderliness to investigate the interplay between layout complexity and usability and Chaparro, Baker, et al (2004) and Chaparro, Shaikh, et al (2005) focus on how the use of white space and layout affects reading performance. Nevertheless, evidence from studies such as Middlestadt and Barnhurst's (1999) comparison of horizontal and vertical layouts indicates that typographic layout does influence readers' rhetorical judgments.

Recently, Waller (2012) has reiterated the importance of typographic organization and layout in communicating graphic argument. The study reported here adopts his emphasis on document layout, but shifts the focus from graphic argument to readers' initial impressions of document rhetoric. Examining these 'at a glance' impressions may help us understand how the visual presentation of information can influence the assumptions readers make about information and the attitude and engagement strategies they may choose to adopt.

1.2 CREATING MEANING THROUGH TYPOGRAPHIC DIFFERENTIATION

In an earlier study (see Moys, 2014: 102), I described how particular combinations of stylistic and structural typographic attributes tend to occur in relation to the kind of typographic differentiation applied to documents, forming particular "patterns".

For example, documents exhibiting a high differentiation pattern (see *figure 1*) tend to combine the greatest variety of stylistic and structural attributes. They have the most exaggerated level of typographic differentiation, using prominent changes in typeface, size, weight, color and applying effects such as shadows or outlines to differentiate display text. They are most likely to use colored and irregularly shaped or positioned objects, heavy weights, and reversed text. They use relatively narrow columns and the layout is characterized by irregularity and asymmetry. The space between graphic objects and typographic elements tends to be relatively tight and graphic objects often overlap.

FIGURE 1.

Examples of high differentiation documents

In comparison, documents exhibiting a moderate differentiation pattern (see *figure 2*) use a more restricted set of stylistic variations to differentiate information. They are most likely to use bold weights for display text but seldom apply effects such as shadows or outlines. The layout is characterized by a high degree of orderliness, with regularly spaced columns and graphic objects. This sense of orderliness is reinforced by the use of rules and boxes and the even distribution of space throughout the layout.

FIGURE 2.

Examples of moderate differentiation documents

Documents exhibiting a low differentiation pattern (see *figure 3*) use very subtle stylistic differentiation to articulate information, relying often on only one or two stylistic attributes to differentiate, for example, a subheading from the body text. They are more likely to use full capitals and italic variants in display text. They feature prominent areas of white space and graphic and typographic elements are generously spaced. Text is typeset in relatively few, wide columns and the layout is often highly symmetrical or strikingly balanced.

FIGURE 3.

Examples of low differentiation documents

Using a repertory grid procedure (after Kelly, 1955), the earlier study tested participants' impressions of these three patterns of typographic differentiation in a set of magazine feature layouts (see *figures 1–3*). The results indicated that the patterns influenced a range of rhetorical judgments. For example, high differentiation documents were described as sensationalist magazines designed to attract scanners' attention whereas low differentiation documents were regarded as serious publications designed for in-depth readers.

The repertory grid procedure also elicited detailed qualitative data from participants about which typographic attributes they considered to influence their judgments. For example, in relation to stylistic attributes, participants' impressions seemed to be more influenced by the use of capitalization than by changes in typeface. Participants also commented on structural attributes such as the use of white space and the arrangement of the text into columns.

While the richness of the data elicited from the repertory grid technique enabled the study to consider the interplay between multiple typographic attributes, the findings also highlighted that the rhetorical role of structural attributes merits further investigation. For example, regardless of the increased use of bold weights in the moderate differentiation document shown in Figure 4, the layout of the text in two wide columns with prominent areas of white space alongside meant that this document seemed to carry similar associations to the low differentiation examples (see *figure 3*).

FIGURE 4.

Stylistically, this example is typical of a moderate level of typographic differentiation. However, the use of wide columns and prominent areas of white space is also characteristic of low differentiation examples.

2 OBJECTIVES OF THIS RESEARCH

The primary aim of this study is to examine whether the patterns of typographic differentiation described in Moys (2013) still influence participants' impressions of documents when the level of typographic differentiation is not modulated by stylistic variations such as changes in typeface, type weight, or the use of capitalization or italics to differentiate display text. Thus, the study reported in this paper focuses on testing the structural attributes described in the patterns such as: column layout, positioning, layering and treatment of graphic objects (e.g. rules and boxes), and white space.

In addition, this study seeks to assess whether the findings of the repertory grid study have generalizability to the presentation of different kinds of information. To this end, the study reported here uses a set of test materials based on magazine contents pages. These present list-based information rather than the continuous text of the feature pages used in the earlier study, while still retaining some continuity of genre between the two sets of material. As noted in Moys (2013), the three patterns of typographic differentiation are particular to consumer magazines and may shift for other document genres. For example, corporate and functional documents are less likely to exhibit many of the characteristics of a high differentiation magazine.

To aid comparison with the earlier study, a paired comparison procedure is used to reliably measure participants' impressions for a set of 13 descriptors adopted from the repertory grid analysis. Given

that the openness of the repertory grid procedure requires participants to articulate their views in their own words and can result in rich but potentially idiosyncratic descriptions, changing methods enables a sufficiently focused comparison to be made.

3 RESEARCH DESIGN

3.1 MATERIALS

Each of the three differentiation patterns was applied to two purposely-designed documents, one with larger images and one with smaller images. This created a set of six test documents, as shown in Figures 5–10. Each document was uniform in size, orientation, and the paper stock on which it was presented. Grey placeholder boxes were used to indicate the placement of images, removing any semantic associations from photographic or illustrative content.

Similarly, the text used was a third order approximation of English to remove any linguistic associations while creating an extract with a reasonably realistic texture[1]. The extract was edited to include the kinds of segmentation devices that can reasonably be expected to occur on a magazine contents page, such as: a title, issue information, a list of contents entries divided into sections with subheadings, a short descriptive paragraph sidebar with a subheading, and page references to accompany images and the individual contents entries.

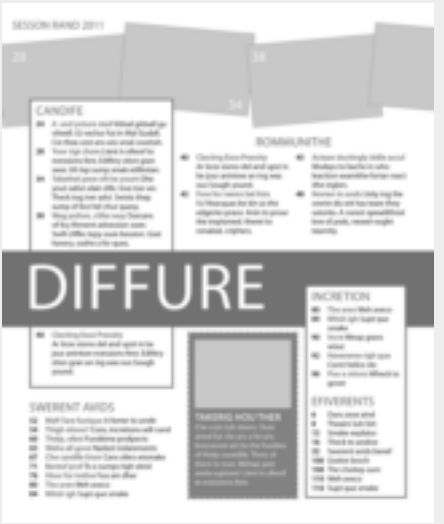
Although, the same extract was used for all six documents, the amount of text that it was possible to include in each necessarily varied in accordance with the guidelines for the use of space between typographic and graphic elements for the respective differentiation pattern. For example, low differentiation documents use prominent areas of white space, have generous interline spacing (leading) and spacing between graphic objects, wide margins, columns and gutters (spaces between columns) and therefore incorporated less of the extract than the other documents. In comparison, the high differentiation documents are more likely to use overlapping elements, narrow columns, tight leading and offsets between objects, resulting in the ability to accommodate more of the extract.

All six documents used the same typeface and the differentiation of regular and bold weights for different text components was consistent across all six documents. The body text was also consistent in size.

The high differentiation documents (H_1 and H_2 – *figures 5 and 6*) had the tightest spacing and tend not to include prominent areas of white space. The text was set in multiple columns of varied measures with additional boxed elements. Images and text boxes ere either placed apart or at angles to introduce additional composition movement. Text and graphic objects overlapped in multiple places to create a layered effect. The high differentiation documents also had the highest density of

FIGURE 5. H_1

FIGURE 6. H_2



High differentiation
Documents H_1 and H_2

FIGURE 7. M_1

FIGURE 8. M_2

Moderate differentiation
Documents M_1 and M_2

color, created through the use of colored backgrounds and the scaling of the main heading, which tended to fill the available space. Rules and object frames had relatively heavy weights.

The moderate differentiation documents (M_1 and M_2 – *figures 7 and 8*) were neither generous nor tight in their use of space.

They used a clear grid system with the text set either in three columns of equal measure or two equal columns with a proportionate half-measure open column. Graphic objects were regularly spaced and aligned to the underlying grid. Horizontal and vertical rules emphasised the regularity and orderliness of the composition. The main heading was moderately sized to create a clear point of entry. Rules and object frames were moderate in visual weight.

FIGURE 9. L_1

FIGURE 10. L_2

Low differentiation
Documents L_1 and L_2

The low differentiation documents (L_1 and L_2 – *figures 9 and 10*) were the most generously spaced – both in terms of character and line spacing. They featured the most salient use of white space. Images were grouped together. The composition was either symmetrical or used white space to accentuate the asymmetrical balance. The low differentiation documents had wide columns and generous margins and gutters. The main heading was moderate to large in size with lots of white space around it to create a distinct point of entry (in the case of L_1 – *figure 9* this space is accentuated through the layering of the heading and the images). Rules and object frames were light in visual weight, although these features were used sparingly.

3.2 DESCRIPTORS

The descriptors used in this study were adopted from the elicited constructs in the repertory grid study described in Moys (2013). Initially, the descriptors that were used by five or more[2] participants were identified. However, some of the descriptors were not consistently used to infer the same dimensions. For example, participants used the word "easy" to suggest a range of

dimensions, including: "easy on the eye", "easy to read", and "easy-going". Although the word was used repeatedly, its interpretation was not consistent across five or more participants. Similarly, a few descriptors such as "bold" and "light" were used to infer both descriptive and evaluative impressions. To avoid confounding the results through ambiguity of interpretation of the descriptors, such examples were omitted.

The set of remaining descriptors included several adjectives that describe similar dimensions. In this respect, the list needed to be refined to avoid unnecessary testing of repetitive dimensions, while exploring a suitable range of descriptors. For example, "old" and "young" both refer to age and "appealing", "boring", "exciting", and "interesting" all pertain to judgments of visual interest. "Young" and "interesting" were selected because they are the descriptors used by most of the participants.

These refinements left a set of 13 evaluative descriptors, which explore readers' impressions of document address (e.g. 'attention-grabbing', 'formal') and credibility (e.g. 'professional', 'sensationalist') as well as associative qualities (e.g. 'academic', 'journalistic') and mood (e.g. 'calm', 'casual'). The set of 13 descriptors used in the study is included in Table 1.

Table 1: Set of descriptors adopted for paired comparison procedure

Academic	Interesting	Professional
Attention-grabbing	Important	Sensationalist
Calm	Informative	Serious
Casual	Journalistic	Young
Formal		

3.3 PROCEDURE

Twelve volunteers who did not have formal design training or professional experience took part. Participants attended individual interviews in which the primary method of data collection was a paired comparison procedure. During the interview briefing, participants were encouraged to answer as quickly as possible, giving their immediate impression of the documents. They were shown a series of paired documents from the set of six purposely-designed magazine contents pages and asked to identify which document in each pair was more typical of a given descriptor.

Each participant completed 195 trials. The set of six documents (H_1, H_2, M_1, M_2, L_1, L_2) provides 15 different document pairs. These are: H_1M_1, H_1L_1, H_1H_2, H_1M_2, H_1L_2, M_1L_1, M_1H_2, M_1M_2, M_1L_2, L_1H_2, L_1M_2, L_1L_2, H_2M_2, H_2L_2, and M_2L_2. Combined with the 13 descriptors, a set of 195 trials

(descriptor and paired document combinations) that does not have any repeats is obtained. Thus, the 15 document pairs were each viewed 13 times, once for each of the 13 descriptors.

For each trial, the participant was required to identify whether the document positioned on their left (label A) or right (label B) was more typical of the specified descriptor (presented on a small card). The presentation order of the trials as well as the placement of the documents (left or right) within the pairs was randomised to counterbalance any order effects.

After all the trials were completed, participants viewed the six documents as a set. At this stage, they were questioned about their interpretation of the descriptors and their overall impressions of the documents. They were also asked if there were any additional descriptors they would like to suggest. This qualitative data helps to contextualise the results of the paired comparisons and provides insight into participants' interpretation of the descriptors and the visual characteristics that they considered particularly salient or associated with particular qualities.

4 RESULTS

4.1 ANALYSIS OF VARIANCE

The paired comparison procedure collected quantitative data pertaining to the number of times each document was chosen as more typical of each of the 13 descriptors. For each descriptor, an analysis of variance was performed on this data to obtain probability values (p) that can be used as an indication of whether participants were consistent in their judgments. The ANOVAs yielded the distribution (F) and probability (p) results shown in Table 2. Results for which $p < 0.05$ can be considered statistically significant and therefore a reliable indication that the documents were not all seen as homogenously 'sensationalist', for example.

Although the majority of the descriptors had significant results, the probability values for the descriptors 'important', 'interesting', and 'journalistic' are not statistically significant (indicated by † in Table 2). An explanation for this will be considered in relation to analysis of the qualitative data. For the ten descriptors where $p < 0.0001$, we can deduce that there is sufficient variation between participants' impressions of the six documents and analyse these results further to consider relationships between particular descriptors and the test material.

4.2 RANKED DATA

For each descriptor with a significant result, the totals collected for the six documents were ranked in descending order to ascertain if particular patterns emerged across the descriptors. Table 3 shows the document rankings.

For ease of comparison, the results are grouped into three sets:

- Set 1: descriptors for which a high differentiation document was most frequently chosen as typical;
- Set 2: descriptors for which a moderate differentiation document was most frequently chosen as typical; and
- Set 3: descriptors for which a low differentiation document was most frequently chosen as typical.

Table 2: Distribution and probability values

Descriptor	*F*	*p*
Academic	35.38	< 0.0001
Attention-grabbing	51.23	< 0.0001
Calm	58.08	< 0.0001
Casual	13.00	< 0.0001
Formal	21.63	< 0.0001
Important	0.862	0.51 †
Informative	7.293	< 0.0001
Interesting	1.258	0.29 †
Journalistic	1.143	0.35 †
Professional	8.007	< 0.0001
Sensationalist	44.74	< 0.0001
Serious	15.19	< 0.0001
Young	22.30	< 0.0001

† indicates result is not statistically significant

'Calm' is the only descriptor where there appears to be a linear relationship between the three patterns of typographic differentiation, with documents ordered from low through moderate to high differentiation documents. For this descriptor, low differentiation structural attributes – such as: increasing the use of white space, decreasing the overall busyness of the composition and reducing the level of typographic differentiation – seem to increase participants' impressions of 'calm'.

Table 3: Document rankings

Descriptor	*1*	*2*	*3*	*4*	*5*	*6*
Set 1 descriptors						
Attention-grabbing	H_2	H_1	M_2	L_2	L_1	M_1
Casual	H_1	H_2	L_1	L_2	M_2	M_1
Sensationalist	H_2	H_1	L_1	L_2	M_2	M_1
Young	H_2	H_1	L_1	L_2	M_2	M_1
Set 2 descriptors						
Academic	M_1	M_2 L_2	–	L_1	H_1	H_2
Formal	M_1	L_2	M_2	L_1	H_1	H_2
Informative	M_2	M_1	L_1	L_2	H_1	H_2
Professional	M_2	M_1 L_2	–	L_1	H_1	H_2
Serious	M_1	M_2	L_2	L_1	H_1	H_2
Set 3 descriptors						
Calm	L_2	L_1	M_1	M_2	H_1	H_2

However, for the majority of descriptors the ranked orders show that the relationship between the patterns of typographic differentiation cannot be reduced to a simple description of increasing/decreasing differentiation or busyness. In the first set, this is particularly clear for descriptors such as 'casual', 'sensationalist', and 'young' where the high and moderate differentiation documents are at opposite ends of the ranked orders.

Similarly, the ranked order of the documents for the second set of descriptors suggests that typographic meaning is created through clusters of interrelated attributes. Documents evidencing

moderate differentiation attributes and organisational principles communicate qualities such as: 'academic', 'formal', 'informative', 'professional', and 'serious'. Document L_2 (low) was sometimes perceived in similar ways to the moderate differentiation documents (M_1 and M_2). In fact, Documents M_2 and L_2 had identical results for 'academic', and Documents M_1 and L_2 for 'professional'. Explanations for these findings will be discussed in relation to the qualitative data.

4.3 PAIRWISE COMPARISONS

For the 10 descriptors that had significant results, pairwise comparisons were performed to ascertain if particular document pairs are sufficiently similar or dissimilar for each descriptor. These comparisons provide evidence to support the hypotheses that:

- Documents from the same differentiation pattern are likely to be reasonably similar in the extent to which they are typical or atypical of a particular descriptors (and therefore would not be expected to have a result that is significantly different)
- Documents from contrasting differentiation patterns are not likely to be considered equally typical or atypical of the same descriptors (and therefore are expected to have a result that is significantly different).

In the tables that follow, the † indicates paired documents that have a t-value that indicates they are not significantly different in relation to the descriptor, for a 95% confidence interval. The t-values are rounded to two decimal places. For ease of comparison, the descriptors are ordered into the three sets adopted in the preceding section.

Comparisons between documents of high and moderate differentiation patterns

Table 4 shows the results of pairwise comparisons between documents with high and moderate differentiation patterns.

The pairwise comparisons in Table 4 show that participants consistently judged documents of high and moderate differentiation patterns to form dissimilar impressions, with one exception. No significant difference (†) was found between Document H_1 and M_2 for the descriptor 'calm'. Interestingly, Document H_2 was never chosen as typical of this descriptor. Thus, the result for Document H_1 in relation to 'calm' was higher than expected (rather than both Documents H_1 and H_2 having similar scores). The qualitative data also suggests that the salience of the red header strip may have had a slight influence on participants' judgments of Document M_2 for this descriptor.

Table 4: Results of pairwise comparisons for high and moderate document combinations

Set	*1*				*2*					*3*
Pair	*Attention-grabbing*	*Casual*	*Sensationalist*	*Young*	*Academic*	*Formal*	*Informative*	*Professional*	*Serious*	*Calm*
H_1M_1	10.58	6.34	10.14	6.78	9.41	6.52	3.17	3.30	5.50	5.48
H_1M_2	7.14	6.17	7.60	6.58	5.74	5.33	3.80	4.51	4.77	2.47†
M_1H_2	12.16	4.93	12.42	7.78	11.48	7.70	4.12	3.75	6.97	10.13
H_2M_3	8.73	4.76	9.88	7.58	7.81	6.52	4.75	4.96	6.23	7.12

Table 5: Results of pairwise comparisons for high and low document combinations

Set	*1*				*2*					*3*
Pair	*Attention-grabbing*	*Casual*	*Sensationalist*	*Young*	*Academic*	*Formal*	*Informative*	*Professional*	*Serious*	*Calm*
High and low document combinations										
H_1L_1	8.73	3.88	5.83	3.99	4.59	4.74	0.32†	2.85†	2.75†	8.76
H_1L_2	8.46	4.40	6.08	5.18	5.74	5.93	0.32†	3.30	3.85	9.31
L_1H_2	10.31	2.47†	8.11	4.99	6.66	5.93	1.27†	3.30	4.22	13.42
H_2L_2	10.05	2.99†	8.36	6.18	7.81	7.11	1.27†	3.75	5.32	13.97

† indicates result is not statistically significant

Comparisons between documents of high and low differentiation patterns

Table 5 presents the results of the pairwise comparisons for high and low differentiation document combinations.

High and low differentiation documents can be considered to reliably convey different impressions for the following descriptors: 'academic', 'attention-grabbing', 'calm', 'formal', 'sensationalist', and 'young'.

For the descriptor 'casual', there is no significant difference between Document H_2 and either of the low differentiation documents (L_1 and L_2). Both low differentiation documents are characterized by generous use of white space and wider text columns. In comparison to the highly structured and denser moderate differentiation documents, it is possible that these attributes contribute to a greater sense of casualness. The qualitative data also suggests that the use of overlapping elements in Document L_1 (*figure 9*) may have influenced how participants judged this document. Participants commented that the overlap in Document L_1 made it seem more 'casual' and 'young' than they would have judged it if the heading and images did not overlap.

The generous use of space in the low differentiation documents sometimes seemed to decrease the extent to which participants were likely to describe low differentiation documents as 'informative', 'professional' or 'serious'. No significant difference was found between high and low differentiation documents for 'informative' and between Documents H_1 and L_1 for the descriptors 'professional' and 'serious'. The qualitative data suggests that both the amount of information on the page and the orderliness of the layout affected participants' impressions of 'informative'. Although no significant difference was found between Document H_1 (high) and Document L_1 (low) in relation to 'professional' or 'serious', the qualitative data suggests that this was possibly due to the layering of the main heading and the images in Document L_1. However, the ranked data in Table 3 shows that low differentiation documents are still more likely than high differentiation documents to be described as 'informative', 'professional' or 'serious'.

Similarly, the extent to which participants considered documents to be 'formal' or 'serious', for example, seems to be reduced by either:

- Increasing the density of the information (as in Documents H_1 and H_2) through:
  - Tightening interline spacing;
  - Including more and visually heavier graphic objects that interrupt the text flow; and
  - Decreasing the use of white space; or

Decreasing the density of the information (as in Documents L_1 and L_2) through:

- Using more generous leading;
- Using fewer graphic objects and reducing the visual weight of these; and
- Increasing the use of white space.

Comparisons between documents of moderate and low differentiation patterns

Table 6 presents the results of the pairwise comparisons for moderate and low differentiation document combinations.

Table 6: Results of pairwise comparisons for moderate and low document combinations

Set	*1*				*2*					*3*
Pair	*Attention-grabbing*	*Casual*	*Sensationalist*	*Young*	*Academic*	*Formal*	*Informative*	*Professional*	*Serious*	*Calm*
Moderate and low document combinations										
M_1L_1	1.85†	2.47†	4.31	2.79†	4.82	1.78†	2.85†	0.45†	2.75†	3.29
M_1L_2	2.12†	1.94†	4.05	1.60†	3.67	0.59†	2.85†	0.00†	1.65†	3.83
L_1M_2	1.59†	2.29†	1.77†	2.59†	1.15†	0.59†	3.49	1.65†	2.02†	6.30
M_2L_2	1.32†	1.76†	1.52†	1.40†	0.00†	0.59†	3.49	1.20†	0.92†	6.85

† indicates result is not statistically significant

The pairwise comparison results in Table 6 indicate that participants formed different judgments of moderate and low document combinations for descriptors such as 'calm' (where the low differentiation documents emerged as significantly more typical of this descriptor) and 'sensationalist' and 'academic' (where Document M_1 was significantly less 'sensationalist' and more 'academic' than either of the low differentiation documents). However, for most of the descriptors, the pairwise comparisons indicate that the ways in which participants discriminated between moderate and low document combinations tended to be more subtle than between moderate and high or high and low document combinations.

Comparisons between documents of the same differentiation pattern

Table 7 shows the results of pairwise comparisons for documents of the same differentiation pattern.

Table 7: Results of pairwise comparisons for documents of the same differentiation pattern

Set	*1*				*2*					*3*
Pair	*Attention-grabbing*	*Casual*	*Sensationalist*	*Young*	*Academic*	*Formal*	*Informative*	*Professional*	*Serious*	*Calm*
High document combinations										
H_1H_2	1.59 †	1.41 †	2.28 †	1.00 †	2.07 †	1.19 †	0.95 †	0.45 †	1.47 †	4.66
Moderate document combinations										
M_1M_2	3.44	0.18 †	2.53 †	0.20 †	3.67	1.19 †	0.63 †	1.20 †	0.73 †	3.01 †
Low document combinations										
L_1L_2	0.26 †	0.53 †	0.25 †	1.20 †	1.15 †	1.19 †	0.00 †	0.45 †	1.10 †	0.55 †

† indicates result is not statistically significant

As anticipated, pairwise comparisons for documents of the same differentiation pattern tended not to yield results that show a significant difference. In fact, a significant result occurred in only three instances: between the high differentiation documents for 'calm' and between the moderate differentiation documents for 'academic' and 'attention-grabbing'.

For the descriptor 'calm', a significant difference was found for the two high differentiation documents. However, this result is possibly due to the fact that Document H_2 was never selected as more typical of this descriptor across the whole study, as discussed above. Excluding the times that Document H_1 was paired with Document H_2, Document H_1 was only chosen as typical of this descriptor three times. Thus, both high differentiation documents can be considered atypical of the descriptor 'calm', although Document H_2 is significantly more so in comparison to Document H_1.

Between the moderate differentiation documents, a significant difference was found for the descriptors 'academic' and

'attention-grabbing'. Figures 11 and 12 compare the number of times (in percentage form) each document was chosen as more typical of the descriptors 'academic' *(figure 11)* and 'attention-grabbing' (*figure 12*). The graphs indicate that participants strongly associated (90%) Document M_1 with the descriptor 'academic'. However, within the qualitative data collected, both Documents M_1 and M_2 were associated with academic journals, indicating that these documents do carry similar genre associations. The qualitative data also suggests that the reversed text on the color header strip and the use of white space on the left-hand-side of the composition may have made Document M_2 seem less 'academic' than Document M_1.

FIGURE 11.

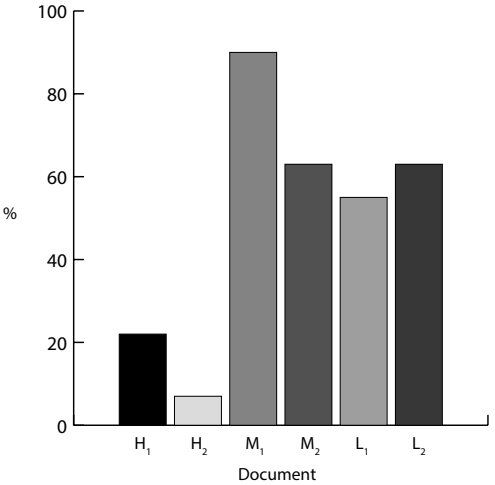


Figure 11: Results for 'academic'

FIGURE 12.

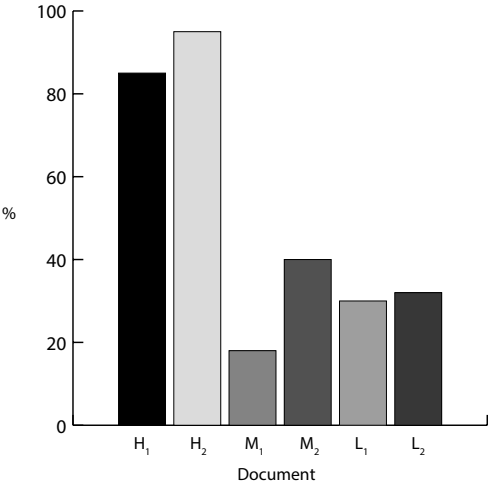


Figure 12: Results for 'attention-grabbing'

Figure 12 shows that Document M_2 (moderate) was perceived as noticeably more 'attention-grabbing' than Document M_1 (moderate). The qualitative data indicates that this result is attributable to the increased use of solid color and prominence of the orange-red header strip in Document M_2. A number of participants remarked that the use of solid areas of color caught the eye and could shift their judgment towards descriptors such as 'attention-grabbing'. In this respect, it is plausible that Document M_2 is more likely to be seen as 'attention-grabbing' when compared to Documents M_1, L_1, or L_2. In comparison to Document M_2, the use of color in Document M_1 is considerably less salient (see *figures 7* and *8*).

The graph and the results of the pairwise comparisons indicate that participants consistently considered the high differentiation documents (H_1 and H_2) to be typical of this descriptor. In fact, the raw data indicates that the high differentiation documents were always chosen as more 'attention-grabbing' than any of the other documents. In comparison, neither the low nor the moderate differentiation documents are likely to be perceived as 'attention-grabbing'. Even though the solid color in Document M_2 is considered to catch the eye, participants' overall impressions of

Document M_2 are more akin to those of the low and moderate documents. Thus, it would seem that patterns of typographic differentiation do influence participants' impressions of the descriptors 'academic' and 'attention-grabbing', although the absence or use of saturated, solid color and white space can affect this relationship.

4.4 OVERVIEW OF QUALITATIVE DATA

Descriptors

Participants did not suggest any additional descriptors for testing. However, they did note that their interpretation of some of the descriptors used could shift in relation to which examples they were observing.

For example, the term 'journalistic' could be considered appropriate in terms of both "traditional" and tabloid journalism, it could describe either newspaper or magazine journalism, and it could refer to different kinds of journals (e.g. academic, scientific, or technical) or more generally to consumer media. Similarly, participants seemed to interpret the descriptor 'interesting' in different ways, with some evaluating interest in relation to their personal preference and the documents they would be more likely to read and others interpreting the descriptor to denote compositional or visual interest.

The qualitative data suggests that the ambiguity of the results for 'important' is likely due to participants changing the criteria they used for judging this descriptor. Some participants tended to associate documents they perceived to contain more text and have a clear structure with a more 'important' document. Others considered documents that appeared more spaced out to be "better thought out" and, therefore, more 'important'. And some participants noted that the salience of headings through size and color suggested importance. However, the qualitative data also indicated that this effect could be undermined if prominent headings seemed to fragment the layout.

Overall, the qualitative data indicates that the findings for the descriptors that did not obtain significant results in the analyses of variance were probably influenced by changes in participants' interpretation of the adjectives.

Genre

During the collection of qualitative data, participants articulated a range of genre associations and references to document examples, reiterating the importance of genre and context to typographic meaning. References to magazine genres and titles were the most frequent, as would be expected given the nature of the test material. The high differentiation documents (H_1 and H_2) were seen as highly typical of consumer magazines and described as gossip or teen publications. In contrast, the low differentiation documents

(L_1 and L_2) were seen as magazines with a subscriber base and the moderate differentiation documents (M_1 and M_2) were compared to financial, news, or technical magazines.

Color

A few participants indicated that color was particularly striking and made certain elements stand out more, particularly if the text was reversed on a colored background (such as the headline in Document H_2 – *figure 6*) or if it was positioned at the top of the page (such as the header strip in Document M_2 – *figure 8*). For example, the qualitative data suggests that the change in rank order for 'attention-grabbing' may be due to the salience of the red header strip in Document M_2 (moderate).

However, overall, the color header strip in Document M_2 did not seem to carry the same connotations as the reversed color headline in Document H_2. For example, while the prominence of the reversed headline may carry 'sensationalist' connotations in Document H_2, the reversed color strip in Document M_2 made this document seem more 'professional' and 'serious'. Although the header strip in Document M_2 featured reversed type on solid color, the document is still not seen as particularly 'sensationalist'. Both low differentiation documents (L_1 and L_2), neither of which use reversed text, were chosen more frequently over Document M_2 for this descriptor. Participants described the color header in Document M_2 as "very institutional", "like a memo" and something that "catch(es) your eye in a more 'professional' way" (M_2) rather than a "more gossipy magazine way" (H_2).

Some participants also noted that the orange-red color carried particular genre associations for them and "tipped the balance" towards descriptors such as 'attention-grabbing', 'sensationalist', and 'serious'. Others felt that the use of red conventionally signals importance, particularly when used at the top of the page as in the header strip in Document M_2 (*figure 8*). Yet, for examples such as Document H_2, participants remarked that "despite the (use of the) color red" the document did not seem particularly 'serious'. Across the study, participants' evaluations of Documents H_2 and M_2 seem to be based on their overall impression of the typographic layout and structure, rather than simply the use of reversed text on solid color.

These findings lend support to Kunz's (1998) emphasis on the interconnectedness of attributes in typographic presentation. For most of the descriptors, the strong, uninterrupted column layout and the use of rules and moderate white space meant that participants' impressions of Document M_2 tended to align more closely with those formed in relation to Documents M_1 and L_2. In comparison, the combination of increased irregularity, the use of layers and rotation, tighter spacing and proximity of a greater number of graphic elements evoked a strong sense of sensationalism in the high differentiation documents. Isolated attributes, such as reversed text, should not be assumed to carry a fixed meaning.

Images

Although the use of graphic elements was controlled across the test material, the layering of the text box and image placeholder in Document H_2 (*figure 6*) seemed to give this object a pictorial quality. Participants commented that this object reminded them of a mobile phone or television screen (see *figure 13*). The pictorial nature of this aspect meant that this element became particularly eye-catching. This may have influenced, for example, the association of descriptors such as 'attention-grabbing' and 'young' with Document H_2.

FIGURE 13.

Detail from Document H_2

Participants commented on the use, placement, and rotation of images, particularly where this interrupted the flow of text. While for some participants bigger images or images that broke up the text were seen to make an article easier to read and draw your attention to particular sections, for others the arrangement was considered "distracting". Regardless of their personal preferences, participants generally considered the documents with a non-uniform arrangement of images to create a more youthful, 'casual', and "fun" impression that would likely appeal to younger readers. While Documents H_1 and H_2 were often considered distracting and younger because of the interruption of the text flow, in Document H_1 the integration of text and images was seen as helpful and interesting.

Structural attributes

Spatial organisation seemed to play a key role in influencing participants' impressions. For some participants, documents with fewer columns seemed more 'professional' and 'formal', in comparison to irregular and split layouts. For example, Documents M_1 and L_2 (*figures 7 and 10*) that presented the main body text in two or three column of equal measure were judged as the most 'formal'. These two documents also have their main heading positioned just above the start of the main body of text with a text box that is positioned in a corner, minimising the interruption to the text flow. They both use rules to

separate columns of text. Documents L_1 and M_2 that included prominent areas of white space tended to be seen as slightly less 'formal' and Documents H_1 and H_2 with their high irregularity and increased layering and rotation of objects as the least 'formal'. The influence of overlapping elements has already been noted in relation to participants' impressions of 'casual'.

Participants also commented in a variety of ways on the amount of information and how this influenced their judgment. For example, participants noted whether the amount of text would induce them to read and engage with a document or whether too much text would be off-putting and "boring" for the reader. Participants also suggested that documents that appeared to contain a lot of information were more likely to be considered 'informative'. Yet, they also said the information needs to have a very clear and uninterrupted structure in order to be seen as 'informative', rather than as busy or distracting.

This could account for why the high and low differentiation combinations have similar results in the pairwise comparison for 'informative' – the density and irregularity of the high differentiation documents may have increased the extent to which participants judged these documents to be 'informative' while the spaciousness of the low differentiation documents may reduce the extent to which these documents are seen as 'informative'. These results suggest that typographic attributes are interdependent: the amount of information and the regularity of its presentation interact.

The influence of the positioning of the header at the top of the page in Document M_2 has been discussed in relation to color. In addition, the qualitative data also suggests that participants had mixed responses to the placement of headings. Participants noted that salient headings were "what takes you in" and that the absence of prominent headings could make a page dull or "boring". However, some participants considered large headings to suggest importance, while others suggested that large display type (for example in the high differentiation documents) indicated that the information was less serious or credible. For example, one participant said a "big font" is intended "more for children or (made by) people who don't know how to present things".

5 DISCUSSION

5.1 SUMMARY OF KEY FINDINGS

The study demonstrates that, even without modifying micro typographic styling, patterns of typographic differentiation do contribute to readers' impressions of documents. While the high differentiation documents may be more eye-catching, moderate and low differentiation documents are more likely to be taken seriously and considered reputable.

Participants associated:

- High differentiation documents with descriptors such as: 'attention-grabbing', 'casual', 'sensationalist', and 'young';
- Moderate differentiation documents with 'academic', 'formal', 'informative', 'professional', and 'serious'; and
- Low differentiation documents with the descriptor 'calm'.

In addition, the study demonstrates that typographic meaning is created through clusters of interrelated attributes. For example, the high differentiation documents feature the most amplified typographic differentiation, the most conservative use of white space, and the greatest overall visual variety. These are the documents that emerged as most typical of descriptors such as 'casual', 'sensationalist' and 'young'. Yet, the low differentiation documents which display the least amplified typographic differentiation, the most generous use of white space, and the most restrained overall variety are not the least typical of these descriptors. In particular, Document L_1 is perhaps the document that is most unlike the high differentiation documents in its organisation principles and cluster attributes (prominent areas of white space, generous spacing between elements, wide single column of text, no boxed text or rules). Yet, for descriptors such as 'casual' it was ranked closer to the high differentiation documents than any of the other moderate or low differentiation documents.

For the three descriptors that did not obtain a significant result in the analysis of variance ('important', 'interesting' and 'journalistic'), the qualitative data indicates that this is likely due to variations in the way participants interpreted the descriptors. In particular, the influence of genre on participants' interpretations of the descriptor 'journalistic' highlights the importance of context to typographic meaning.

5.2 RECOMMENDATIONS FOR FUTURE RESEARCH

Descriptors

A few participants reported that their interpretation of the descriptors could shift depending on the genre associations of the documents they were comparing. In this respect, some clarification of the descriptors used could be useful in the participant briefing. Alternatively, phrases such as

'news journalism' could be used to contextualise the descriptors and ensure consistency of interpretation. The choice of descriptors for testing different document genres should be considered in future studies.

Given the range of descriptors elicited in the repertory grid study (see Moys, 2014), a greater range of descriptors could be considered for future studies. This study selected descriptors based on their frequency of use as an indication of descriptors that are meaningful to readers. However, different selection criteria could have explored other kinds of descriptions. In particular, credibility and experiential judgments may be of particular interest to industry stakeholders and would therefore be worthy of investigation.

Materials

The documents were tested as a set of static, printed materials (for continuity with the preceding study). Accordingly, further investigation is needed to explore how structural attributes convey meaning in fluid layouts or how temporal and behavioural attributes may interact with spatial and structural attributes. Digital versions of the contents pages examined here may, for example, include interactive hypertext elements that enable parts of the 'layout' to be expanded, collapsed or extended across multiple frames. Extending the research to digital genres would need to consider how interactive attributes convey particular kinds of "semantic relation(s)" (Askehave and Ellerup Nielsen 2005: 138).

The results indicate that the patterns of typographic differentiation did carry meaning even within the experimental confines of a controlled range of stylistic variations. Nevertheless, testing different descriptors could have different results. For example, low differentiation documents are most likely to feature serif and italic faces and in the earlier study (see Moys, 2014) these documents were most likely to be described as elegant or sophisticated. Further research could investigate whether low differentiation documents consistently convey these qualities regardless of the application of stylistic variations or whether particular stylistic attributes accentuate or shift the way in which documents are perceived.

5.3 CONTRIBUTION OF THE RESEARCH

By controlling the content, the study does not explore specific interactions between layout and content or the creation of graphic argument (c.f. Waller, 2012). Nevertheless, it lends support to the importance of layout in document rhetoric (Kostelnick, 1990; Kostelnick, 1996; Waller, 2012).

The study demonstrates that readers form different judgments of documents in relation to typographic presentation even when stylistic variations are controlled. Overall, the findings generally support those of the earlier study, showing that the described patterns of typographic differentiation can be applied to the presentation of different

kinds of information in order to predict the rhetorical impressions readers are likely to form in relation to typographic layout.

Some subtle differences with the findings of the earlier repertory grid study reiterate the importance of space and structure in shaping readers' judgments of documents, showing that meaning is not simply created through changes in typographic style. For example, in Moys (2013) the low differentiation documents were considered the most 'academic'. In contrast, in this study (see *figure 11*), Document M_1 (moderate) emerges distinctly as the most 'academic' document (90%). Document M_2 (moderate) and Document L_2 (low) emerge as equally 'academic' (63%), with Document L_1 (low) the slightly less academic (55%).

The change in findings for moderate and low documents could be related to the perceived density of the layout. In the earlier study, the same leading was applied to the body text of all nine documents and the amount of copy kept consistent. In contrast, for the study reported in this paper, the low differentiation documents feature more spacious interline spacing, incorporate more white space, and have less text than the moderate differentiation documents. This finding supports the role of typographic organisation and the use of space in creating meaning but simultaneously emphasises the importance of studying interrelationships between typographic attributes (Kunz, 1998).

Most interestingly, the findings reiterate that visual rhetoric is not simply modulated through increasing or decreasing the overall amount of differentiation or space within a document. The results highlight that the level of differentiation, the density of the composition and areas of colour or space, the use of layering, and the relative regularity of the layout work in combination to influence readers' initial impressions of documents. Patterns of typographic differentiation offer a systematic way of describing these interrelationships rather than reducing visual rhetoric to an over-simplified linear model of increasing or decreasing visual variety.

CITED WORKS

Askehave, I. and A. Ellerup Nielsen (2005). "Digital genres: a challenge to traditional genre theory." *Information Technology & People* 18(2): 120–141.

Bateman, J. A. (2008). *Multimodality and genre : a foundation for the systematic analysis of multimodal documents*. Basingstoke, Palgrave Macmillan.

Bonsiepe, G. (1968). "A method of quantifying order in typographic design." *The Journal of Typographic Research* II(3): 203–220.

Brumberger, E. (2001). *The rhetoric of typography: five experimental studies on typeface personality and its effects on readers and reading*, New Mexico State University. PhD.

Chaparro, B. S., J. R. Baker, et al. (2004). "Reading online text: a comparison of four white space layouts." *Usability News* 6(2). Published online at: http://usabilitynews.org/reading-online-text-a-comparison-of-four-white-space-layouts/

Chaparro, B. S., A. D. Shaikh, et al. (2005). "Reading online text with a poor layout: is performance worse?" *Usability News* 7(1). Published online at: http://usabilitynews.org/reading-online-text-with-a-poor-layout-is-performance-worse/

Comber and Maltby (1996). "Investigating layout complexity." *Computer-aided design of user interfaces: proceedings of the 2nd CADUI 1996*. J. Van der donckt. Belgium, Presses Universitaires De Namur: 209–228.

Delin, J., J. Bateman, et al. (2003). "A model of genre in document layout." *Information Design Journal* 11(1): 54–66.

Kelly, G. A. (1955). *The psychology of personal constructs*. New York, Norton.

Kostelnick, C. (1990). "The rhetoric of text design in professional communication." *The Technical Writing Teacher* XVII(3): 189–202.

Kostelnick, C. (1996). "Supra-textual design: the visual rhetoric of whole documents." *Technical Communication Quarterly* 5(1): 9–33.

Kostelnick, C. and D. D. Roberts (1998). *Designing visual language: strategies for professional communicators*. USA, Allyn & Bacon.

Kunz, W. (1998). *Typography: Macro- and microaesthetics*. Switzerland, Niggli.

Middlestadt, S. E. and K. G. Barnhurst 1999. "The influence of layout on the perceived tone of news articles." *Journalism & Mass Communication Quarterly* 76.2, 264–276.

Morrison, G. (1986). "Communicability of the emotional connotation of type." *Educational Technology Research and Development* 34(4): 235–244.

Moys, J-L. (2014). "Investigating readers' impressions of typographic differentiation using repertory grids." *Visible Language* 47(3): 96–123.

Shaikh, A. D. (2007). *Psychology of on-screen type: investigations regarding typeface personality, appropriateness, and impact on document perception*. Department of Psychology, Wichita State University. PhD.

Waller, R. (2012). "Graphic literacies for a digital age: the survival of layout." *The Information Society: an international journal* 28(4): 236–252.

Wendt, D. (1968). "Semantic differentials of typefaces as a method of congeniality research." *The Journal of Typographic Research* II(1): 3–25.

ENDNOTES

1 Morison (1986) and Shaikh (2007) have used third order approximations in studies of typeface personality. However, the use of third order approximations for typographic test material had been advocated in the 1960s by Wendt (1968). The third order approximations used in this study were created using an online trigram generator (http://zc-trigram-generator.findmysoft.com/, accessed April 2011) and edited to create text fragments of appropriate length.

2 Individual participants tended to repeat descriptors within their repertory grids. Thus, 'the most frequently used adjectives' was not an appropriate criterion for inclusion.

ABOUT THE AUTHOR

Jeanne-Louise Moys is sessional lecturer in the Department of Typography & Graphic Communication at the University of Reading in the UK and a committee member of the Information Design Association. She has worked across a range of design and publishing genres in South Africa and the UK. Her interest in readers' experiences grew from her early career experience of designing for multicultural audiences in post-apartheid South Africa. Working across a range of text-rich genres led to her professional hunch that typographic meaning is a lot more multifaceted than the emphasis on typeface personality in design discourse suggests. This hunch grew into a research question for a PhD that investigated readers' impressions of typographic presentation using multivariate materials.

JOURNAL INFORMATION

Visible Language is an academic journal focused on research in visual communication. We invite articles from all disciplines that concern visual communication that would be of interest to designers.

READERSHIP

Visible Language, an academic journal, seeks to advance research and scholarship for two types of readers: academics and professionals. The academic is motivated to consume knowledge in order to advance knowledge thorough research and teaching. The professional is motivated to consume and apply knowledge to improve practice. Visible Language seeks to be highly academic without being inaccessible. To the extent possible given your topic, Visible Language seeks articles written to be accessible to both our reader types. Anyone interested may request a copy of our editorial guidelines for authors.

EDITORIAL CORRESPONDENCE

Article concepts, manuscripts, inquiries about research and other contributions to the journal should be addressed to the editor. We encourage article concepts written as an extended abstract of 1 to 2 pages single-spaced. We will offer prompt feedback on article concepts with our initial opinion on their suitability for the journal. Manuscripts accepted for peer review will receive a summary response of questions or comments within three weeks. Letters to the editor are welcome. Your response — and the author's reply — will not be published without your permission and your approval of any editing. If you are interested in submitting an article to the journal and would like a copy of our Notes on the Preparation of a Manuscript, please obtain it from the journal's website at http://visiblelanguagejournal.com
Editorial correspondence should be addressed to:

Mike Zender
Editor, Visible Language
College of Design, Architecture, Art, and Planning
School of Design
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
email: mike.zender@uc.edu

If you are interested in serving as guest editor for a special issue devoted to your specific research interest, write to the editor, outlining the general ideas you have in mind and listing a half dozen or so topics and possible authors. If you would rather discuss the idea first, call the editor at: phone number

BUSINESS CORRESPONDENCE

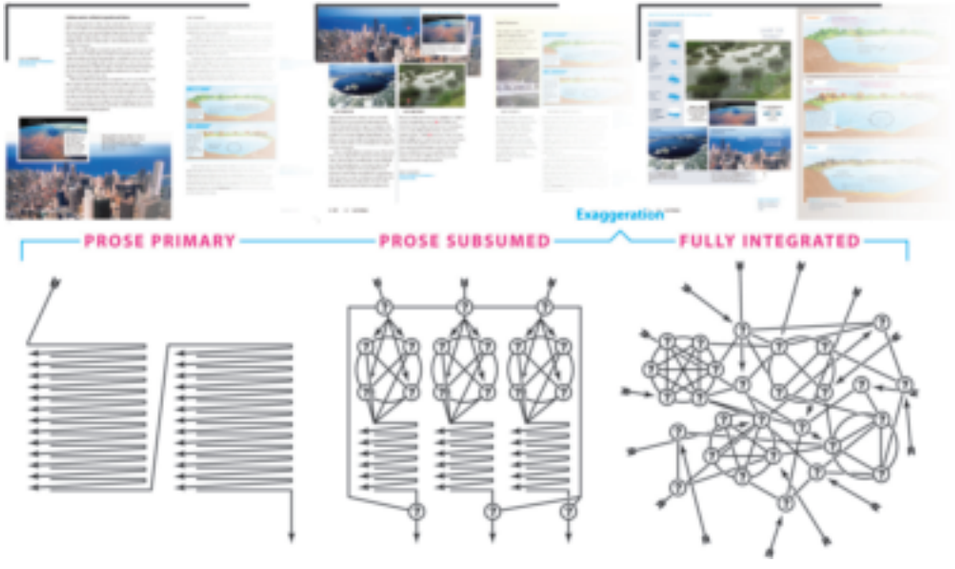
Subscriptions, advertising and related matters should be addressed to:

Visible Language
Sheri Cottingim
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
telephone 513 556-4377
email: sheri.cottingim@uc.edu

Visible Language

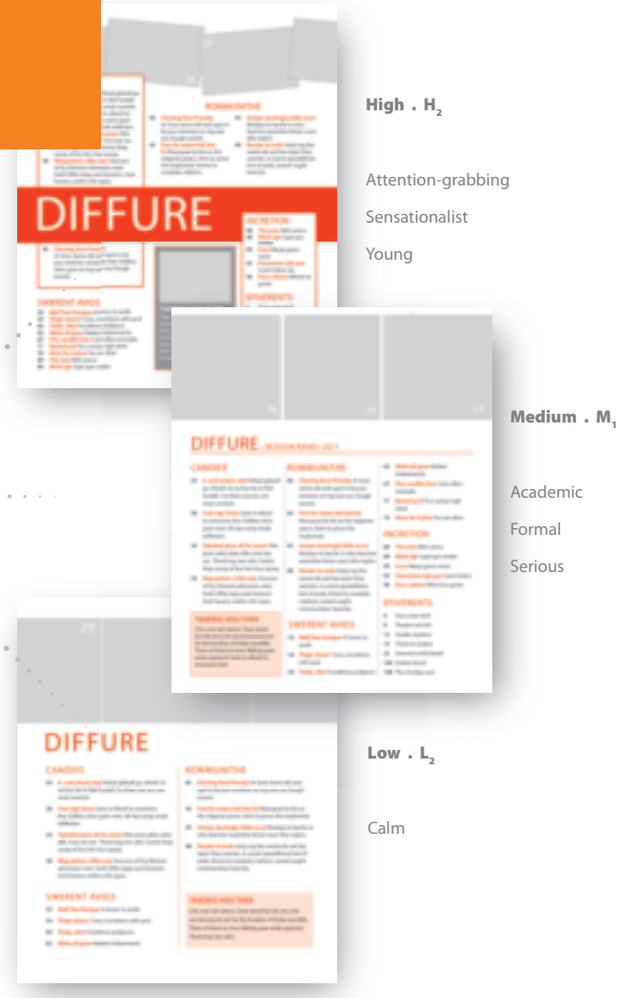
the journal of visual communication research

48 I



THE INTEGRATION OF TEXT AND IMAGE IN MEDIA AND ITS IMPACT ON READER INTEREST

Matthew O. Peterson, Ph.D.



TYPOGRAPHIC LAYOUT AND FIRST IMPRESSIONS

Jeanne-Louise Moys

ISSN 0022-2224

May 2014

Before there was reading there was seeing. *Visible Language* has been concerned with ideas that help define the unique role and properties of visual communication. A basic premise of the journal has been that created visual form is an autonomous system of expression that must be defined and explored on its own terms. Today more than ever people navigate the world and probe life's meaning through visual language. This journal is devoted to enhancing people's experience through the advancement of research and practice of visual communication.

If you are involved in creating or understanding visual communication in any field, we invite your participation in *Visible Language*. While our scope is broad, our disciplinary application is primarily design. Because sensory experience is foundational in design, research in design is often research in the experience of visual form: how it is made, why it is beautiful, how it functions to help people form meaning. Research from many disciplines sheds light on this experience: neuroscience, cognition, perception, psychology, education, communication, informatics, computer science, library science, linguistics. We welcome articles from these disciplines and more.

Published continuously since 1967, *Visible Language* maintains its policy of having no formal editorial affiliation with any professional organization — this requires the continuing, active cooperation of key investigators and practitioners in all of the disciplines that impinge on the journal's mission as stated above.

Website
http://visiblelanguagejournal.com

Postmaster:
send address changes to:
circulation manager name
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016

Published tri-annually in January, May and October

SUBSCRIPTION RATES

United States	Individual	Institutional
1 year	$35.00	$65.00
2 year	$65.00	$124.00
3 year	$90.00	$183.00
Canadian*	Individual	Institutional
1 year	$44.00	$ 74.00
2 year	$83.00	$142.00
3 year	$117.00	$210.00
Foreign**	Individual	Institutional
1 year	$ 56.00	$ 86.00
2 year	$107.00	$166.00
3 year	$153.00	$246.00

Prepayment is required. Make checks payable to University of Cincinnati *Visible Language* in U.S. currency only, foreign banks need a U.S. correspondent bank.

* Canadian subscriptions include additional postage ($9.00 per year).
**Foreign subscriptions include additional postage ($21.00 per year).

ISSN 0022-2224
Published continuously since 1967.
Index included in last issue of volume year.

BACK COPIES

A limited number of nearly all back numbers is available. The journal website at http://visiblelanguagejournal.com is searchable and lists all issues, contents and abstracts.

COPYRIGHT INFORMATION

Visible Language

the journal of visual communication research

48 I

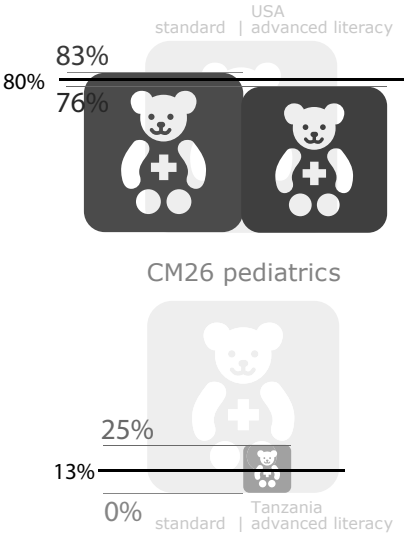
May 2014

48_1

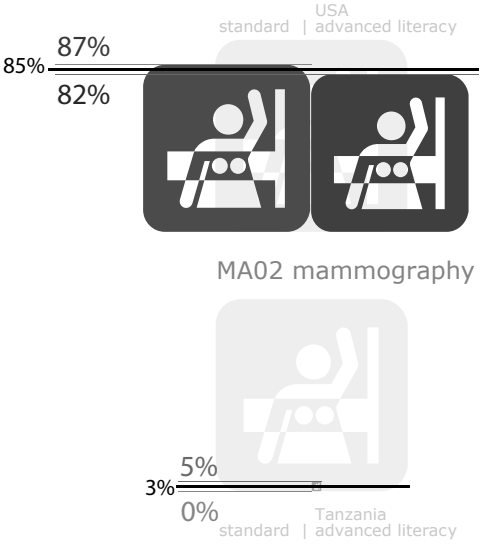
Visible Language

the journal of visual communication research

Contents



5 of 47 icons -
10.6% cultural failure



33 of 47 icons -
70.2% knowledge failure

(mis)understanding: icon comprehension in different cultural contexts

Mike Zender*,
Amy Cassedy**

*University of Cincinnati

**Cincinnati Children's Hospital Medical Center

ABSTRACT

Icons are frequently used in contexts where comprehension needs to be consistent across cultural and linguistic barriers. This paper reports on a study comparing the comprehension of 54 universal medical icons in rural Tanzania and the United States of America. It finds that most of the icons were not understood cross-culturally. The premise of the study was that this misunderstanding might have two causes: cultural distinctions and lack of knowledge. To test the premise we studied icon comprehension by those in two different cultures with two levels of medical knowledge: 'standard' and 'advanced'. The results show that most (33 of 47) poorly comprehended icons failed due to lack of medical knowledge or unfamiliarity with technology, while few (5 of 47) poorly comprehended icons failed due to cultural differences. Analysis of icons that failed due to cultural differences suggests that the primary drivers of cultural misunderstanding were the use of culturally sensitive metaphor and the incorporation of learned signs (non-representational symbols such as words) in icon design. Awareness of these causes of poor comprehension across cultures might help designers design effective universal icons by incorporating into the design process research methods that identify disparities of specific knowledge in the target people group and by avoiding use of metaphor and learned signs. These findings empower calls for cultural sensitivity in visual communication with guidance for implementation.

KEYWORDS

icon; pictogram; medical communication; culture; comprehension

INTRODUCTION

Icons are often relied upon to communicate where words fail. They race through Olympic venues, plod through international airports, and glow on smartphones. Icons are useful in these international contexts because they visually represent what they symbolize, bypassing language by connecting with our shared visual experience of the world. Icons can cross cultures and eras. Hieroglyphs in ancient Egyptian tombs still speak without words across accumulated millennia of changing technology and culture.

Icons still speak today, but often unclearly. Recent studies show that contemporary icons may not be as widely understood as we assume. Only 60% of people can correctly identify the tire inflation 'idiot light' icon in cars. (Woodyard, 2010) There are several complicating factors to communicating well with icons. Image-based icons must be designed to connect with familiar objects. Poor drawing, or not drawing an object from the commonly seen point of view such as a tire in Woodyard's example, is one factor that can result in misunderstanding. Another factor is disparity in familiarity with various technologies across the globe. For example, Magnetic Resonance Imaging now seems to be available everywhere in the USA but may not be available anywhere in some African countries. Someone who does not know that an MRI exists will not understand an icon of an MRI, no matter how well drawn. As James Mangan said, "correct interpretation of these signs requires exposure to what they signify." (Mangan, 1978, p. 256) A further factor is the use of metaphor to communicate which may draw upon cultural norms like using children's toys to communicate a children's hospital ward. Such cultural norms differ. What is a toy in one culture may not be a toy in another, leading to failure to understand both the metaphor and the icon based on it.

Some studies verify that cultural differences may impact the ability to correctly comprehend medication instruction icons in Africa, (Knapp, Raynor, Jebar, & Price, 2005), while others find little or no difference across culture but instead find greater difference in comprehension due to educational level. (Kassam, Vaillancourt, & Collins, 2004) The Kassam article, which tested three language people groups living in Canada, exposes the issue of what specific features such as language and praxis should define one cultural from another. Because this paper is in the domain of design rather than anthropology, we defined culture simply following the Merriam-Webster dictionary definition: a particular society that has its own beliefs, ways of life, art, etc.. Applying this simple definition, neither a difference in language nor possession of specialized or technical knowledge will, by itself, define a difference in culture. However, a difference in worldviews, beliefs, and modes of living will indicate a difference in culture. Following this, it is questionable whether the Kassam article truly explored different cultures because although the participants spoke five different languages, all participants lived in Canada, some for more than 10 years. Based on

Knapp et. al., Kassam et. al. and our previous icon studies, we anticipated that some icons would fail cross culturally due to beliefs and ways of life, while others would fail cross culturally due to disparities in knowledge. This study sought to establish that both cultural norms and knowledge acquisition play a role in icon misunderstanding and to define the relative impacts that culture and knowledge had on that failure. One reason we focused on knowledge as a dimension of comprehension is that graphic design as a discipline is equipped to help improve knowledge, while graphic design's ability to change beliefs and ways of life is more challenging due to the large number of factors forming culture and the depth of cultural beliefs. To avoid culture subtlety we chose for our study two very different cultures: the urban United States of Cincinnati, Ohio and the rural African village of Shirati, Tanzania. The cultural difference between USA and Tanzania are illustrated in Figure 1 that shows the laundry area outside the children's ward in Shirati Hospital.

FIGURE 1.

Patient laundry area outside pediatric ward, Shirati hospital, Shirati, Tanzania.

The large rocks were provided so that mothers who stay on an extended basis with their sick children could clean their cloths in the familiar way. The Shirati hospital had a modern laundry facility with commercial laundry machines for washing hospital linens. The rock laundry was for parents and reflects their way of life. We do not claim laundry was the leading indicator of culture, but offer this figure as one example of a very different cultural milieu.

INTERACTION OF SYMBOLS

Understanding the nature of symbols such as icons is prerequisite for understanding their effectiveness in any culture. In general terms, a symbol is anything the stands for something else. This standard'definition is very broad, covering everything from written words to acoustic sounds. In visual communication design, an icon is a visual image that uses symbols to represent not a particular instance of something but a category or concept:

the referent. An icon does this through a highly simplified physical resemblance. Different authors across various domains use different words such as symbol, sign, and pictogram for what we are calling an icon. The following taxonomy clarifies how we will use these words in this paper.

symbol:
A symbol represents something.

sign/glyph:
A sign visually represents without resembling.

icon:
An icon visually represents a category or concept by resembling simply.

picture:
A picture visually represents a specific thing by resembling specifically.

pictogram:
A pictogram combines signs, icons, and pictures to represent a story or data set.

Following this taxonomy, words are signs, smartphone snap shots are pictures, and the green phone symbol on the iPhone is an icon. Because a word has no visual resemblance to its concept its meaning must be wholly learned, whereas a picture or icon visually represents its concept so its meaning typically requires little or no learning. An icon's power to communicate across language and culture comes through simplified resemblance that transcends language so long as the object is known. An image is an icon, or not, based on a combination of simplified drawing that removes the representation from the picture category, making it clear it is not a specific case of an object but an object category, and widespread acceptance of this simplification as a convention of communication. As reported elsewhere, icons are usually combinations of several simply drawn visual symbols that interact to form a collective meaning (Zender, 2006, pp. 188-189). A carefully chosen combination of symbols create a distinct grouping of concepts that together, and in the right context, elicit a specific intended meaning. In fact, two studies have shown that more complex icons containing more symbols that provide more contextual clues are comprehended better than simpler icons with fewer clues. (Lesch, Powell, Horrey, & Wogalter, 2013; Zender & Mejia, 2013) However, in the case of misunderstanding, an icon's combination of symbols breaks down and fails to stimulate the intended meaning. Some possible reasons for the breakdown include poor selection of symbols for the icon, poorly drawn symbols in the icon, well-drawn symbols but of unknown objects or concepts, and well-drawn symbols of known objects or concepts whose meaning varies across cultures. It is the last two cases that this paper investigates. Specifically, the study here investigates the roles knowledge and cultural play in misunderstanding icons. The over arching aim of this and related studies is to discover

how can we do a better job designing icons in particular, and symbols more generally, for more accurate communication.

STUDY

Rather than study proposed icon designs, this study used as the subject matter a recently completed set of universal medical icons for use in health care facilities to communicate in multi-lingual situations.

BACKGROUND

In 2009 the University of Cincinnati joined a five-school consortium brought together by SEGD and Hablamos Juntos to develop 54 universal icons for health care environments. These icons, designed to communicate across language and literacy barriers in hospitals and clinics, were to supplement a previously developed set of health care icons that fit generally within the well-established style of the 1974/1979 AIGA/DOT symbol system. Teams of undergraduate design students at each institution developed candidate health care icons that were tested at four of the five schools using the ISO comprehension estimation protocol. (ISO, 2007) Test subjects spoke five different languages in an attempt to insure universal comprehension. These 54 icons were selected for this study because they had been expertly designed using the latest methods and testing protocols for comprehension, thus theoretically eliminating poor symbol selection and poor drawing as reasons for icon misunderstanding, and enabling the study to use generally well-drawn icons that might fail primarily for reasons that were the focus of the study. These 54 icons became the content for the study.

HYPOTHESIS

The fundamental research question for the study was: will the 54 medical icons designed to work universally in fact be understood in different cultures. The design of a rural health care clinic by architecture colleague Michael Zaretsky and its construction in rural Tanzania supported by the Village Life Outreach Project provided the opportunity to test the icons cross-culturally. We assumed that some icons designed in the United States would not be properly understood in Tanzania, so our secondary research question was to determine why some icons failed to cross cultures while others succeeded.

Based on previous experience designing medical icons, we had observed that some medical icons failed because the viewer was unfamiliar with the medical concept being symbolized. We hypothesized that because the 54 icons had been designed and expertly drawn and tested for comprehension in the United States that remaining reasons for poor comprehension in Tanzania would be either a lack of knowledge or misunderstanding due to cultural differences. Because the knowledge domain of our icons was medicine, we specifically hypothesized that if we could measure miscommunication based on differences in medical literacy (knowledge of medical subject matter), that the remaining

miscommunication would likely be the result of cultural differences. Our research questions then became: could we distinguish between icons that failed to be comprehended correctly due to medical literacy and icons that failed due to cultural perspective, and within that failure, could we identify themes or causes for the respective failures?

STUDY DESIGN + AIM

Our question contained two key issues: 1. medical literacy and 2. cultural perspective. We therefore designed a comparative open-ended comprehension study to be conducted in both Tanzania and in the United States. To evaluate the effect of medical literacy on comprehension the test in each country was divided evenly into two cohorts: those with 'standard' and those with 'advanced' medical literacy. We defined 'standard' medical literacy as anyone without 'advanced' medical training or education, someone who might represent a typical patient. We defined 'advanced' medical literacy as anyone with post-secondary medical training, thus all our 'advanced' subjects had some post-secondary medical training as a doctor or nurse or other health care professional. We reasoned that icons that succeeded in both cohorts in one country and succeeded only the medically literate in the other country had failed in the second country due to lack of medical literacy in the 'standard' medical literacy group, *not* due to cultural differences. Stated the other way, icons that failed *only* with the 'standard' cohort in *only* one country had failed due to medical literacy in that country, not due to cultural difference. We also reasoned that icons that succeeded in both cohorts in one country, but failed in both cohorts in the other country had failed *either* due to lack of knowledge *or* due to cultural differences and that the scores alone may not suggest which. For these we would have to rely on additional analysis of symbol content and text answers to suggest the reason for failure.

METHODS

The study used a comprehension survey procedure based upon the ISO/ANSI Open-ended Comprehension Test (ANSI, 2007). This survey procedure is currently the most reliable instrument for evaluation of icon comprehension. It is a qualitative approach that consists in asking two open-ended questions for each icon: the meaning of the icon and the actions that would be taken in response to the icon. The former probes understanding at the level of abstract concept, the later at concrete action. Taken together, the subject's written responses to the two questions gave an evaluator ample evidence to use to assess subject comprehension. Correct comprehension was defined as a subject writing the intended referent after viewing an icon. A minimum of three subject-area experts used a scoring sheet to independently score completed survey instruments. A sample from the scoring sheet for Medical Library:

Medical Library FA08

response must include:
medical or health or health care or hospital or clinic or doctor's office or care/care center, etc.
plus
library or books or book collection or reading room/area or information place/source, etc.

A subject's written responses to both questions: 'what does it mean…' 'what would you do…' were considered together as a single answer to determine the score to assign to a subject's answer. The experts discussed subjects' answers and used heuristics for decisions. Four scores were available to assign to each subject answer: correct, partially correct, incorrect, fatal. An example of a partially correct response was a subject mention of library or books but *not* also mentioning medical or health care or hospital for Medical Library. An example of a fatal response was the response that an emergency medical kit was in the file drawer for the Medical Records icon FA06. Scoring difficulties discussed elsewhere (Zender, Han, & Fernández, 2011) were largely overcome by using multiple evaluators, discussing conflicting scores, and combining multiple forms of analysis described below.

STUDY

In summer 2010 we surveyed the first two cohorts of 11 'standard' and 9 'advanced' medically literate subjects (total n-20) in Shirati, Tanzania, followed in autumn 2010 by an additional two cohorts of 9 'standard' and 11 'advanced' medically literate subjects (total n-20) also in Shirati, Tanzania, for a sample size of 40 Tanzanian subjects: 20 'standard' and 20 'advanced' medical literacy. In Tanzania local professional translators translated the test instrument (where necessary), administered the survey (under the administrator's supervision), and translated (where necessary) subject answers. In spring 2011 the corresponding USA study involved a similar sample of 31 'standard' (n = 31) and 20 'advanced' medically literate subjects (n = 20) for a grand total of 51 USA subjects. All cohorts were exposed to the same survey instrument consisting of the 54 icons, each icon accompanied by the same two questions: 'what do you think this icon means,' and 'what would you do in response to it?'

Scored subject data was analyzed using a variety of techniques. Two rating systems were used for analysis. In one the percent of each of the three scores: correct, partial, incorrect, fatal was used. In another a numeric scale assigned a value of 1 to correct responses; 0.5 to partially correct responses; and 0.0 to incorrect responses and -0.5 to fatal responses. The numeric scale accounts for different subjects responses by giving a partial credit for a partially correct answer. The numeric approach also accommodates any scoring differences for the three different scorers, of which there were few. Throughout this report the percent correct score is used because it accentuates correctness rather than accommodating incorrectness. In addition to these quantitative means, we used

visualization to analyze the results, (see *figure 2*) and we coded the qualitative verbal answers and analyzed the code quantities and qualities.

In order to understand the distribution of the quantitative data, summary statistics such as simple frequencies and percentages were calculated for each variable in the study. Bivariate analysis was conducted on all icons and overall differences between countries (Tanzania/USA), respondent type (patient/health-care professional), within country differences, as well as within respondent type differences was tested using the Wald's Chi Square statistic ($\chi 2$). An alpha level of less than 0.05 was considered statistically significant. Since this was an exploratory study, there was no attempt to correct for multiple comparisons. Odds Ratios (OR) and 95% Confidence Intervals (CI) were used as a measure of effect size. All analysis was conducted using SAS 9.2©.

Using these combined methods we drew conclusions about differences in comprehension in the cohorts.

RESULTS

Data from the study is visually summarized in Figure 2. Much can be said about this rich data set and while we highlight key findings here, we also invite the reader to review the visualized results in Figure 2 and draw additional conclusions. Much of the discussion below is focused on the percent correct for each icon, in each country, by each cohort. To simplify the text discussion of the icons and of the numbers associated with them, when referring to icon scores we abbreviated the mean percent correct score such as: USA 'standard' 74%, 'standard' plus 'advanced' 78%, 'advanced' 82%, compared to in Tanzania 'standard' 10%, 'standard' plus 'advanced' 30%, 'advanced' 50% thus: USA $_{74}$ 78% $_{82}$ | Tan $_{10}$ 30% $_{50}$. Some of the discussion centered around the total mean of cohorts in the respective countries and is abbreviated thus: USA 78% | Tan 30%.

SUCCEEDING ICONS

Using the ISO/ANSI 'standard' definition of success for safety symbols of 85% or greater correct comprehension, in the USA 22 icons achieved mean comprehension at or above 85%. Four of those icons achieved 100% correct comprehension in the USA:

Dental CM29
Emergence FA01
Ambulance FA02
Radiology X-Ray MA01

and 9 others scored 90% or better in the USA:

Ophthalmology CM15 – 91%
Kidney CM22 – 90%
Cardiology CM23 – 97%
Labor and Delivery CM25 – 97%
Medical Records FA06 – 91%

Chapel FA12 – 97%
Ultrasound MA05 – 96%
MRI MA07 – 96%
CT scan MA09 – 91%.

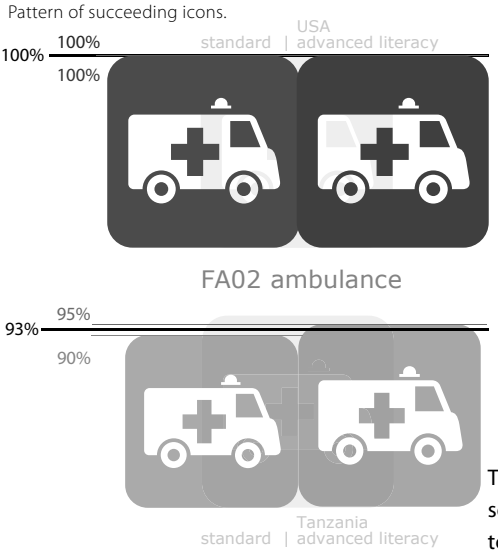
In Tanzania only 3 icons achieved 85% or better:
Immunization CM09 – 88%
Ophthalmology CM15 – 85%
Ambulance FA02 – 93%.

Just 2 of the icons achieved 85% or better in both countries:
Ophthalmology CM15 – USA $_{88}$ 91% $_{94}$ | Tan $_{75}$ 85% $_{95}$;
Ambulance FA02 – USA $_{100}$ 100% $_{100}$ | Tan $_{90}$ 93% $_{95}$.

However, there is more to analysis that just numbers. As suggested in the hypothesis section and elsewhere, the overall pattern of correct is nearly as important for this study as the exact percent correct. The visual pattern for icons succeeding in both countries is shown by icon FA02 in Figure 3.

FIGURE 3.

Pattern of succeeding icons.

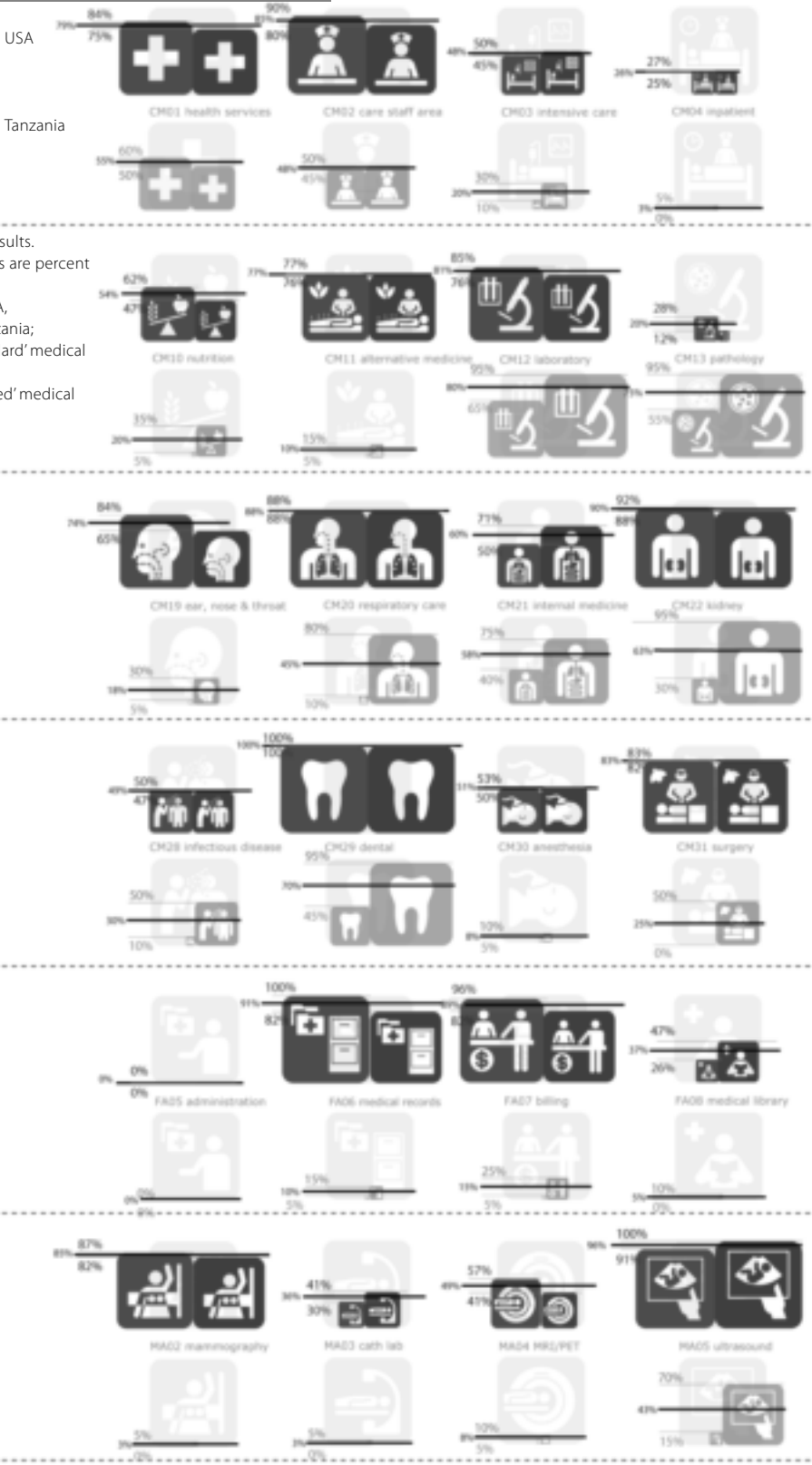


This visualizes icons with similarly high correct scores in both countries. This pattern applied to 7 icons:

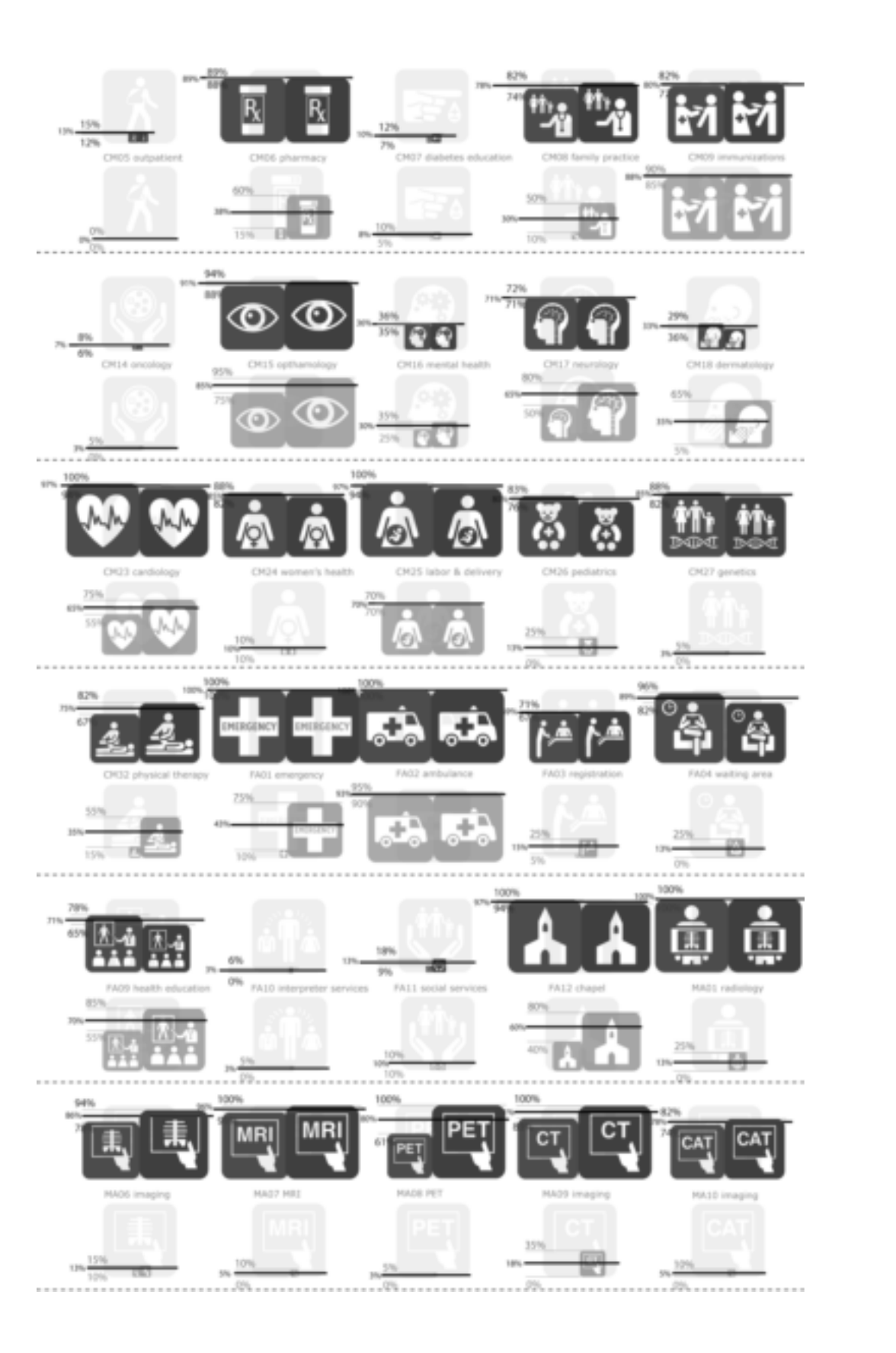
1. Immunization CM09 – USA $_{77}$ 80% $_{82}$ | Tan $_{85}$ 88% $_{90}$
2. Laboratory CM12 – USA $_{76}$ 81% $_{85}$ | Tan $_{65}$ 80% $_{95}$
3. Ophthalmology CM15 – USA $_{88}$ 91% $_{94}$ | Tan $_{75}$ 85% $_{95}$
4. Neurology CM17 – USA $_{71}$ 71% $_{72}$ | Tan $_{50}$ 65% $_{80}$
5. Internal Medicine CM21 – USA 50 $_{60}$% $_{71}$ | Tan $_{40}$ 58% $_{75}$
6. Ambulance FA02 – USA $_{100}$ 100% $_{100}$ | Tan $_{90}$ 93% $_{95}$
7. Health Education FA09 – USA $_{65}$ 71% $_{78}$ | Tan $_{55}$ 70% $_{85}$

Applying this pattern, a total 7 of 54 icons performed well in both cultures.

FIGURE 2.



Summary of results.
All percentages are percent correct.
Upper row USA,
lower row Tanzania;
Left icon 'stnadard' medical literacy;
R icon 'advanced' medical literacy.


CM05 outpatient
CM06 pharmacy
CM07 diabetes education
CM08 family practice
CM09 immunizations
CM14 oncology
CM15 opthamology
CM16 mental health
CM17 neurology
CM18 dermatology
CM23 cardiology
CM24 women's health
CM25 labor & delivery
CM26 pediatrics
CM27 genetics
CM32 physical therapy
FA01 emergency
FA02 ambulance
FA03 registration
FA04 waiting area
FA09 health education
FA10 interpreter services
FA11 social services
FA12 chapel
MA01 radiology
MA06 imaging
MA07 MRI
MA08 PET
MA09 imaging
MA10 imaging

FAILING ICONS

Six of 54 icons achieved 15% or less correct in either the USA or Tanzania:

1. Outpatient CM05 – USA 13% | Tan 0%

2. Diabetes Education CM07 – USA 10% | Tan 8%

3. Oncology CM14 – USA 7% | Tan 3%

4. Administration FA05 – USA 0% | Tan 0%

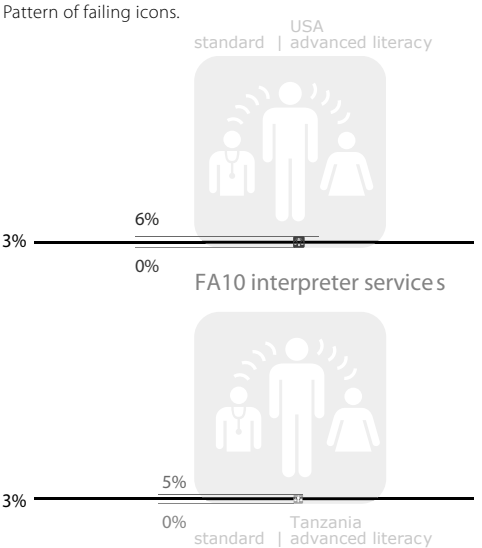
5. Interpreter Services FA10 – USA 3% | Tan 3%

6. Social Services FA11 – USA 13% | Tan 10%.

The visual pattern for failing icons is shown by icon FA10 in Figure 4.

FIGURE 4.

Pattern of failing icons.



This visualizes icons with low correct scores in both countries. Another icon had a similar pattern of low scores in both countries, but higher than 15%:

7. Mental Health CM16 – USA 36% | Tan 30%.

Clearly, these seven icons did not communicate well in either culture. In total, 7 of 54 icons succeeded and 7 of 54 icons failed, leaving 40 icons with misunderstanding either due to knowledge or culture.

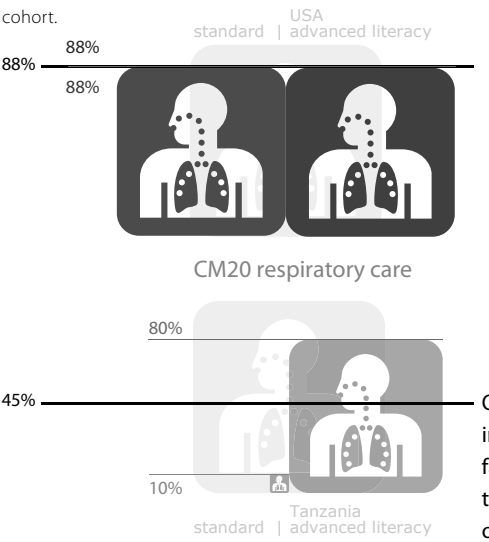
FAILING ONLY IN THE TANZANIAN 'STANDARD' COHORT

Differences in responses by Tanzanian's with 'standard' and 'advanced' literacy seemed to be driving most of the extreme results between countries (see *table 1*) with 'standard' subjects being unable to correctly comprehend many icons. For example, 'advanced' literacy subjects were 44 times more likely to

correctly comprehend the Kidney icon CM22 compared to 'standard' literacy [OR=44.3 (CI=4.8-410.9), $\chi^2 = 11.14$, p=0.0008], and 36 times more likely to understand the Respiratory Care icon CM20 compared to 'standard' literacy [OR=36.0 (CI=5.8-223.5), $\chi^2 = 14.79$, p=0.0001]. The visual pattern for icons succeeding with both cohorts in the USA and with the 'advanced' cohort in Tanzania, but failing with the 'standard' cohort in Tanzania, is shown by icon CM20 in Figure 5.

FIGURE 5.

Pattern of icons failing only in the Tanzanian 'standard' cohort.



Our hypothesis had suggested that the pattern in Figure 5 would be an indicator of an icon that failed due to lack of medical knowledge, not cultural difference. This pattern was seen at the level of significance in 12 of the 54 icons:


1. Pharmacy CM06 – USA 88% 89% | Tan 15% 60%


2. Family Practice CM08 – USA 74% 82% | Tan 10% 50%


3. Nutrition CM10 – USA 47% 62% | Tan 5% 35%


4. Respiratory Care CM20 – USA 88% 88% | Tan 10% 80%


5. Kidney CM22 – USA 88% 92% | Tan 30% 95%


6. Infectious Disease CM28 – USA 47% 50% | Tan 10% 50%


7. Dental CM29 – USA 100% 100% | Tan 45% 95%


8. Surgery CM31 – USA 83% 82% | Tan 0% 50%


9. Physical Therapy CM32 – USA 67% 82% | Tan 15% 55%


10. Emergency FA01 – USA 100% 100% | Tan 10% 75%


11. Chapel FA12 – USA 100% 94% | Tan 40% 80%


12. Ultrasound MA05 – USA 91% 100% | Tan 15% 70%.

The pattern was also apparent, but not to the level of statistical significance, in one additional icon:

13. Cardiology CM23 – USA 94% 100% | Tan 55% 75%

Following our hypothesis, we believe these 13 icons failed due to differences in knowledge, not due to differences in cultural. One disputable icon, Emergency FA01, will be discussed below.

There were 2 icons with an unusual pattern of greater success in Tanzania than the USA:

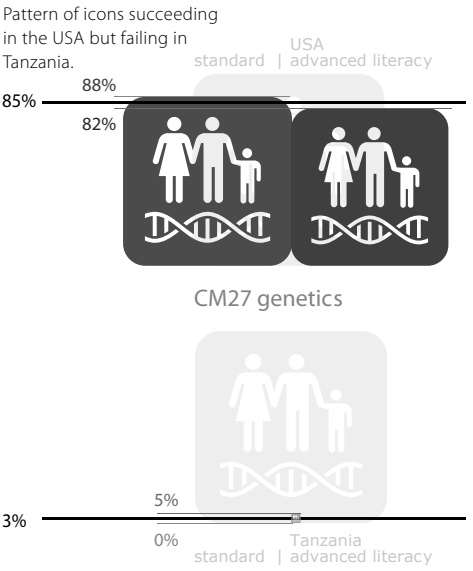
1. Pathology CM13 – USA 12% 28% | Tan 55% 95%
2. Dermatology CM18 – USA 36% 29% | Tan 5% 65%

For the Pathology icon CM13 residents of Tanzania were significantly more likely to comprehend compared to respondents in the US sample [OR=0.09 (CI=<0.001-0.3), $\chi^2 = 20.92$, p=<0.0001]. Due to the great disparity in knowledge in Tanzania, we considered these also to have failed due to knowledge disparity, for a total of 15 failing due to knowledge. In total, 7 icons succeeded, 7 failed, and 15 failed due to knowledge, leaving 25.

SUCCEEDING IN THE USA, FAILING IN TANZANIA

As noted in the hypothesis section, icons that succeeded equally in cohorts with both 'standard' and 'advanced' medical literacy in the USA, but that failed with both cohorts in Tanzania might have failed *either* due to lack of medical knowledge *or* due to cultural differences. Such differences in icon comprehension appeared when data from both 'standard' and 'advanced' medical literacy were pooled together. The visual pattern for icons succeeding in one country but not in the other is shown by icon CM27 in Figure 6.

FIGURE 6.



This pattern visualizes icons succeeding with both cohorts in the USA but failing with both cohorts in Tanzania. Results in Table 2 show the US sample significantly more likely to correctly comprehend icons compared to the Tanzanian sample. Significant differences were found on 38 of the 54 icons. Differences were especially extreme [OR=30 or higher] for 15 of 54 icons. Some of the most extreme examples of this were:

MRI MA07 – USA 96% | Tan 5%

[OR=361.0 (CI=48.3-999.9), $\chi^2 = 32.95$, p<0.0001]

Genetics CM27– USA 85% | Tan 3%

[OR=227.4 (CI=26.1-999.9), $\chi^2 = 24.13$, p<0.0001]

Mammography MA02 – USA 85% | Tan 3%

[OR=220.9 (CI=25.3-999.9), $\chi^2 = 23.90$, p<0.0001].

After removing the 14 icons that either succeeded or failed in both countries, and the 15 icons that failed due to knowledge there remain 25 icons that failed for *either* knowledge *or* cultural reasons:

1. Health Services CM01 – USA 75% 84% | Tan 60% 50%

2. Care Staff Area CM02 – USA 90% 80% | Tan 45% 50%

3. Intensive Care CM03 – USA 45% 50% | Tan 10% 30%

4. Inpatient CM04 – USA 25% 27% | Tan 0% 5%

5. Alternative Medicine CM11 – USA 77% 76% | Tan 5% 15%

6. Ear, Nose, Throat CM19 – USA 84% 65% | Tan 5% 35%

7. Women's Care CM24 – USA 88% 82% | Tan 10% 10%

8. Labor & Delivery CM25 – USA 100% 94% | Tan 70% 70%

9. Pediatrics CM26 – USA 83% 76% | Tan 0% 25%

10. Genetics CM27 – USA 88% 82% | Tan 0% 5%

11. Anesthesia CM30 – USA 50% 53% | Tan 5% 10%

12. Registration FA03 – USA 67% 71% | Tan 5% 25%

13. Waiting Area FA04 – USA 96% 82% | Tan 0% 25%

14. Medical Records FA06 – USA 100% 82% | Tan 5% 15%

15. Billing FA07 – USA 96% 82% | Tan 5% 25%

16. Medical Library FA08 – USA 26% 47% | Tan 0% 10%

17. Radiology MA01 – USA 100% 100% | Tan 0% 25%

18. Mammography MA02 – USA 82% 87% | Tan 0% 5%

19. Cath Lab MA03 – USA 30% 41% | Tan 0% 5%

20. MRI/PET MA04 – USA 41% 57% | Tan 5% 10%

21. Imaging MA06 – USA 78% 94% | Tan 10% 15%

22. MRI MA07 – USA 91% 100% | Tan 0% 10%

23. PET MA08 – USA 61% 100% | Tan 0% 5%

24. CT Imaging MA09 – USA 83% 100% | Tan 0% 35%

25. CAT Imaging MA10 – USA 74% 82% | Tan 0% 10%

TABLE 1. Within Country Differences - Tanzania

	Odds Ratio	Lower 95th CI	Upper 95th CI	chi square	Prob
CM01 (health services)	0.7	0.2	2.3	0.40	0.5257
CM02 (care staff area)	1.2	0.4	4.2	0.10	0.7516
CM03 (intensive care)	3.9	0.7	22.1	2.30	0.1297
CM04 (in-patient)	<0.001	<0.001	>999.9	0.00	0.9594
CM05 (outpatient)					
CM06 (pharmacy)	8.5	1.9	38.8	7.63	0.0058
CM07 (diabetes)	2.1	0.2	25.3	0.35	0.5557
CM08 (family practice)	9.0	1.6	49.4	6.40	0.0115
CM09 (immunization)	1.6	0.2	10.7	0.23	0.6347
CM10 (nutrition)	10.2	1.1	93.3	4.25	0.0393
CM11 (alternative med)	3.4	0.3	35.4	1.01	0.3142
CM12 (laboratory)	10.2	1.1	93.3	4.25	0.0393
CM13 (pathology)	15.5	1.7	139.6	6.00	0.0143
CM14 (oncology)	>999.9	<0.001	>999.9	0.00	0.9594
CM15 (opthamology eye)	6.3	0.7	60.2	2.58	0.1081
CM16 (mental health)	1.6	0.4	6.3	0.47	0.4917
CM17 (neurology)	4.0	1.0	16.3	3.75	0.0528
CM18 (dermatology)	35.3	3.9	321.9	9.98	0.0016
CM19 (eye, ear, nose)	8.1	0.9	75.5	3.41	0.0649
CM20 (respiratory care)	36.0	5.8	223.5	14.79	0.0001
CM21 (internal medicine)	4.5	1.2	17.4	4.76	0.0291
CM22 (kidney)	44.3	4.8	410.9	11.14	0.0008
CM23 (cardiology)	2.5	0.6	9.4	1.72	0.1896
CM24 (women's health)	1.0	0.1	7.9	0.00	1.0000
CM25 (labor & delivery)	1.0	0.3	3.9	0.00	1.0000
CM26 (pediatrics)	>999.9	<0.001	>999.9	0.01	0.9403
CM27 (genetics)	>999.9	<0.001	>999.9	0.00	0.9594
CM28 (infectious disease)	>999.9	<0.001	>999.9	0.00	0.9508
CM29 (dental)	23.2	2.6	208.5	7.88	0.0050
CM30 (anesthesia)	2.1	0.2	25.3	0.35	0.5557
CM31 (surgery)	>999.9	<0.001	>999.9	0.00	0.9452
CM32 (physical therapy)	6.9	1.5	31.4	6.30	0.0121
FA01 (emergency)	27.0	4.6	159.7	13.21	0.0003
FA02 (ambulance)	2.1	0.2	25.3	0.35	0.5557
FA03 (registration)	6.3	0.7	60.2	2.58	0.1081
FA04 (waiting area)	>999.9	<0.001	>999.9	0.01	0.9403
FA05 (administration)					
FA06 (medical records)	3.4	0.3	35.4	1.01	0.3142
FA07 (billing)	6.3	0.7	60.2	2.58	0.1081
FA08 (medical librbay)	>999.9	<0.001	>999.9	0.00	0.9594
FA09 (health edu)	4.6	1.0	21.0	3.96	0.0466
FA10 (interpreter serv)	>999.9	<0.001	>999.9	0.00	0.9594
FA11 (social services)	1.0	0.1	7.9	0.00	1.0000
FA12 (chapel)	6.0	1.5	24.7	6.16	0.0130
MA01 (rdiology X-ray)	>999.9	<0.001	>999.9	0.01	0.9403
MA02 (mammography)	>999.9	<0.001	>999.9	0.00	0.9594
MA03 (cath lab)	>999.9	<0.001	>999.9	0.00	0.9594
MA04 (MRI)	2.1	0.2	25.3	0.35	0.5557
MA05 (untrasound)	13.2	2.8	62.6	10.58	0.0011
MA06 (imaging X-ray)	1.6	0.2	10.7	0.23	0.6347
MA07 (MRI)	>999.9	<0.001	>999.9	0.01	0.9427
MA08 (PET)	>999.9	<0.001	>999.9	0.00	0.9594
MA09 (CT imaging)	>999.9	<0.001	>999.9	0.00	0.9538
MA10 (CAT scan)	>999.9	<0.001	>999.9	0.01	0.9427

In all, 25 of 54 icons fit in the pattern of equal success in cohorts with both 'standard' and 'advanced' medical literacy in the USA, but failure with both cohorts in Tanzania. Additional analysis of icon content below helped determine whether these failed due to lack of medical knowledge or due to cultural differences.

OTHER COMPARISONS

The comparisons above directly inform the hypothesis. To support these, additional comparisons for significance were performed to establish reliability thresholds. Comparisons by respondent type ('standard' vs. 'advanced' medical knowledge) were calculated by pooling data from both countries. While health-care professionals were significantly more likely to identify icons, the number of icons as well as the size of the odds ratios were far less than the country differences seen in Table 2. For example, the icon with the highest odds ratio was for Dental Services CM29 with health-care professionals being 12 times more likely to correctly identify this icon compared to patients in this sample [OR=12.0 (CI=1.5-98.0), χ^2=5.37, p=0.0204]. By comparison, the highest OR in country differences in Table 2 is 361.0 with 26 other icons having an OR higher than 12.0. The reason for the less extreme results may be due to the fact that there were no significant differences between patient and health-care professional in the US sample. These results were used to confirm the significance of differences in the following section.

Between countries contrast were calculated for both 'standard' as well as 'advanced' medical literacy. US 'standard' were more likely to correctly identify most of 'standard' of Care icons compared to their Tanzanian counterparts. No Tanzanian 'standard' literacy subjects correctly identified the icons for Outpatient Services CM05, Pediatrics CM26, and Infectious Disease CM28. Relatedly, US health-care professionals recognized icons at a significantly higher rate than the Tanzanian sample. This was especially apparent in the highly specialized services such as:

Radiology X-ray imaging MA01
[OR=90.7 (CI=8.5-964.0), χ^2=13.97, p=0.0002],
Mammography MA02
[OR=88.7 (CI=8.3-944.8), χ^2 = 13.8, p=0.0002], and
Genetics CM27
[OR=88.7 (CI=8.3-944.8), χ^2 = 13.8, p=0.0002].

Conversely, 'advanced' medical literacy subjects of Tanzania were significantly more likely to correctly identify the:

Pathology icon CM13
[OR=0.007 (CI=<0.001-0.1), χ^2 = 15.19, p=<0.0001]
and the
Dermatology icon CM18
[OR=0.20 (CI=0.1-0.9), χ^2 = 4.44, p=0.0351].

TABLE 2. Country Differences

	Odds Ratio	Lower 95th CI	Upper 95th CI	chi square	Prob
CM01 (health services)	3.4	1.3	8.5	6.50	0.0108
CM02 (care staff area)	6.9	2.5	19.1	14.13	0.0002
CM03 (intensive care)	3.6	1.4	9.2	6.85	0.0089
CM04 (in-patient)	13.7	1.7	110.0	6.07	0.0138
CM05 (outpatient)	>999.9	<0.001	>999.9	0.01	0.9362
CM06 (pharmacy)	13.0	4.2	40.2	19.80	<0.0001
CM07 (diabetes)	1.2	0.3	5.9	0.07	0.7925
CM08 (family practice)	7.9	3.0	21.1	17.26	<0.0001
CM09 (immunization)	0.5	0.2	1.8	1.03	0.3100
CM10 (nutrition)	5.1	1.9	13.5	10.47	0.0012
CM11 (alternative med)	29.7	8.5	103.9	28.18	<0.0001
CM12 (laboratory)	1.1	0.4	3.3	0.03	0.8721
CM13 (pathology)	0.09	0.0	0.3	20.92	<0.0001
CM14 (oncology)	3.0	0.3	30.1	0.87	0.3505
CM15 (opthamology eye)	1.7	0.4	6.4	0.57	0.4522
CM16 (mental health)	1.3	0.5	3.3	0.30	0.5824
CM17 (neurology)	1.3	0.5	3.4	0.39	0.5323
CM18 (dermatology)	0.9	0.4	2.3	0.03	0.8736
CM19 (eye, ear, nose)	15.1	5.1	44.5	24.19	<0.0001
CM20 (respiratory care)	9.0	2.9	27.8	14.78	0.0001
CM21 (internal medicine)	1.0	0.4	2.5	0.01	0.9247
CM22 (kidney)	5.5	1.6	18.7	7.65	0.0057
CM23 (cardiology)	21.5	2.7	173.8	8.30	0.0040
CM24 (women's health)	52.5	13.6	202.1	33.17	<0.0001
CM25 (labor & delivery)	17.1	2.1	139.5	7.06	0.0079
CM26 (pediatrics)	28.9	8.6	97.3	29.46	<0.0001
CM27 (genetics)	227.4	26.1	>999.9	24.13	<0.0001
CM28 (infectious disease)	3.8	1.4	10.2	7.05	0.0079
CM29 (dental)	>999.9	<0.001	>999.9	0.01	0.9389
CM30 (anesthesia)	12.9	3.4	48.8	14.32	0.0002
CM31 (surgery)	14.6	4.9	43.1	23.49	<0.0001
CM32 (physical therapy)	5.1	2.0	13.1	11.20	0.0008
FA01 (emergency)	>999.9	<0.001	>999.9	0.01	0.9457
FA02 (ambulance)	>999.9	<0.001	>999.9	0.004	0.9531
FA03 (registration)	12.2	4.1	36.3	20.27	<0.0001
FA04 (waiting area)	64.7	16.1	260.9	34.40	<0.0001
FA05 (administration)					
FA06 (medical records)	114.0	23.8	545.0	35.19	<0.0001
FA07 (billing)	52.4	13.6	201.8	33.13	<0.0001
FA08 (medical librbay)	21.0	2.6	169.5	8.16	0.0043
FA09 (health edu)	1.1	0.5	3.0	0.06	0.8049
FA10 (interpreter serv)	1.0	0.1	16.6	0.00	1.0000
FA11 (social services)	1.3	0.3	5.2	0.12	0.7240
FA12 (chapel)	26.0	3.2	208.8	9.39	0.0022
MA01 (rdiology X-ray)	>999.9	<0.001	>999.9	0.01	0.9336
MA02 (mammography)	220.9	25.3	>999.9	23.90	<0.0001
MA03 (cath lab)	21.0	2.6	169.5	8.16	0.0043
MA04 (MRI)	12.3	3.2	46.6	13.71	0.0002
MA05 (untrasound)	25.7	5.4	121.6	16.77	<0.0001
MA06 (imaging X-ray)	39.7	11.1	142.3	31.90	<0.0001
MA07 (MRI)	361.0	48.3	>999.9	32.95	<0.0001
MA08 (PET)	134.3	16.1	>999.9	20.50	<0.0001
MA09 (CT imaging)	42.4	11.4	158.2	31.15	<0.0001
MA10 (CAT scan)	65.4	13.2	325.4	26.11	<0.0001

ANALYSIS OF RESULTS

The number and significance of differences in icon comprehension between the US and Tanzania, and differences between cohorts in Tanzania, demonstrated that we were able to distinguish various reasons for icon failure, at least in regards to medical knowledge. As stated earlier, additional analysis was needed to help establish more precisely the causes of various levels of understanding, particularly in icons that succeeded in one country but not in the other.

SUCCEEDING ICONS

As noted, only 3 icons achieved an 85% correct comprehension level in both countries, and only 2 of the same icons achieved 85% success in both countries. So 52 of the 54 icons failed to perform at 85% across cultures. However, 7 of 54 icons fit a *general pattern* of success in both cultures. These 7 icons used familiar objects: a person getting a shot, a microscope, an eye, a brain, bowels, an emergency vehicle, a teacher.

FAILING ICONS

Four of the 7 failing icons were descriptions of a service activities: Diabetes Education CM07; Administration FA05, Interpreter services FA10; Social Services FA11, and one, Mental Health CM16, represented a state of being. We have discussed elsewhere that icons have difficulty communicating actions and states of being (Zender, 2006). This challenge can be overcome by a multi-frame pictogram or an animation, but these approaches were not a part of this icon system.

FAILING ONLY IN THE TANZANIAN 'STANDARD' COHORT: LACK OF KNOWLEDGE

The substantial number of icons (15 of 54) misunderstood due to lack of knowledge confirmed our hypothesis that domain knowledge was a driver of miscommunication across cultures and that this distinction could be identified and measured. Analysis of the referent concepts and the individual symbols used to represent them suggested two broad reasons for misunderstanding.

Eight of the 15 icon referents in this category were related to established medical specialties: Family Practice CM08, Dermatology CM18, Respiratory Care CM20, Kidney CM22, Cardiology CM23, Dental CM29, Surgery CM31 and Physical Therapy CM32. We suspected that because Tanzania had few doctors and hospitals that these medical specialties were not common knowledge but were known to medical professionals. The marginally significant Cardiology icon was instructive in this regard. The difference between 'standard' and 'advanced' literacy for the Cardiology icon was not as pronounced [OR=2.5 (CI=0.6-9.4), x^2=1.72, p=0.1896] as the other specializations such as Kidney [OR=44.3 (CI=4.8-410.9), χ^2=11.14, p=0.0008].

We suspected that Cardiology might be one of the first medical specializations to emerge in a developing medical culture and therefore more widely known than other specializations.

Six of these 15 icon referents were related to concepts that were likely to be part of advanced medical training: Pharmacy CM06, Nutrition CM10, Pathology CM13, Infectious Disease CM28, Emergency FA01, and Ultrasound MA05. More specific analysis of the individual symbols used in the icons revealed that the Pharmacy icon CM06 combined a symbol of a pill bottle with a typographic sign for prescription, the "Rx." This icon was comprehended equally well by those with and without 'advanced' medical knowledge in the USA: 56% and 44% respectively, while in Tanzania there was a significant difference in comprehension: 60% correct by those with 'advanced' medical knowledge but only 15% correct by those with 'standard' medical knowledge. We concluded that in this case Tanzanian 'advanced' medical knowledge subjects understood the Rx sign due to medical training, but that this knowledge was not common for other Tanzanians. Dowse found the same issue with the Rx sign in South Africa (R. Dowse & Ehlers, 2001, p. 91). Detailed analysis of the incorrect text answers was instructive for the Emergency icon FA01 as well. Most of the incorrect answers said simply "cross" but a substantial number elaborated by saying "a cross with some writings (sic)." This suggested a lack of comprehension because the word Emergency was in English leaving non-English speakers out. We believed because the key difference was the word/sign "EMERGENCY" that the icon was understood with 'advanced' medical knowledge subjects not so much because of medical knowledge but because their English reading ability was higher. Hence we did NOT consider it a lack of medical knowledge despite its fit to the pattern and instead counted it as a failure due to reliance on a learned sign: the word "EMERGENCY," which we associated with cultural difference below.

The remaining icon that failed due to knowledge, according to our pattern for interpretation, was Chapel FA12. The majority of the correct answers responded to the question "what does it mean" with the answer "church" which was an acceptable synonym for chapel on the scoring sheet, but for the question 'what actions would you take,' most responded "I don't know." This suggested that 'standard' respondents were familiar with a church but unfamiliar with a church in a medical context. This, we believe, accounted for the number of wrong answers. In summary, 14 of these 15 icons failed due to knowledge disparities.

SUCCEEDING IN THE USA, FAILING IN TANZANIA: KNOWLEDGE OR CULTURE?

As stated previously, additional analysis of icon content helped determine whether the 25 of 54 icons that succeeded in the USA but failed in Tanzania failed due to lack of medical knowledge or due to cultural differences.

Content analysis showed that 18 of the 25 fell into two categories: 9 were 'advanced' technical imaging technologies, 9 were hospital services. Indeed, all but one of the technical imaging related icons, MA01 through MA10, had an odds ratio between 21 and 361, meaning USA respondents were between 21 and 360 times more likely to correctly identify imaging icons: Radiology MA01, Mammography MA02, Cath Lab MA03, MRI/PET MA04, Imaging MA06, MRI MA07, PET MA08, CT MA09, and CAT MA10. Only Ultrasound in this category was not included because it was part of the category of failure only with 'standard' literacy subjects in Tanzania. Ultrasound comprehension could have been influenced by the presence of a portable ultrasound training team in Shirati Hospital at the time of the first survey. We believed the 9 imaging icons (excluding Ultrasound) failed because neither 'standard' nor 'advanced' subjects were familiar with these technologies, not because of cultural misunderstanding.

The other large category comprised 9 hospital services: Health Services CM01, Care Staff Area CM02, Intensive Care CM03, Inpatient CM04, Registration FA03, Waiting Room FA04, Medical Records FA06, Billing FA07, and Medical Library FA08. Content analysis alone was unhelpful so we looked for patterns in the written answers for these icons. The responses for the Health Services icon CM01 showed great consistency around "a cross, a church." In Tanzania a cross was clearly associated more with church than health care. Prevalence and type of with religious practice would qualify this as a cultural difference. For the icon Billing FA07 it was instructive that many people responded "two people standing" but no one specifically responded "dollar" or "money". This suggested that the dollar sign was unfamiliar in Tanzanian culture and the cause of the icon failure. For the Care Staff Area icon CM02 respondents consistently wrote "nurse," which was of course a literal description of the symbol. Apparently care staff areas were not prevalent in Tanzania, a failure of knowledge. For the Registration icon FA03 several people answered "two people writing/talking" or "a patient getting treatment." Because the individual symbols were understood but the combination was not, we interpreted this as a lack of familiarity with this form of registration. For Waiting Room FA04 many respondents answered, "a picture of a person seated reading" or something similar. This suggested to us that sitting and reading were not equivalent to Waiting Room in Tanzanian hospitals. The Medical Records icon FA06 also had many similar wrong answers such as "a first aid box in a file cabinet." The icon for Medical Library FA08 had a similar pattern of wrong answers that were literal descriptions of the individual symbols, "a person reading and a cross," that did not add up to the correct conclusion: Medical Library. We found many cases where incorrect answers described the individual symbols the subjects observed in the icons: "person reading and a cross," but whose individual symbols did not lead to the intended abstract concept. This way of answering was pronounced for the failed Outpatient icon CM05 where most incorrect answers were something like "man walking with broken arm." Answers with

a concrete description instead of an abstract concept may give insight into the icon decoding process. This is corroborated by Beaufils et. al. in their study of cognitive processes used to comprehend icons in that they scored lower – less correct - for participants' answers that gave a concrete (literal) meaning to symbols and scored higher – more correct - for a given abstract meaning. (Beaufils et al., 2014)

Icons for Intensive Care CM03 and Inpatient CM04 were not particularly well understood in either culture and probably should be included in the failing ions category. They were included here because the difference between USA and Tanzanian comprehension was significant.

Two of the icons with extreme differences in comprehension in the USA compared to Tanzania could be considered highly technical, such as Genetics CM27 and Anesthesia CM30, and thus failures due to knowledge. Similarly, Alternative Medicine CM11, Ear, Nose, Throat CM19, Women's Health CM24, Labor & Delivery CM25, and Pediatrics CM26 could be considered medical specializations and failures due to knowledge. However, while Ear Nose &Throat CM19 seemed to be a straightforward example of a medical specialization and thus a knowledge failure, closer analysis of answers suggested a more nuanced explanations for the others. For the Alternative Medicine icon CM11 most answers described the examination of a patient on a bed. The needles were not mentioned. This may suggest that homeopathic or 'alternative' remedies were so familiar in Tanzania they were not labeled alternative! If true, 'alternative' in one culture may not be 'alternative' in another culture. However, the evidence was unclear whether a different view of technology was in play here or not so we left it attributed to knowledge disparity. The Women's Health icon CM24 combined a symbol of a woman with the sign for female (a circle with a plus sign). This icon was well understood by both groups in the USA but was not in Tanzania. Comparing the Women's Health icon to the similar Labor & Delivery icon helped clarify the reason for the miscommunication. The Labor & Delivery icon CM25 combined the same a woman symbol as CM24 but with the symbol of a baby placed in the woman's belly instead of the circle woman sign. The Labor & Delivery icon was comprehended well in Tanzania at 70% correct, compared to Women's Health at 10%. We concluded therefore that the learned sign for woman was the cause of the miscommunication in Women's Health. The Pediatrics icon CM26 answers revealed a different problem. The bear was frequently called a "cat," and never called a "bear." While it was twice called a "toy" and twice called "an idol," it was clear that there were no bears roaming Africa at the time of this study and that stuffed bears were not children's toys. This icon was misunderstood because it used an animal and a metaphor non-existent in Tanzania: a cultural failure.

In summary, of the 25 icons that succeeded in the USA but failed in Tanzania, 9 failed due to imaging knowledge; 5 of 9 hospital services icons failed due to knowledge, 2 failed generally, and 2 failed due to culture; 2 technical icons failed due to knowledge; and 3 of

5 specialization icons failed due knowledge and 2 failed due to culture. In total, 5 of these icons failed due to cultural misunderstanding.

Limitation

The study did not collect educational level other than presence or absence of advanced medical training.

DISCUSSION: SUMMARY OF KEY FINDINGS

The study demonstrated that different causes of cultural icon misunderstanding can be identified: 33 failed due to knowledge, and 5 failed due to culture.

KNOWLEDGE FAILURES

The 33 icons that failed due to knowledge disparities were grouped in three categories: 1. medical technologies including imaging, 2. medical knowledge including practice specializations and training experiences, and 3. medical operations and services. Icons of unknown technologies, medical specializations, and hospital services cannot be effective until the nature and devices associated with those things having first been known.

Others have noted that education level played a role in medical icon comprehension and our study confirmed that. (Ros Dowse & Ehlers, 2003) But whereas Kassam quoting Dowse suggested that education was important to comprehension because it built visual literacy, this study has suggested that more education may be important because it transmits more domain knowledge. This was born out by our findings that for 7 icons using familiar objects, scores in Tanzania were nearly equal for those with standard and advanced levels of knowledge, but for 15 icons of medical specialties or technical objects standard and advanced scores in Tanzania were significantly different. Educational level alone does not account for this. Others have connected level of general education to level of health literacy (Kickbusch, 2001). General skills and knowledge of medical practices, not a special visual literacy, are the likely drivers of improved icon understanding due to education that others have observed. Dowse seemed to suggest this, saying, "Every single respondent with less than 5 years schooling displayed extremely poor comprehension of medication information." (R. Dowse & Ehlers, 2001, p. 91) It is hard to imagine that those with low or no reading literacy, and who are therefore almost totally dependent on reading visual images and symbols for their daily survival, have low visual literacy. Based on interviews with three subjects in Tanzania we would argue that people with low education are highly skilled at reading images and that lack of medical knowledge, not low visual literacy, was the driver of much observed icon misunderstanding.

Others have also observed that both knowledge and culture are two key actors involved in interpretation of icons (Beaufils et al., 2014). This study confirmed that and demonstrated that these two can be distinguished and measured.

Finally, others have noted that one of the key difficulties interpreting symbols was that some have an arbitrary connection to their referent (Beaufils et al., 2014). This paper defined a taxonomy that connected image function to image resemblance: sign = no resemblance; icon = simple resemblance; picture = specific resemblance. The taxonomy noted that signs must be learned. Learned signs are not stable across cultures that use different languages and signs because a sign's meaning is not inherent in its visual form: no resemblance. Our taxonomy is a continuum, not fixed points. An image can move from one category to another as it is used and adapted over time. For example, the cross was originally an icon of Roman capital punishment, then a sign of Christian salvation through Jesus' death, then a medical sign of life saving treatment. It evolved from icon to sign to sign and took on new meanings. We scored a Tanzanian response to a cross sign "incorrect" because the response was "Church" not "Health Services." However, this misunderstanding was as much a statement about intention of the icon designer who intended the cross sign to mean health care not church, as about the respondent who was not privy to the designer's intention. A symbol functioning as an icon is visually related to what it resembles, but a symbol functioning as a sign is fixed only by cultural convention.

CULTURAL FAILURES

Detailed analysis of written answers and symbol content exposed 5 icons that appeared to have failed for identifiable cultural reasons. We grouped these into two general causes: use of metaphor, and use of learned signs.

Use of Metaphor

The Pediatrics icon CM26 combines a symbol of a bear with a cross. The bear symbol was not literally representative of children, but was used as a metaphor for children and combined with the cross sign to represent medical. The Pediatrics icon was comprehended 28 times better in the USA than in Tanzania. In Tanzania not one person with 'standard' medical literacy comprehended this icon correctly. However, medical care for children was a familiar concept in Tanzania. As pictured above, the children's ward was very active with its own laundry space for parents. Yet the Pediatric icon failed. Post survey interviews with subjects confirmed that Tanzanians did not give stuffed bears as toys to their children. We interpreted this as an example of cultural miscommunication based on use of metaphor.

Use of Learned Sign

Introducing a learned sign into the Woman's Health icon CM24 created the potential for miscommunication where the sign had not been learned. This

was similar to the cause of failure in the Billing icon FA07 where the learned dollar sign was not understood, and the Emergency icon FA01 where the word was not understood. For the Health Services icon CM01 the cross sign was (mis)understood as a church, not a hospital.

SIGNIFICANCE

Icons are often misunderstood when applied across cultures. This study has shown that lack of medical domain knowledge was a key driver of this misunderstanding for this icon set. Most (33 of 47) of the icons that failed in this study failed for this reason. Just 5 failed due to broader cultural differences.

Many have called on designers to be more sensitive to the cultural context of their work (Ros Dowse & Ehlers, 2004; Grenier et al., 2011; Kassam et al., 2004). More precise understanding of the causes of cultural misunderstanding is one step in this direction. The following principles for cross-cultural icon design may be use to respond to this.

Learn What They Know

Learn how familiar the target culture is with the concepts to be communicated, particularly where technology and specialized knowledge are involved. Concepts and objects that are absent the cultural consciousness will have to be introduced by scaffolding them onto familiar concepts.

Cultural Metaphor

Recognize when a proposed icon employs metaphor to communicate, and when it does, check to see if the metaphor is present in the target culture.

Learned Signs

Avoid learned signs. This includes words and other learned symbols such as Rx and the sign for female. If signs can't be avoided, they may be disambiguated as part of a system of icons that explains the unfamiliar sign.

Identify and Redesign Icons for Difficult Referents

Use an iterative, participatory design and testing process to improve icons that are simply failures of design.

In 1978 James Mangan articulated a list of steps similar to the one above for creating effective cross-cultural images (Mangan, 1978, pp. 265-266). This study has identified domain knowledge as a dominant reason for cultural (mis)understanding. This knowledge sharpens designers focus onto the main reasons for icon misunderstanding and the suggested means to address it.

ACKNOWLEDGEMENTS

Thanks to Village Life Outreach Project for logistical support of the studies in Tanzania, and special thanks to study administrator Maren Carpenter-Fearing who administered the second cohorts of subjects in Tanzania; to Keith Crutcher for advice on the manuscript; to the many participants in the study who gave freely of their time; and to graduate design students in course DSGN7042 who administered the surveys in the USA.

CITED WORKS

American National Standard Criteria for Safety Symbols, Z535.3 C.F.R. (2007).

Beaufils, E., Hommet, C., Brault, F., Marque, A., Eudo, C., Vierron, E. Mondon, K. (2014). "The effect of age and educational level on the cognitive processes used to comprehend the meaning of pictograms." *Aging Clinical And Experimental Research, 26*(1), 61-65. doi: 10.1007/s40520-013-0179-6

Dowse, R., & Ehlers, M. S. (2001). "The evaluation of pharmaceutical pictograms in a low-literate South African population." *Patient Education and Counseling, 45*(2), 87-99. doi: 10.1016/S0738-3991(00)00197-X

Dowse, Ros, & Ehlers, Martha. (2004). "Pictograms for conveying medicine instructions: comprehension in various South African language groups." *South African Journal of Science, 100* (November/December 2004), 687-693.

Dowse, Ros, & Ehlers, Martina Sabrina. (2003). "The influence of education on the interpretation of pharmaceutical pictograms for communicating medicine instructions." *International Journal of Pharmacy Practice, 11*(1), 11-18.

Grenier, Sylvain, Vaillancourt, Régis, Pynn, Debra, Cloutier, Michel C., Wade, Julie, Turpin, Pierre Marc, Preston, Cindy. (2011). "Design and development of culture-specific pictograms for the labelling of medication for first nation communities." *Journal of Communication In Health care, 4*(4), 238-245. doi: 10.1179/1753807611Y.0000000007

Graphical symbols - *Test methods Part 1: Methods for testing comprehensibility*, 9186-1 C.F.R. (2007).

Kassam, Rosemin, Vaillancourt, LCol Régis, & Collins, John B. (2004). "Pictographic instructions for medications: do different cultures inter-

pret them accurately?" *International Journal of Pharmacy Practice, 12*, 199-209.

Kickbusch, I. S. (2001). "Health literacy: addressing the health and education divide." *Health Promotion International, 16*(3), 289-297. doi: 10.1093/heapro/16.3.289

Knapp, Peter, Raynor, David K., Jebar, Abel H., & Price, Sarah J. (2005). "Interpretation of Medication Pictograms by Adults in the UK." *The Annals of Pharmacotherapy, 39*(July/August), 1227-1233.

Lesch, Mary F., Powell, W. Ryan, Horrey, William J., & Wogalter, Michael S. (2013). "The use of contextual clues to improve warning symbol comprehension: making the connection for older adults." *Ergonomics, 56*(8), 1264-1279.

Mangan, James. (1978). "Cultural Conventions of Pictorial Representation: Iconic Literacy and Education." *Educational Communication and Technology, 26*(3), 245-267.

Woodyard, Chris. (2010, Aug 03, 2010). "One third of drivers can't recognize this idiot light." *USA Today*.

Zender, Mike. (2006). "Advancing Icon Design for Global Non Verbal Communication: Or What Doest the Word Bow Mean?" *Visible Language, 40*(2), 177-206.

Zender, Mike, Han, Bing, & Fernández, Oscar. (2011). "*Did it work? Evaluative research methods for graphic design*." Paper presented at the IASDR 2011, Delft, Netherlands.

Zender, Mike, & Mejia, G. Mauricio. (2013). "Improving Icon Design: through focus on the role of individual symbols in the construction of meaning." *Visible Language, 47*(1), 66-89.

ABOUT THE AUTHORS

Mike Zender is Professor of Design at the University of Cincinnati. Professor Zender received his MFA (terminal degree) in 1977 from Yale University and In 2004 he was a Medical Informatics Course Fellow at the Marine Biology Laboratory, Woods Hole, MA. Mike's research on communication through simple symbols such as icons and pictograms, specifically in the medical domain, and the application and testing of these in global cross-cultural communication, is uncovering principles for communication design and design research.

Amy Cassedy is a Research Associate in the Division of Biostatistics and Epidemiology at Cincinnati Children's Hospital Medical Center and an Adjunct Assistant Professor in the Department of Sociology at the University of Cincinnati.

JOURNAL INFORMATION

Visible Language is an academic journal focused on research in visual communication. We invite articles from all disciplines that concern visual communication that would be of interest to designers.

READERSHIP

Visible Language, an academic journal, seeks to advance research and scholarship for two types of readers: academics and professionals. The academic is motivated to consume knowledge in order to advance knowledge thorough research and teaching. The professional is motivated to consume and apply knowledge to improve practice. Visible Language seeks to be highly academic without being inaccessible. To the extent possible given your topic, Visible Language seeks articles written to be accessible to both our reader types. Anyone interested may request a copy of our editorial guidelines for authors.

EDITORIAL CORRESPONDENCE

Article concepts, manuscripts, inquiries about research and other contributions to the journal should be addressed to the editor. We encourage article concepts written as an extended abstract of 1 to 2 pages single-spaced. We will offer prompt feedback on article concepts with our initial opinion on their suitability for the journal. Manuscripts accepted for peer review will receive a summary response of questions or comments within three weeks. Letters to the editor are welcome. Your response — and the author's reply — will not be published without your permission and your approval of any editing. If you are interested in submitting an article to the journal and would like a copy of our Notes on the Preparation of a Manuscript, please obtain it from the journal's website at http://visiblelanguagejournal.com
Editorial correspondence should be addressed to:

Mike Zender
Editor, Visible Language
College of Design, Architecture, Art, and Planning
School of Design
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
email: mike.zender@uc.edu

If you are interested in serving as guest editor for a special issue devoted to your specific research interest, write to the editor, outlining the general ideas you have in mind and listing a half dozen or so topics and possible authors. If you would rather discuss the idea first, call the editor at: phone number

BUSINESS CORRESPONDENCE

Subscriptions, advertising and related matters should be addressed to:

Visible Language
Sheri Cottingim
Office of Business Affairs
College of Design, Architecture, Art, and Planning
University of Cincinnati
PO Box 210016
Cincinnati, OH 45221-0016
telephone 513 556-4377
email: sheri.cottingim@uc.edu