Visible Language **39.2**

Dynamic Visual Formation

Isabel Meirelles

Abstract

Computational media bring new complexities to the visual realm and the creation of visual forms. With the objective of examining theoretically and experimentally the creative process of image-making in the computer environment, a "system of dynamic visual formation" is proposed. The central argument is that images are no longer fixed, unique and eternal. Rather, what is created in computational media is a variable spatio-temporal module. Spatial and temporal properties of the system are defined. Ultimately, the investigation searches for the most elemental constituents of dynamic visual formation moving towards a theory of dynamic visual language.

Art does not reproduce the visible but makes visible.
Paul Klee, *The Thinking Eye*

It is common sense that every new visual environment, every new medium, requires a different approach to the creation of visual elements. The forces at play are different for each case and thus demand new ways of conceiving in the visual realm. What is new in the creative process of image-making in computational media? Does Visual Language—and the basic elements as traditionally taught—suffice for the creation of visual forms in interactive media? And if not, what would the basic elements be? What and how can we to teach visual language both for and in a dynamic environment?

Visual language

Point, line and plane (and volume for 3D environments) are considered the basic elements of visual language—and of Geometry—and have been discussed in most books on visual language since the Bauhaus courses in the '20s. Among the sources used are Kandinsky's *Point and Line to Plane,* Klee's *The Thinking Eye* and Dondis' *A Primer of Visual Literacy.* However, the distinction between point, line and plane is no longer necessary or even valid in a dynamic environment, where the spatial structure is a process which changes in time. In other words, each point, line and plane is now one of many states of a "dynamic visual formation."

Traditionally, a visual element is described by seven basic attributes: shape, scale, orientation, position, tone, color and texture. The relations among the attributes create inner and outer qualities of elements, which provoke meaning-making. In the static visual world the seven attributes suffice for the creation of spatial structures and even the indication or inducement of spatial and temporal relations. Works of art by Kandinsky, Klee, Vasarely and Soto are examples. However, in a dynamic environment the attributes as such are not enough. The dimension of time must be incorporated in such a way that space and time can no longer be isolated.

Dynamic media

Computational media are of a different nature and require a different approach to the creation of visual elements. Essential to the creation of visual elements in the computer environment is previous knowledge of certain fundamentals:

• "traditional" visual language used in the creation of static spatial structures;
• "temporal" visual language used in the creation of spatio-temporal structures (such as in films);
• perception of visual forms whether static or temporal.

But these fundamentals are not enough. There is a need for new modes of visual sensitivity and conceptualization, as well as new modes of perception and creation of visual forms. With the objective of examining theoretically and experimentally the creative process of image-making in the computer environment, a "system of dynamic visual formation" is proposed. The main argument is that images are no longer fixed, unique and eternal. Rather, what is created in computational media is a variable spatio-temporal module.

The principle governing the proposed system is that the paradigm of the creative process of image-making has changed. This shift has its origin in the 1920s. Four concepts are central to it:
• spatial and temporal relationships within works of art;
• transformable qualities of works of art;
• participation of the viewer;
• relationships with other arts and with scientific fields of knowledge.

Although listed in isolated form, the concepts are interconnected and in many cases interdependent. Works of art which, for example, incorporate motion (optical or actual) already involve the concept of change and thus transformation. Similarly, many works inciting psychological movement require the active participation of the viewer in order for the perception of motion to be felt.

The concepts emerged in the art scene in Europe and Russia in the beginning of the 20th century and were largely explored in the '50s and '60s, when the term "kinetic art" was then coined and accepted by the artistic community (Popper, 1968, 95). It is also relevant to point out that those works undermined the traditional division of painting and sculpture.

Also taken into account are the five principles of New Media Language proposed by Lev Manovich (2001): numerical representation, modularity, automation, variability and cultural transcoding.

Like the inquiries on visual form carried out by Kandinsky and Klee in the beginning of the 20th century this investigation searches for the most elemental constituents of "dynamic visual formation."

It is relevant to point out that, since this is new territory, there was a need to create new vocabulary that would name and describe the proposed system and its constituents.

The system of dynamic visual formation

In order to explain the proposed system it is first necessary to consider the difference between "visual form" and "visual formation."

A visual form is a stable spatial structure. It is a time-independent spatial whole. Because there is no change with time, it is described only by spatial parameters.

Visual formation engages the spatiality of visual form with a temporal dimension. It is time-dependent in that it changes in time, such that later parts are dependent on earlier ones in the continuous process of formation. Its dimensions of time and space cannot be isolated.

What is proposed is a "dynamic visual formation." The term "dynamic" indicates the possibility of modifying the process already changing in time. In this sense, it is not a fixed process. Rather it is a dynamic, an ever changing spatio-temporal whole: it is always in the course of becoming, of forming and trans-forming.

The fundamentals of the system of dynamic visual formation comprise the study of two processes *(figure 1):*
1) The system of "inter-actions": the exchange of spatial and temporal information by two agents—an "active subject" and a programmed system.
2) The dynamic visual formation: the properties of visual attributes and the basic element—"rhythmic unit."

set of rules

spatial + temporal information:
properties of visual attributes

rhythmic unit

active subject   programmed system
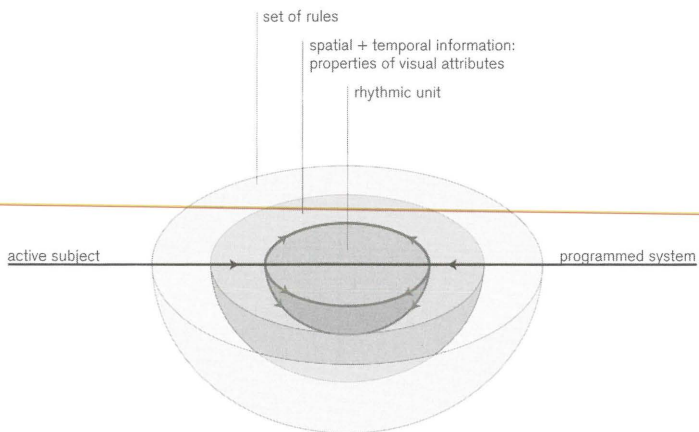
Figure 1. Diagram of the System of Dynamic Visual Formation.

System of inter-actions

A dynamic visual formation is created and re-created in the process of exchanging information between two agents.

Two conditions are necessary for the exchange of information. The first condition is the existence of two independent and active agents, one of which is a programmed system, and the other an active subject. By active subject I mean a person who is intellectually and physically engaged in the process of exchanging information. The subject is not an observer, but the creator of his or her own visual experience. One agent requires the existence of the other. The second condition is the mutual interdependence of the actions of the two agents. The exchange of information fulfilling the conditions will be called "inter-actions."

A person (e.g., artist, designer, programmer), or a group of people, develops a program that circumscribes the possibilities for inter-actions by means of which visual formations can emerge. The program that constitutes the system of inter-actions–defines "what," "how" and "how much" spatial and temporal information to exchange in the process of dynamic visual formation. It is limited by a set of rules organized in three categories:

1) Rules of Formation: The Content
The rules of formation define what information is given a priori and what information to exchange in the creative process.

2) Rules of Action: The Methods
The rules of action define how information is exchanged, including the active mechanisms of exchange (the ways of inter-acting), where they are active, and how to activate them for each pair of inter-actions defined by the rules of formation.

3) Rules of Influence: The Quantities
The rules of influence define how much information is exchanged. These quantities specify the extent of spatial and temporal information for each pair of inter-actions covered by the rules of action.

Dynamic visual formation

The process of dynamic visual formation is described by two interdependent components. One is the most elemental constituent in the process of dynamic visual formation: the properties of visual attributes of the basic element. And the second is the basic element: the variable spatio-temporal module called "rhythmic unit."

The properties of visual attributes

The properties of visual attributes are what constitutes the information being exchanged in the system of inter-actions, in other words, what creates the basic element.

What is proposed is a set of properties for each of the seven basic attributes (shape, scale, orientation, position, tone, color and texture). Properties are variables that function as independent data settings. It is by means of inter-actions that data values are set. Data values are numerical. Properties are subjected to the rules of usage: "what," "how" and "how much" information to exchange in the process.

The properties are grouped in three separate but interdependent categories: spatial, temporal and kinetic. Changes of value of one property affect the value of other property or properties in the two other categories (see examples below). The properties proposed are (detailed descriptions below):

- **spatial:** origin;
- **temporal:** starting point and duration;
- **kinetic:** velocity, amplitude and reference point.

Although we are dealing with the same attributes as traditionally in visual language, because now they have properties in spatial and temporal dimensions, each attribute creates a "rhythm" which defines the role it plays in the basic element.

Since the system attempts to search and, at the same time, to explore the most elemen-tal constituents of dynamic visual formation, the "loop" is used as its organizing prin-ciple. In this sense, the attribute's rhythm is a loop, a compound of a cycle and an interval *(figure 2)*. The Cycle is a periodic recurrence of the relationship among all properties of the three categories of the attribute. And the Interval is the period of time between the recurrence of the attribute's cycle: it is the amount of time set for the cycle to initi-ate again from its Starting Point. For example, if the value for interval is zero what is experienced is a continuous repetition of the cycle, where the Starting Point looses in significance. The rhythm of the attribute is equal to the cycle. If the value is 1 second, the perception of the rhythm is that of a cycle which repeats every second. In other words, there is a rest, a pause of 1 second at the Starting Point of the cycle *(figure 3)*.

attribute's rhythm						
cycle						interval
spatial category	kinetic category			temporal category		
origin	velocity	amplitude	reference point	starting point	duration	

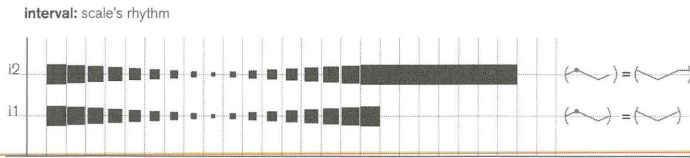Figure 2. Configuration of an attribute's rhythm

interval: scale's rhythm

Figure 3. Diagrammatic representation of two scale rhythms with different values
for interval: i1=0 and i2=x.

Spatial property: origin

The spatial property of Origin is related to the dimensions of space. It describes the spatial qualities of attributes. Thus, Origin is defined by the same parameters which are used to define the seven attributes in the static world of visual representation. For example, the attribute of position is described by values in the "x" and "y" coordinates (2D environment).

Origin is a variable and it defines the spatial quality at zero point in time of the cycle of that attribute. Not only can its value be changed by means of inter-actions, but most importantly, the value is only one state of many in the formation and trans-formation process of the rhythm. In other words, unlike in a static visual element, spatial parameters are no longer fixed *(figure 4)*.

origin for attribute of tone: scale rhythm
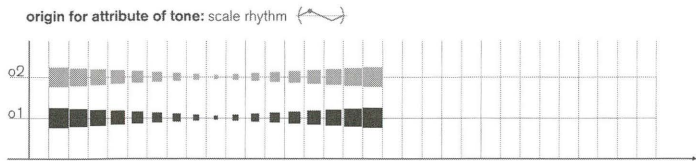
Figure 4. Diagrammatic representation of two scale rhythms with different values for Tone origin: o1 = 100% and o2 = 50%. The values for all other attributes (scale, shape, etc.) are the same.

Temporal Properties: Starting Point and Duration

Temporal properties are related to the dimension of time. They define the temporal qualities of attributes. Two temporal properties are proposed: Starting Point and Duration.

Starting Point is a variable and defines the moment in time for the cycle to initiate. In other words, it defines the zero point in time of the cycle. There are many possible points in the cycle which can be set as the Starting Point, of which four are defined in the system *(figure 5)*. It is relevant to point out that the cycle does not differ by changing the value of Starting Point. This is an example of inter-actions which affect the formation process without modifying its structural whole. In other words, it is the perception of the rhythm which is affected. However, if the interval is larger than zero, changes of values alter both the perception and the visual quality of the entire rhythm, where the Starting Point plays a major role *(figure 6)*.

Duration is a variable that describes the period of time for the completion of the cycle.

starting point: scale's rhythm
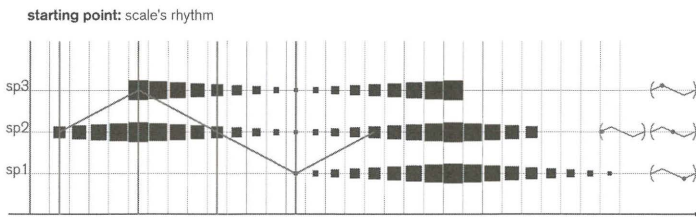
Figure 5. Diagrammatic representation of three scale rhythms showing the four values for Starting Point defined in the system.

starting point: scale's rhythm
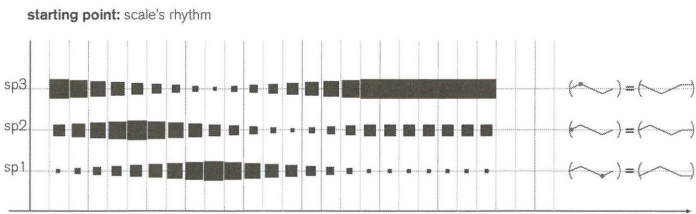
Figure 6. Diagrammatic representation of three scale rhythms with intervals larger than zero and different values for Starting Point.

Kinetic properties: Velocity, Amplitude and Reference Point

Kinetic properties define the spatio-temporal dependent qualities of attributes. In other words, the properties in which the dimensions of space and time are inseparable. Three properties are proposed: Velocity, Amplitude and Reference Point.

Velocity is the speed and direction of the process of change of an attribute. It is a variable that defines the rate at which the attribute's formation occurs *(figure 7)*. Because the motion is cyclical it always happens in opposite directions. The starting direction depends on the value of Starting Point (temporal property; also see *figure 5*).

Amplitude is a variable that defines the extent of the process of change of an attribute. Depending on the value of Starting Point (temporal property), the value for Amplitude might set a maximum value or a minimum value or both *(figure 8)*.

Reference Point is the point in the spatial structure in relation to which the motion refers. It is a variable that defines the anchor for the attribute's motion. Any point in the spatial structure can be set to be the Reference Point. In the present system five points are assigned, which are related to Shape Origin (spatial property) set to a square: four corners and the center *(figure 9)*. For example, if the value is set to the upper left corner for the attribute of orientation, the square will rotate around that corner.

velocity: scale's rhythm
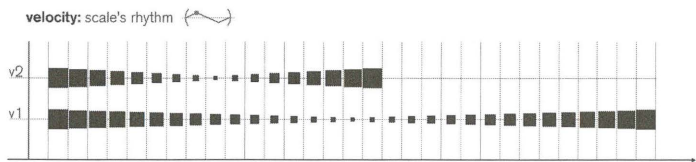
Figure 7. Diagrammatic representation of two scale rhythms with different values for Velocity:
v1 < v2

amplitude: scale's rhythm

Figure 8. Diagrammatic representation of two scale rhythms with different values for Amplitude.
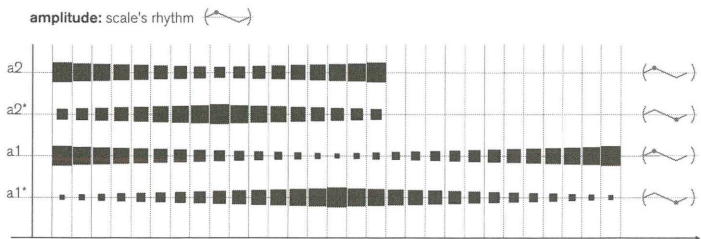a1 and a2 determine minimum values; a1* and a2* maximum values.
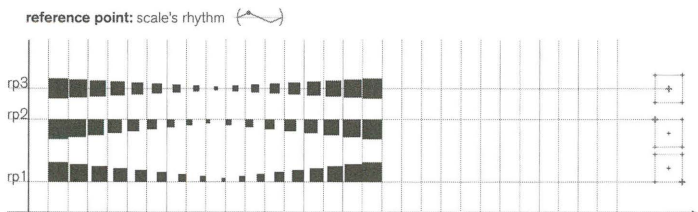
reference point: scale's rhythm

Figure 9. Diagrammatic representation of three scale rhythms with different values
for Reference Point.

The basic element—"rhythmic unit"
What is proposed as the basic element is a "rhythmic unit." The rhythmic unit is a variable spatio-temporal module. As described earlier, it is created and recreated by means of a system of inter-actions. The rhythmic unit is characterized by a rhythm (again a "loop") which is the visual formation process. The rhythm is a compound of a **rhythmic cycle** and an **interval.**

The rhythm of the rhythmic unit is complex. Unlike a static visual element, inner and outer qualities of a dynamic visual formation are created in a dual system of relationships *(figure10).* The first is the internal relationship among the spatial, temporal and kinetic properties of each attribute, which makes each attribute's rhythm. The second, apparently similar to the static form, is the relation among all attributes, which now is more complex since each attribute already has internal relationships.

rhythmic unit							interval
rhythmic cycle							interval
shape's rhythm	scale's rhythm	orientation's rhythm	position's rhythm	tone's rhythm	color's rhythm	texture's rhythm	
cycle interval	cycle interval	cycle interval	cycle interval	cycle interval	cycle interval	cycle interval	
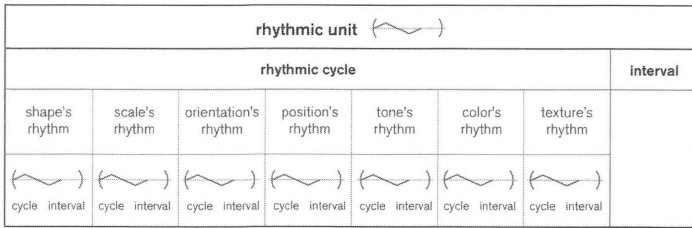
Figure 10. Configuration of the Rhythmic Unit.

Rhythmic Cycle is a whole made of the relationship among all attributes' rhythms: it is a periodic recurrence of the relationship among the cycle and interval of all attributes. It is relevant to point out that the rhythmic cycle is not an "and-summative" whole. Rather a "gestalt," a total structure, a system of relationships in which variables are integrated and coordinated. All attributes are equally important constituents of the whole. They all contribute to the formation of the rhythmic structure, where each individual property has a role. Any change of values of an attribute's property (by means of inter-actions) effects the structure of the whole which is the rhythmic unit's rhythm. The same way that any change in the whole rhythmic structure effects the individual parts, the attributes' rhythms.

The Starting Point of the rhythmic cycle is the **Point of Equilibrium:** the relationship among all Starting Points of all attributes' cycles. It is a variable and it defines the zero point in time of the rhythmic cycle.

Interval is the period of time between the recurrence of the rhythmic cycles. It is the amount of time set for the process of visual formation to start again from its Point of Equilibrium. The interval is a variable and the value can be set by means of inter-actions.

Experiments

Can the elemental constituents of the proposed system allow for the creation of dynamic visual formations in computational media? Can complexity be created out of the exchange of the most elemental spatial and temporal information? How does the creation and re-creation of dynamic visual formations happen in practice?

A series of experiments were created as exploratory environments of the system of dynamic visual formation in the context of rhythmic visual patterns originated by the ever changing relationships among rhythmic units. Ultimately, experiments with the system were used as catalysts for the creative process of visual formation in the computer environment. The experiments are constrained by a rigorous use of the most elemental formal and algorithmical parameters. It is relevant to remark that the constraints are a point of departure and not of arrival.

Spatial qualities are those of elemental 2D geometric forms, in the same way that the algorithms used in the exchange of information are the most basic ones. No colors other than black and white are used. The binary system of positive-negative or negative-positive is explored. However, the attribute of tone plays a major role and allows for gray scale rhythmic combinations.

All rhythmic patterns are modular and serial structures: a rhythmic unit is used as a variable spatio-temporal module that is repeated and organized in a rigid regular grid. All experiments are serial in two respects. One is the way in which units are organized in the structure of the pattern. And the other is the creation of serial rhythmic patterns in the process of inter-actions, when the oneness of units and regularity of the patterns are disrupted and obliterated by qualitative and serial variants and trans-formations.

The same rigor was applied for the choice of input and output devices used in the experiments. Again, they are the most elemental ones: the mouse and the monitor screen. The challenge was to explore the constraints imposed by their very nature and capabilities.

Experiments can be experienced in the url: http://www.atsweb.neu.edu/dva/m.meirelles/dvf/

Analysis of experiments

> I ask myself: what are all the things I perceive for which
> I have no concepts? Or conversely: what are all the things
> I don't perceive because I have no concepts for them?
> Karl Gerstner, *The Spirit of Colors*

One of the central issues explored in the experiments is the perception of motion vis-a-vis the conception of dynamic visual formation, more specifically in relation to the perception and creation of rhythmic visual patterns.

The perception of motion is dependent on a system of references where the distinction between "thing" and "framework" is essential. Koffka explains that in a totally homogeneous field a moving point would not be perceived in motion due to lack of frameworks of reference. In this condition, the point "would be exposed to the *same* stresses everywhere, all positions being dynamically indistinguishable from each other" (Koffka, 1935, 281, italics in original).

The influence of object and field factors plays a major role in the perceptual (aka phenomenal) experience of motion as well as of time. For example, experimentation has showed that "under equal stimulus conditions large objects move (phenomenally) more slowly than small ones" and that "the apparent velocity is the smaller, the greater the field" (Koffka, 1935, 290, 294).

Because "dynamic visual formations" are processes changing in time which are modified by means of inter-actions, so are the relationships among them. Both are time-dependent occurrences. The relationships among rhythmic units create spatial and temporal tensions which are influenced by two interdependent factors. One is the nature of visual formations: the role of the attributes' properties in forming and trans-forming rhythms. And the other is the system of references: the distinction between "object" and "framework." In rhythmic visual patterns—as can be seen in the experiments—the system of references is plural. In these cases, a rhythmic unit is relative not only to the space where it happens (field), but to the other concurrent rhythmic units, also working as frames of references.

Thus, two influential factors play a major role in the perception and conception of rhythmic patterns: the nature of the rhythm and the system of references. The scrutiny of rhythmic patterns is centered in the examination of the limits of these factors. On one hand, their influence in creating uniform and homogeneous patterns, and on the other, the complexities and ambiguities that arise in the entire pattern.

The role of the nature of rhythms in the creation of patterns can be explored in the experiment *3 patterns*. The experiment compares three patterns, each constructed of a singule rhythmic unit repeated 49 times in a regular grid. Differences among the patterns are created by the nature of the units' rhythms. In other words, the attribute playing a major role in forming the rhythm is different in each case: rhythmic unit "a:" tone; rhythmic unit "b:" orientation; rhythmic unit "c:" scale.

The concept of "synchrony" is used as a means to explore differences among the patterns. The term "synchrony" is used to express a coincidental rhythmic occurrence. One or more rhythmic units are in synchrony when their rhythms happen exactly at the same time and with exactly the same values for the properties. The different levels of synchrony point to the differences in perception and creation of the patterns caused by the influence of the nature of the rhythm and that of the system of references.

For example, in the case of all units "a" in synchrony (all units with coincidental values for the attribute of tone) the whole pattern is perceived as one singular uniform unit, to the extent that individual rhythms are obliterated. This is not the case for the other two patterns, where the individuality of the units are maintained, even though the patterns are quite uniform *(figure 11)*.

On the other extreme, that of almost no synchrony among units, in the case of the rhythmic unit "b" (orientation), the pattern is less complex than the other two, where the contrasts of scale and tone are more evident and thus create stronger spatial and temporal tensions among the formations *(figure 12)*.
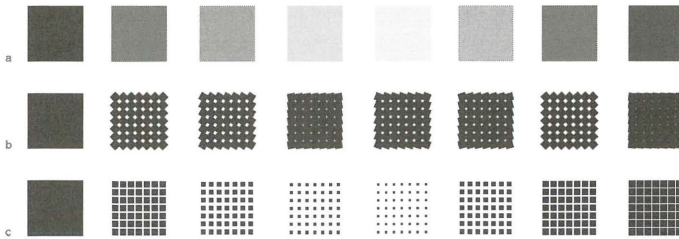
Figure 11. Sequence of images captured from the experiment *3 patterns* showing all units in each pattern in synchrony
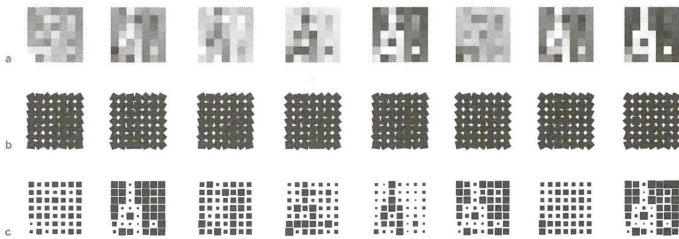
Figure 12. Sequence of images captured from the experiment *3 patterns* showing asynchronous units in each pattern

Similar analysis could be made for units with complex rhythms—where more than one attribute plays a major role in forming the unit's rhythm. In the experiment *dve1*, the original pattern is constructed of a rhythmic unit repeated four times in a regular grid. Because the value of origin for the attribute of scale can be changed, patterns might have up to 64 units. The number of units depends on the values set in the process of inter-actions. Independent of the number (from 4 to 64) the units are always organized in a regular grid.

When rhythms are complex, the concept of synchrony can be explored in diverse ways. One is to have all units with coincidental rhythm for a given attribute, for example only the tonal values being exactly the same. In this case, although there is uniformity in the tonal motion the scale motion is still asynchronous, which keeps some level of complexity in the entire pattern *(figure 13)*.

The other is to synchronize both rhythms of scale and tone attributes. Although this option creates a more homogeneous pattern, there is still some diversity due to the values of origin for the attribute of scale being different. This happens because the spatial property of origin influences the temporal property of duration *(figure 14)*.

Figure 13. Sequence of images captured from the experiment *dve1* showing units in synchrony for Tone rhythm, and differences in the Scale rhythms.
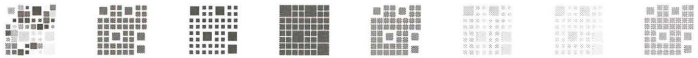

Figure 14. Sequence of images captured from the experiment *dve1* showing units in synchrony for both Tone and Scale rhythms, and differences in values for the spatial origin of Scale.

Further complexities and ambiguities caused by the nature of rhythms and the system of references can be examined in the experiment *dve2*. The experiment is constructed out of the same rhythmic unit repeated 25 times and organized in a regular grid. The unit's rhythm is complex: a compound of the rhythms for the attributes of shape (in "x" and "y" coordinates) and of position (in "x" and "y" coordinates).

When the experiment is first opened two attributes play a major role in forming the unit's rhythm: "x" shape and "x" position. The pattern has very strong vertical tensions and the impression is of five elongated rectilinear units changing width in relation to its left side–which is an illusion, not a true fact, since the value for reference point is set to the center point. The rhythm in relation to the "x" position is almost not perceived.

By using the mechanism of mouse over, the units (restoring the values for the zero point in time of its cycle) become "independent" (and no longer perceived as one long unit). Now ambiguities are more complex in that not only the nature of the rhythm (as before), but also the frames of reference influence the uncertainties in the entire pattern. It is my perception that the units appear from time to time to "join forces" with other units and become "one horizontal longer unit" and then to "separate" again and restore its own independence. My personal impression is that it is quite difficult to identify individual units in the entire pattern unless the value for tone is less than 100%. In this sense, the oneness of units is obscured and at times obliterated by the nature of the rhythm and the uncertainties in the frames of reference. The regularity of the grid is not completely lost due to the strong horizontal tensions created by the relationships among units, even though the pattern is now complex and multiform *(figure 15)*.

If a third attribute contributes to forming the whole rhythm, the uncertainties are enlarged. For example, by setting values for amplitude and velocity of "y" position. In this way, the regularity of the grid is completely lost. The perception of individual units is now very distinct, to the extent that each unit seems to perform a different rhythm that appears to be spatially and temporally independent from all the other concurrent rhythms *(figure 16)*.
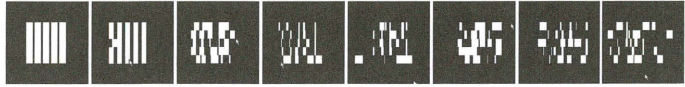
Figure 15. Sequence of images captured from the experiment *dve2* showing the process of inter-actions with a series of units and the creation of new horizontal stresses among units.
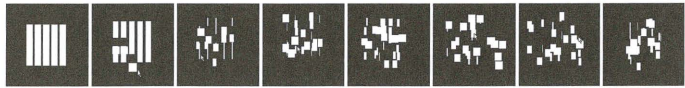
Figure 16. Sequence of images captured from the experiment *dve2* showing the process of inter-actions with a series of units and the obliteration of the perception of the original grid.

Because of the number of variables in this experiment, there are a series of possibilities in the creation of rhythmic patterns by means of inter-actions. Each value that is changed creates new spatial and temporal relationships and tensions among the units. The formation and trans-formation processes disrupt the regularity of the pattern to the extent that new dimensions are discovered. It is an invitation to dynamic visual explorations.

Conclusions

A "system of dynamic visual formation" was proposed with the objective of examining theoretically and experimentally the creative process of image-making in the computer environment. Ultimately, the investigation searched for the most elemental constituents of dynamic visual formation.

The central question to the entire research was whether the proposed system—the basic element and its properties—suffice for the creation of visual formations in computational media. In other words, if the conceptual framework sustaining the system would address and allow for any imaginable visual output. Even though this might seem ambitious, I believe there is a need for such a quest, if we are willing to teach how to create visual forms for and in the computer environment.

I don't think I have answered the question yet. There is a need for further exploration and experimentation with the system in order to determine what is essential and what needs to be incorporated. In this respect, the series of experiments—of which three were analyzed here—are still incipient and so far explore only rhythmic visual patterns.

On the other hand, it is possible to argue that the experimentation with the system already suggests a few essential points towards a theory of dynamic visual formation:

1) The creative process of image-making in computational media is not an individual's isolated activity. Rather, it is a collaborative and participatory endeavor, where at least three agents are necessary: the creator of the program (who defines the system of inter-actions); an active subject and a programmed system (whose inter-actions create dynamic visual formations).

2) In the computer environment a visual element is a variable spatio-temporal module always in the course of becoming, of forming and trans-forming.

3) The most elemental formal and algorithmical parameters create complex and even unpredictable rhythmic visual patterns. The exchange of basic spatial and temporal information in the process of visual formation produces spatio-temporal complexity.

4) Two interdependent factors play a major role in the perception and the creation of rhythmic patterns: the nature of a dynamic visual formation's rhythm, and the system of references.

Those issues in fact point to directions for further research. There are two major areas of future research. The first concerns the system of inter-actions: the exchange of spatial and temporal information between the agents. It includes the nature of algorithms and of input and output devices. The experiments were constrained by elemental algorithms and used the most basic input and output devices. It is quite obvious that those two areas can be easily implemented with more complex algorithms or with other existing input and output devices. The direction, however, does not seem to depend on purely technological decisions. Rather what should be considered is the conceptual gains effected by the decisions. For example, the search for a more "gesture-based" system of inter-actions could determine the choice of input devices. Is it possible to integrate degrees of freedom that artists are used to in everyday practice? To what extent would this affect the creative process of visual formation? What kinesthetic properties should be considered in the creative process?

The second area for further research concerns the quality of spatial and temporal information used in the system of dynamic visual formation. This is the area presently being carried out: the system presented here is under implementation by including the attributes of color and texture in a series of new experiments. The purpose is to further explore the correlation of properties for different attributes and to study their roles in the creation and re-creation of rhythmic visual patterns.

References

Dondis, D.A. 1973. *A Primer of Visual Literacy.* Cambridge, Mass: MIT Press.

Gerstner, Karl. 1981. *The Spirit of Colors.* Cambridge, Mass.: MIT Press.

Kandinsky, Wassily. 1982. *Kandinsky, Complete Writings on Art.* Lindsay, K.C. and P. Vergo, editors. Boston, Mass: Hall.

Kandinsky, Wassily. 1979. *Point and Line to Plane.* New York, N.Y.: Dover Publications.

Klee, Paul. 1969. *The Thinking Eye: the Notebooks of Paul Klee,* Spiller, J., editor. New York : G. Wittenborn.

Koffka, Kurt. 1935. *Principles of Gestalt Psychology.* New York, N.Y.: Harcourt Brace and Co.

Manovich, Lev. 2001. *The Language of New Media.* Cambridge, Mass.: MIT Press.

Popper, F. 1968. *Origins and Development of Kinetic Art.* London: Studio Vista.

Author Note

Isabel Meirelles teaches graphic design at the Department of Visual Arts, Northeastern University in Boston. She has a BArch in Architecture and Urban Planning from Febasp, Sao Paulo, Brazil, a MArch in History and Theory of Architecture from the Architectural Association School of Architecture in London and a MFA in Graphic Design from Dynamic Media Institute, Massachusetts College of Art, Boston. Her primary research interest is twofold: to examine both theoretically and experimentally new complexities in the visual realm brought by computational media and its connections with other fields of knowledge; and to investigate the field of information design, the ways in which we perceive and organize information. Professional experience includes work as an architect and urban designer, head of museum departments and art director in publication and interactive design. A recent interactive new media project is the CD-ROM Presence: The Ephemeral in Focus, awarded Honorable Mention by the American Association of Museums 2004 Museum Publications Design Competition.

Abstract

One of the most important, and most studied, aspects of human perception is the act of reading. Reading has received much attention from researchers, both from a human information processing (HIP) approach and as a common, practical act that needs to be optimized, especially in the realm of human-computer interaction (HCI). One of the text variables that has been studied for over 100 years is line length, at times referred as line width. Psychologists, typographers and others working in the field of reading and advertising have demonstrated the effects of line length on readability of text. Two of the questions addressed in past studies include, How long should a column of text be, to optimize readability of the text?, and Which view is more preferred by readers ñ multiple narrow columns or one wide column with the same amount of information content?

Optimal Line Length in Reading –
A Literature Review

Nanavati and Randolph G. Bias

Research has led to recommendations that line length should not exceed about 70 characters per line. The reason behind this finding is that both very short and very long lines slow down reading by interrupting the normal pattern of eye movements and movements throughout the text. In a world of personal digital devices (PDAs), one-inch cell phone displays and of wide-screen TVs and full-wall computer displays, the question of line length has renewed timeliness. Studies reviewed here show that different aspects of reading performance such as comprehension, reading speed, method of movement (e.g., paging and scrolling) and eye movements are affected by changes in line length. In addition to that, various typographic factors such as font type and size, line and character spacing as well as different screen structures such as varying number of columns and screen sizes also affect readability. These factors have an effect on optimal line length for the text read from printed or on-screen material.

Introduction

One of the most important, and most studied, aspects of human perception is the act of reading. Reading has received much attention from researchers, both from a human information processing (HIP) approach (see Gough, 1972) and as a common, practical act that needs to be optimized, especially in the realm of human-computer interaction (HCI) (see Gould et al., 1987).

From both the HIP approach and the HCI-optimization approach, it has been interesting to examine what variables in the reading act lead to enhanced or inhibited reading performance. These variables can be divided into the following types:

• Variables within the reader (e.g., visual acuity, education level, reading experience, familiarity with the reading material).

• Typographic variables in the text itself, such as font type and size, line and character spacing, line length and column structure.

• Variables within the environment (e.g., ambient lighting, visual angle, tailorability of the reading material and medium of the text, i.e., printed or on-screen reading material).

• And within the on-screen text, display characteristics such as contrast, flicker, aspect ratio and image polarity (Gould et al., 1987; Dillon, 1992).

The advent and evolution of computer technology, including computer terminals, the personal computer, selectable fonts and variable-width fonts, has spawned a wealth of studies design to identify the settings—controllable through typographic and environment variables—that will optimize reading performance and enjoyment.

One of the text variables that has been studied for over 100 years is line length, at times referred as line width. The effects of line length upon readability of text have been demonstrated by psychologists, typographers and others working in the field of reading and advertising. Two of the questions addressed in past studies include, How long should a column of text be, to optimize readability of the text? and Which view is more preferred by readers—multiple narrow columns or one wide column with the same amount of information content? Research has led to recommendations that line length should not exceed about 70 characters per line (Spencer, 1968). The reason behind this finding is that both very short and very long lines slow down reading by interrupting the normal pattern of eye movements throughout the text. In a world of personal digital devices (PDAs), one-inch cell phone displays and of wide-screen TVs and full-wall computer displays, the question of line length has renewed timeliness.

Studies reviewed here show that aspects of reading performance such as comprehension, reading speed, method of movement (e.g., paging and scrolling) and eye movements are affected by changes in line length. In addition to that, various typographic factors such as font type and size, line and character spacing as well as different screen structures such as varying number of columns and screen sizes also affect readability. These factors have an effect on optimal line length of the text read from printed or on-screen material.

A fundamental finding

"Line lengths that are recommended as optimal for print are not the most legible on screen, when reading rate is used as a measure of legibility" (Dyson and Kipping, 1998). The differences between reading hard copy and reading a computer display—differences such as lighting source, glare potential, whether the text can be moved, ability (or lack of it) to dynamically reflow the text, potential for color change or text motion and visual angle, mean that a century of research data on optimal line lengths in print may or may not be applicable to the on-line reading situation.

Measuring line length

Early studies in the field of reading measured line lengths in terms of picas. One pica is one sixth of an inch. Some of the studies measured line length in the units of inch, centimeter or millimeter. More recent studies tend to measure line length as total number of characters in a line or characters per line (cpl). Units like picas and millimeters can be interpreted in cpl for better comparison of line lengths. Bringhurst (1992) presented a method for determining the number of characters in a line for any font face and size: measure the length of the lowercase alphabet and use a copyfitting table (abbreviated version shown in *table I*) that shows for a given alphabet length, the average number of characters in that line.

Another approach to measuring line length is based on the size of the text and the distance at which the text is read or visual angle of a line. As called out in Gould et al. (1987), visual angle of a line of characters is the angle formed by the width of the line as it relates to the reading distance of that line from the reader's eye. This is useful in the sense that printed text and text on screen are normally viewed at different distances from the reader, with printed text viewed at a shorter distance than the text on screens.

POINTS	LINE LENGTH IN PICAS							
	10	14	18	22	26	30	35	40
80	40	56	72	88	104			
85	38	53	68	83	98	113		
90	36	50	64	79	86	107		
95	34	48	62	75	89	103		
100	33	46	59	73	86	99	116	
105	32	44	57	70	82	95	111	
115	29	41	53	64	76	88	103	
120	28	39	50	62	73	84	98	112
125	27	38	48	59	70	81	94	108
130	26	36	47	57	67	78	91	104
135	25	35	45	55	65	75	88	100
140	24	34	44	53	63	73	85	97
145	23	33	42	51	61	70	82	94
150	23	32	41	51	60	69	81	92
155	22	31	39	49	58	67	79	90
160	22	30	39	48	56	65	76	87
165	21	30	38	46	55	63	74	84
170	21	29	37	45	53	62	72	82
175	20	28	36	44	52	60	70	80
180	20	27	35	43	51	59	68	78
185	19	27	34	42	49	57	67	76
190	19	26	33	41	48	56	65	74
195	18	25	32	40	47	54	63	72
200	18	25	32	39	46	53	62	70
220	16	22	29	35	41	48	56	64
240	15	20	26	32	38	44	51	58
260	14	19	24	30	35	41	48	54
280	13	18	23	28	33	38	44	50
300	12	17	21	26	31	35	41	47
320	11	16	20	25	29	34	39	45
340	10	15	19	23	27	32	37	42

Table I: Copyfitting table—average character count per line (Reproduced by author permission, Bringhurst, 1992.)

Read down, in the left column: lowercase alphabet length in points. *Read across,* in the top row: line length in picas.

Early findings

Early research into the effects of line lengths on readability concluded that moderate line lengths are optimal for efficient reading compared to very short or very long lines. A number of early studies conducted by Tinker and Paterson (1940) (summarized in Tinker, 1963) had shown that optimal line length of text read from printed material was 13 picas or somewhere in between 59 to 97 mm or 52 cpl, (as interpreted by Rayner and Pollatsek, 1989), which was moderate in between very long lines of 36 picas and very short lines of 5 picas. The explanation given for the legibility of moderate line length was that, if the lines are too long, it is difficult to make an accurate return sweep to the beginning of the next line and if they are too short, readers cannot make maximum use of horizontal cues and other more peripheral information in each fixation.

Burt, Cooper and Martin (1955) also found better reading speeds and ease of reading at medium line lengths between 3.5 to 5.5 inches (approximately 89 to 140 mm). The reason given for this was similar to Tinker and Paterson's findings (1940).

Summary of Studies

This section of the paper reviews literature relevant to the factors affected by line length and factors affecting line length of printed or on-screen text.

Factors affected by line length

User responses from various studies concerning the readability of text from computer screens as well as printed material address subjective and objective attributes of reading affected by line length of the text. Subjective aspects of reading are attitudinal responses of test participants while reading text during the experiment and objective factors are the behavioral and cognitive aspects of readability.

Subjective factors

These factors include readers' attitudes toward line lengths while reading. People's subjective preferences like reported ease of reading and satisfaction, affected by line lengths, usually did not correlate with their performance in range of experiments conducted to explore the effects of line length on readability (see Dyson and Kipping, 1998). This means that what people reported as easy to read was not objectively measured faster to read nor did it yield better comprehension in the experiment. Some of these subjective factors were studied based on various objective measures. For instance Luckiesh and Moss (1941) measured the subjective factor of ease of reading shorter lines (13 to 17 picas) based on objective measure—number of blinks that occurred during reading.

Ease of reading

In 1998, Dyson and Kipping conducted two experiments to explore the effects of line length and paging versus scrolling on reading from computer screens. In the first experiment each participant, out of 24, was asked to read six documents with different line lengths of approximately 25, 40, 55, 70, 85, 100 cpl. The time taken to read each document was recorded to evaluate reading speed at various line lengths. Participants were asked to compare documents at each line length with every other line length and say which they thought was easier to read. Ease of reading was not explicitly defined so participants could take any aspect of reading such as reading speed or effective comprehension as a measure for ease of reading. As a result of the experiment, Dyson and Kipping found that moderate lines of 55 cpl were reported easier to read than very short lines of 25 cpl or very long lines of 100 cpl using scrolled or paged movements to navigate through the document. The experiment reported that participants did not base their judgments on feedback from their speed of reading or degree of understanding; instead they monitored their level of comfort or discomfort when reading particular line length. Dyson and Kipping also noted that perception of ease of reading from computer screen at 55 cpl might be influenced by experience in reading certain types of printed material with line length in between 50 to 70 cpl. Burt, Cooper and Martin (1955) also found reports of ease of reading at medium line length of 55 cpl in printed text. The reason given behind this was difficulty in picking up the beginning of the next line (return sweeps) in long lines and inability to take in large phrases of text with single fixation and make maximum use of peripheral vision in short lines. On the other hand based on a judgment of faster readers Dyson and Kipping (1997b) found that a three column format with shorter lines was easier to read.

In the second experiment by Dyson and Kipping (1998), relatively bright white areas in the test document were replaced with a light gray background color to test whether or not the amount of glare might have factored into their findings. Results of this experiment remained consistent with the previous one in concluding ease of reading among different line lengths read. As a result of these two experiments Dyson and Kipping concluded that line length should be considered a significant factor as a criterion for judging ease of reading.

User preference and satisfaction

Shaikh and Chaparro (2004) considered user preference, satisfaction and reading efficiency as dependent variables of interest in their experiment investigating the effects of four different line lengths (35, 55, 75, and 95 cpl) on reading speed and comprehension. In their study user preference for line length was found to be inconclusive. Users either loved or hated the extreme line lengths (35 and 95 cpl) and judged the medium line length as optimal. Shaikh and Chaparro found no effects of line length or passage type (short news stories or long narrative) on the overall user satisfaction. No direct effects of line length were found on reading efficiency.

Objective factors

Various factors associated with reading such as reading speed, comprehension, eye movements, method of movements for navigation and peripheral vision are affected by line length of text read. These aspects of reading acted as a base for various studies carried out on the effects of line length on readability.

Comprehension

Out of many variables identified as possible influence on comprehension, one is the amount of text that can be viewed simultaneously on-screen (Dyson and Kipping, 1997a). The amount of text displayed on screen can be manipulated by varying line lengths. A number of studies have been conducted to identify the effects of line lengths on comprehension. Dyson and Kipping (1997b) studied the effects of one wide column (about 80 cpl) against three narrow columns (about 25 cpl each) of text documents on various comprehension tasks. Eighteen subjects were asked to silently read three documents displayed on screen. Comprehension was assessed by how successfully participants were able to answer questions about whether or not a series of questions was answerable or not based on what they had read. Researchers found no overall difference in comprehension. However, based on a judgment of a group of subjects they noted that the three column format with short line length, similar to those found in printed magazines, improved comprehension for fast readers. As a possible explanation for this, Dyson and Kipping stated that different types of readers may adopt different reading patterns that affect their comprehension. For example faster readers might have benefited from three narrow columns with short lines due to their ability to easily scan through narrow columns and extract meaning.

In another comprehensive study conducted by Dyson and Kipping in 1998, they found similar results with no significant effect of line length on comprehension. As reported above, they had 24 participants read six documents with different line lengths of approximately 25, 40, 55, 70, 85 and 100 cpl. There was no evidence of speed-accuracy trade-off between speed of reading and comprehension in the study. They reported that faster reading rates did not appear to be at the expense of comprehension. As a second experiment of the same study, relatively bright white areas in the test documents were replaced with a light gray background color, but results for comprehension remained the same. On the other hand Dyson and Haselgrove (2001) had 36 participants read documents edited to a total length of approximately 1000 words with varying line lengths of 25, 55 and 100 cpl at fast and normal reading speeds. They hypothesized that readers may vary their reading rate to maintain a relatively constant level of comprehension across different line lengths (i.e., speed-accuracy trade-off). They attributed the reason behind this contrast to the requirement for participants to recall details and make inferences in their study which was missing in Duchnicky & Kolers (1983) and Dyson & Kipping (1998). Dyson and Haselgrove (2001) concluded that overall

comprehension was reduced when reading fast, but that the type of information recalled was not dependent on the speed of reading. They explained the effects of line length on comprehension in terms of mechanics of reading (nature of eye-movements), mechanics of scrolling and consolidation of what is read. Both very short and very long line lengths can slow down reading by disrupting the normal pattern of eye-movements. They recommended a medium line length of 55 cpl compared with long lines of 100 cpl for better comprehension due to the interruptions that occurred with long line lengths because of locating the beginning of a line, following a return sweep and disrupted concentration.

McMullin et al. (2002) measured the effects of surrounding information and line length on text comprehension. They found that participants got distracted by the additional column or paragraph of information and performed slightly less well on comprehension questions than when information was surrounded by white space. They favored the use of white spaces over multiple columns, as white spaces helped prevent the influence of distracting and unimportant information and decreased the need to scan across the entire screen, which could be tiring for the viewer's eye span. Results from this study supported the research by Dyson and Haselgrove (2001) showing that very small line lengths (25 cpl) lead to less comprehension than line lengths of 55 and 100 cpl. Shaikh and Chaparro (2004) noted a speed-accuracy trade-off while participants read faster at the longest line length of 95 cpl and they actually comprehended less.

Reading rate
Numerous studies have been conducted to explore the effects of line length on reading rate on printed as well as on-screen material. Early research on legibility of different line lengths in print has been measured by reading rate with only a check on comprehension (Tinker, 1963). Past studies have shown that other factors like method of movement (scrolling and paging) and eye movements with fixation frequency, pause duration, regression frequency, and saccade duration also affect reading rate as a function of line length (Paterson & Tinker, 1940; Luckiesh & Moss, 1941). Using reading performance test results Tinker and Paterson (1940) found that very short line widths (seven picas) and excessively long line widths (36 picas) for six point type size, produced delayed reading rates in comparison with moderate line widths of about 14 picas. They summarized the reasons behind delaying rates in reading long lines as an inaccurate return sweep, from the end of one line to the beginning of the next, causing the reader to re-read the same line or omit the next, and in reading short lines as inability to make maximum use of horizontal cues. Another study

conducted by Tinker and Paterson in 1940, compared reading performances on speed of reading for 80 mm lines (form A) with varying line lengths of 59, 97, 114, 136, 152, 168 or 186 mm (form B), both at 10 point type size. They found 80 mm lines yielded faster reading than any of the other line lengths and concluded that optimum line length lies somewhere in between 59 to 97 mm. They also suggested that line lengths should be kept within the limits of 75 to 90 mm if speedy reading is desired. They showed the relation between speed of reading and the ability of the reader to establish and maintain regular rhythmic eye movements. They claimed that less efficient reading with very short lines is due to inability to establish rhythmical eye movements.

An experiment comparing form A and form B (see Tinker and Paterson, 1940) also revealed that negative or positive change in optimum line length affects fast readers in terms of retardation in speed of reading and improves speed of reading for slow readers. On the other hand Luckiesh and Moss (1941), who studied the influence of line-length on readability by measuring the rate of involuntary blinking of the eyes, patterns of eye movements in reading and rate of normal reading on 10-point text type with line lengths of 13, 17, 21, 25, 29 picas, found no influence of line-length on speed of reading. Yet they considered shortest line length of 13 picas superior to lines of longer lengths for increased readability, attributed to fewest blinks during reading shortest lines. However, rate of blinking as an appropriate criterion of readability is questionable and produces unreliable data (Tinker, 1963; Perera, 2004). Tinker (1963) considered rate of blinking as a measure of legibility questionable, if no experimenter other than Luckiesh and his colleagues can duplicate the results in Luckiesh and Moss (1941).

Duchnicky and Kolers (1983) found that 80-character lines were read faster than 40-character lines based on their experiment on readability of text scrolled on visual display terminals as a function of three different variables: line lengths with respect to window size, number of characters in each line and window heights. They had subjects read 30 noncontiguous passages of about 300 words with two character densities (40 or 80 cpl), five window heights of 1, 2, 3, 4 or 20, lines and three different line lengths as approximately full, two-thirds and one-third the screen width. They found that text in 80-character lines were read 30% faster than text in 40-character lines, lines of two-thirds or full-screen widths were read 25% faster than lines of one-third screen widths and text in one-line or two-line windows was read 9% slower than text in 20-line windows. These findings led to the conclusion that speed of reading scrolled text is increased with increase in line length, character density and window height, with window height having less effect than the other two.

Dyson and Kipping (1998) conducted two experiments to explore the effects of line length on reading from a screen. Participants were asked to silently read a series of documents displayed on the screen with line lengths of 25, 40, 55, 70, 85, 100 cpl and scrolling and paging as methods of movement. As participants read and moved through the document with the down and up keys, the time taken to read each document was recorded. The total reading time per document was used as a measure of reading rate. They found that 100 cpl was read significantly faster than 25 cpl in both scrolled and paged conditions. In the scrolled condition they argued that less time had to be spent on scrolling through the documents, as there were fewer lines of text (also Dyson and Haselgrove, 2001). This pattern of scrolling exploited the full size of the window, as more lines of text were read before moving further down in the document allowing for processing of larger chunks of text without interruption from scrolling movement. In the paged condition, the shorter time could be attributed to fewer key presses at 100 cpl than at 25 cpl while navigating through the document. Another possible explanation given for faster reading at 100 cpl was that this line length reduces some glare from the screen, as text was covering most of the screen. Line lengths that did not fill the screen had an area of white to the right of the text which might have proved distracting. This glare is increased when lines are short, but can be reduced by having multiple columns of short lines. Glare had also been reported as a problem in a comparison of different CRT displays by Gould et al. (1987). As a follow up experiment of the same study Dyson and Kipping (1998) had hypothesized that by reducing the glare from the screen using a gray background, shorter lines were read at similar rates to longer lines. This study had revealed that participants may adjust their reading patterns according to the line length they are reading and the amount of scrolling required, thus achieving better reading rates. In 1997, Dyson and Kipping (1997a) concluded that text displayed in a single wide column (80 cpl) was read faster than a three column (25 cpl each) format, when both used paged movement to advance through the document. In contradiction to Dyson and Kipping (1997a), they further concluded that long lines, when read at fast speeds, facilitated skim reading as more information displayed at a time could aid readers to skim through it easily. In their experiment on the effects of line length on reading speed and comprehension, Shaikh and Chaparro (2004) found that while reading rate was fastest in 95 cpl, reading efficiency did not differ based on line length. This implies a speed-accuracy trade-off as participants read faster at 95 cpl but actually comprehended less.

Method of movement

Method of movement refers to how a reader navigates through the text on screen. Studies have been conducted to analyze the effects of line length on comprehension and reading rate influenced by method of movement like scrolling and paging. However, research that has investigated methods of moving through the text on screen has not produced any clear findings. Readers use various scrolling patterns measured by the time spent in pauses, time in scrolling, the length of the first pause and the number of scrolling movements (Dyson and Haselgrove, 2001). Dyson and Kipping (1997a) found that paging was faster than scrolling, as time spent in the physical action of scrolling accounted for the slower reading rate. In 1998, Dyson and Kipping studied the effects of paging versus scrolling on reading from the screen. They noted no significant difference in reading rate in scrolled or paged conditions as well as no interaction between line length and method of movement in their findings. They concluded that reading from a screen permits a number of ways of scrolling through text which may result in a range of reading patterns. These patterns may be influenced by line length and result in varying degree of efficiency of reading. Dyson and Haselgrove (2001) had believed that scrolling time may be longer with short lines as they require more time to bring up an equivalent amount of new text than with longer lines. They reported the influence of line length on time spent in pauses, with the longest pauses at the longest lines resulting in slower reading rate. Participants might have felt that they can allocate more time in pauses as they did not need to spend much time in scrolling long lines. More detailed explanation of the effects of line length on comprehension and reading rate with scrolling and paging from these studies can be found in the previous two sections of this paper.

Eye movements

Measuring eye movements has been an important part of determining readability of text from print as well as screen media. While reading, the eye does not continuously scan a line. Rather, it stops (fixates) for about one-quarter of a second and then rapidly jumps to another place on the line; these stops and jumps are known as fixations and saccades (Gould et al., 1987) and information is perceived during the fixations. When the end of the line is reached readers traverse back to read the beginning of the new line. These return movements are referred to as return sweeps. Backward moves within a line to re-examine material not clearly perceived or understood are know as regressions (Tinker, 1963). Various studies have measured the effects of eye movements on readability of text in terms of fixation frequency, fixation duration, perception time, regression frequency and saccade duration. Eye-movement records from different studies (summarized in Tinker, 1963) show that more fixation pauses of greater duration occur while reading very short lines (Tinker, 1963). Further, reading very long lines is less efficient due to regression following the return sweep to the beginning of the new line. Early research conducted by Tinker and Paterson (1940) reported that while reading excessively short lines, fixation frequency was increased with fewer words per fixation, pause duration was lengthened

and perception time (combination of fixation frequency with pause duration) was greatly increased. Luckiesh and Moss (1941) noted that eye muscle mechanisms would become less capable of making small and precisely controlled steps in each fixation as lines get longer, making reading more difficult due to erratic eye movements.

Kolers et al. (1981) measured the readability of CRT displays based on eye movements. Their results show that increasing the number of characters per line (from 40 to 80) by halving their widths increased the number of fixations but the total number of fixations per passage was decreased. In addition, the number of words extracted per fixation was larger, the duration of each fixation was longer and the total reading time was shorter. Duchnicky and Kolers (1983), replicating findings from Kolers et al. (1981), supported this conclusion that very short lines elicit more and longer fixation pauses. With short line length readers may not be able to make use of much information in each fixation and may decrease their saccade length which could slow down the reading (Rayner and Pollatsek, 1989). Luckiesh and Moss (1941) observed the unique characteristic of eye movements during the reading of lines of different lengths. Data from their study revealed that the average number of fixations as well as average fixation-span was increased as the line-length was increased, which was contradictory to Duchnicky and Kolers (1983). Lynch and Horton (1999) showed that wider lines of text require readers to move their heads slightly or strain their eye muscles to track over the long text, so readability suffers because readers may lose track of the next line on the long trip back to the left margin. On the other hand, Kolers et al. (1981) favored long line lengths of 80 cpl as more information was extracted from the text with each fixation, for more efficient reading. They achieved this by halving letter widths, so they didn't increase the actual line width here.

Peripheral vision
Readability of text read from screen or print material is also affected by how readers make use of horizontal or vertical peripheral cues generated by reading end of lines. Readability is decreased due to inability of readers to make maximum use of these cues. Tinker and Paterson (1940) observed retardation in the rate of reading an excessively short line, due to inability to make maximum use of horizontal peripheral cues in reading printed text. They found that efficient use of peripheral vision during the first fixation of a line yielded premonitions of coming word, phrase positions, sentence breaks and meanings with an optimum line length (i.e., moderate line length of 13 picas), which could guide and shorten future fixations. They also noted that peripheral vision in a vertical as well as a horizontal direction provides visual cues of words in the line succeeding the one in direct fixation to facilitate reading. Furthermore, peripheral vision accompanying the first fixation in a line is less effective with long lines than with short lines. (For printed text Burt, Cooper and Martin, in 1955 concluded that short lines prevent the eye of the trained reader from taking in large phrases with a single fixation and making optimal use of peripheral vision.

Factors affecting line length

A number of studies have been conducted assessing the factors that affect line length of material read from screen as well as print media. The studies so far in this field discuss the structure of the screen including number of columns, vertical or horizontal screen organization, screen size and typographical factors like font type and size, line leading, visual angle, spacing and justification.

Print versus screen media

Variation in optimal line length between print media and on-screen text may be due to a range of reading patterns generated with a number of different ways of scrolling and paging to navigate through the document (Dyson and Kipping, 1998). Kruk and Muter (1984) found that the text was read significantly slower from screen than from book when both had a comparable amount of information at 39 cpl. Various experiments were conducted by Gould et al. (1987) in an attempt to answer why reading is slower from CRT displays than from paper. They reported that lines of characters on CRT displays are often wider than the same lines printed on paper. They found that at optimum conditions reading performance from a display could get to within 90% of reading print. They also noted significantly slower proofreading speed with computer display compared with paper. On the other hand, Kahn and Lenk (1998) recommended that optimal line length should not vary between print and screen media. Results from print research do not address specific characteristics of computer display such as monitor refresh rate, screen resolution, gamma rate (non-proportional rate at which a monitor screen "gets brighter" as the electrical signal applied to it increases) and reflected vs. emitted light. The consensus is that print results should not be applied to Web sites or other computer-based platforms without consideration of the unique aspects of computer monitors (McMullin et al., 2002).

Screen structure

Readability of text is also affected by the way information is organized onto the screen. Number of columns (one wide vs. multiple narrow columns), large or small screen sizes, vertical or horizontal screen organizations and window height are some of the driving factors affecting line length. In a number of studies, rate of reading and comprehension were studied based on changes in line lengths depending upon these factors.

Number of columns

Line length measures can be confounded with the amount of information on the screen. So experiments studying the effects of number of columns on line lengths should compare one wide column with multiple narrow columns where each page has the same amount of information and therefore the same amount of potential glare. As stated earlier, Dyson and Kipping (1997) tested the ease of reading specific formats used for on-line publications. A single column format with relatively long line length of 80 cpl was compared with a three column format with a line length of 25 cpl. They found that a single wide column was read faster with no overall difference in comprehension than three narrow columns when both used paged movement to advance through the document (also (Dyson and Kipping, 1998). In contrast to reading performance in this study, subjective judgments of ease of reading rated the three-column format easier to read for fast readers, reportedly due to their ability to easily scan through narrow columns and extract meaning. Efficient reading with three column format for speedy readers was also supported by Brown (1970), who explained that the three-column format reduces the need for lateral eye movements and thereby allows the readers to scan vertically down the text. On the other hand Duchnicky and Kolers (1983) indicated that text in narrow columns was read significantly more slowly than text in wider columns on CRT displays.

There are several other contradicting recommendations made regarding the effects of line length in single or multicolumn display. Kahn and Lenk (1998) recommended a line length of 66 to 75 cpl for a single column layout and 30 to 40 cpl per column for a multiple column layout for on-screen reading. Lynch and Horton (1999) offered a psychological reason behind narrow columns: at normal reading distance the eye's span of acute focus is only about 3 inches wide. Bringhurst (1992) in his book *Elements of typographic style* proposed an optimal line length for a single column page between 45 to 75 cpl (with 66 cpl being ideal) and for a multiple column page between 40 to 50 cpl per column. Reynolds (1980) recommended a single column of text on the screen while Mourant, Lakshman and Chantadisai (1981) preferred three-column layout attributable to reduced visual fatigue. Luckiesh and Moss (1941) stated that lines of 13 picas in length are superior to greater line lengths for print, but it should not be assumed that two 13-pica columns on the same page would be superior to a single column of 26 picas, due to a possibility that the second column might be a distraction in reading of the first, and vice versa. This argument was also supported by visual masking data (Turvey, 1973), that when a visual stimulus is too near to other stimuli, the ability to perceive each is diminished. This collection of apparently contradictory studies seems to point towards readers' preference for relatively narrow columns of text, but performance superiority, at least for fast readers, on single,

Screen size

Dyson and Kipping (1998) concluded that shorter reading time occurred for on screen material at 100 cpl. As a possible reason for this they cited de Bruijn et al. (1992) who investigated the effects of screen size (12 versus 15 inch) on learning of text presented on the monitor of a personal computer. De Bruijn et al. (1992) found that subjects using a 15 inch screen needed less learning time than subjects using a 12 inch screen, with no difference in learning performance. Learning was assessed by summarization and a multiple choice test after the reading task to measure the amount of information retained. They interpreted this advantage in terms of an efficient integration process in constructing the semantic representation of text from a single page. Dyson and Kipping (1998) also noted that lines that did not fill the screen had an area of white to the right of the text that could distract readers resulting in slower reading. Hansen and Haas (1988) considered page size as one of the primary factors that influence the behavior of users as they read and write with computers. In their experiments they utilized small (personal computer screen holding 24 lines of 80 characters each) and large (workstation screen holding 46 lines of about 80 characters each) page sizes. Their various reading experiments showed reading from workstation superior to reading from personal computers. They further noted that page size could affect reading tasks by limiting the context for the visible text, thus burdening short term memory. If the page size is small, it takes time to scroll through the document and that interferes with concentration. Still, optimum display size remains an issue that requires further empirical attention (Dillon et al., 1990).

In all these studies it is important to note the effects of screen resolution on length of line for onscreen reading. Bridgeman et al., (2003) in their study on effects of screen size, screen resolution and display rate on computer based test performance, noted that screen resolution is critical because it impacts both the size of text on the screen and the amount of information that can be displayed. Also, a higher resolution allows more words per line and more lines per screen than can be displayed in a lower resolution.

Typographical factors

The literature reviewed so far in this paper addressed psychological aspects of reading from screen and print media. Studies cited showed how these factors relate with line length and readability of text. This section of the paper describes typographical factors like font size and type, line leading, spacing and visual angle and their effects on readability with a stress on line length.

Font type and size

Early research conducted by Tinker and Paterson (1940) (summarized in Tinker, 1963) showed a strong relation between line width and font type sizes. They recommended line widths ranging between 17 and 37 picas for 12-point type size, between 17 and 27 picas for 10 points, between 13 and 25 picas for 8 points, and 9 to 25 picas for 6 points. Of course if amount of information is held constant, line length increases with increase in font type sizes. They also noted that 6- and 8-point type sizes were read slower than the larger type sizes (9, 10, 11, 12 points) due to the increased number of fixations, fewer words per fixations, longer pause duration and longer perception time with more regression. (Especially the 'fewer words per fixation' for smaller fonts is counterintuitive, and, we would guess, wrong. But that is what they reported.) Bringhurst (1992) recommended that length of line (in point unit) should be around 30 (between 20 to 40) times the point size of a font type in a conventional book page. For example for 12-point font size line length should be 360-point.

Line leading

The amount of space between lines of text, line leading, is closely related to line length effects. These effects were studied for both print as well as on-screen reading. For on-screen reading, Grabinger and Osman-Jouchoux (1996), in their recommendation for basic typography, noted that space between the lines helps readers maintain vertical position in the text. Readers prefer short lines of about eight to ten words or 45 to 60 characters long that are single spaced. The longer the line, the harder it is to maintain position on that line, therefore more leading is needed. Hedrick (2002) also recommended more leading as the line gets longer. But he strongly disagreed with the use of double line spacing. Kolers et al. (1981) studied the effects of space between lines on reading performance. Their results showed that single spacing produced more fixations per line, that is fewer words per fixation. Therefore total reading time was longer with single spacing than with double spacing. Lynch and Horton (1999) presented a line spacing rule: make line spacing at least 1/30 of the line length to give a two-degree downward angle for finding the next line (HFS, 1988). Though additional line leading allows longer line length without sacrificing leg-

ibility (Lynch and Horton, 1999), line length of more than 75 to 80 cpl with proper leading is too long for continuous reading (Bringhurst, 1992). For printed text, Tinker (1963) presented limits of variation in line width and leading ('safety zones') that may be used for a given type size without loss of legibility from various experiments and surveys in printing practice.

Visual angle

The visual angle of line of characters is the angle formed by the width of the line of characters as it relates to the reading distance of that line from the reader's eye. More precisely, the visual angle of a line of characters is twice the arc tangent of the ratio of one-half the width of that line to its distance from the reader's lens (Gould et al., 1987). Experiments conducted by Gould and Grischkowsky (1984) had a line on a CRT display 1.52 times the width of the same line printed on paper in 10-point type size. In these experiments the visual angle subtended at viewer's eye by a 69-character line on CRT display was greater than the visual angle subtended by a 69-character line on paper with equal reading distance from both the media. Thereafter Gould et al. (1987) conducted a number of experiments in an attempt to explore multiple variables, with reading distance (visual angle) as one of them, for why reading is slower from CRT displays than from paper. Their reading distance experiment concluded that people sit farther away from a CRT screen than from paper which led to smaller visual angle by CRT display than by paper. From the visual angle experiment they noted that speed and accuracy of proofreading were reduced at extreme visual angles. At small visual angles, characters were hard to discriminate and at large visual angles more eye fixations were required due to fewer characters perceived peripherally. From various experiments conducted by Gould et al. (1984; 1987) it is not clear whether people are sitting farther from the screen because lines are longer or lines are kept longer on screens because people sit farther from the screen. McMullin et al. (2002) noted that line length subtended by visual angle increases with the distance from the viewer's eyes to the reading material. Since computer monitors are viewed at greater distance than most of the printed material, eyes can transverse greater line lengths with motions that are equivalent to the recommended line lengths in print.

Word spacing and text justification

One of the early psychological studies in typography was conducted by Burt, Cooper and Martin in 1955. They concluded that short lines require widely varying spaces between words and increase the number of broken words at the end of lines. Bringhurst (1992) also supported this by stating that line length less than 38 or 40 cpl will lead to uneven spaces and more hyphenation at the end of the line which is hard to read. Hedrick (2002) recommended that no more than three lines in a row should be hyphenated. He also found difficulties encountered in reading lines of text with uneven spaces and hyphenation.

Campbell (1981) studied the effects on reading speed of two different types of right justification in computer printed text, one with fixed character spacing and the other with variable character spacing, where extra space is distributed proportionally between and within the words of the line. He noticed that right justification with variable character spacing produced faster reading than fixed character spacing or unjustified text. On the other hand Gregory and Hartley (1987) recommended the use of left-justified text on the screens when line lengths are approximately 20 cpl.

Length and complexity of passages

Surber (1992) had 52 college students read a long or a short passage and then take a multiple choice test containing main point and detail questions to investigate the effects of length of the passage on retention of information, reading speed and highlighting patterns. He found that subjects took almost twice the amount of time per word to read short passages compared with long passages. He hypothesized the reason behind this was engagement in a greater depth of processing for more elaboration. The conclusion of efficient reading with narrative passages compared to short news articles was also supported by Shaikh and Chaparro (2004). They found that test participants read faster and comprehended more while reading narrative passages compared with reading short news articles. We believe these findings had more to do with the participants' motivation and their understanding of the task, than with lower-level reading processes.

Design Guidelines

Here we distill the studies reviewed in this paper and derive recommended design guidelines for printed and on-screen text considering line length as a major factor affecting readability of the text.

Design guidelines for printed text

• Medium line lengths are optimal for efficient reading of printed text compared to very short lines of 5 picas and very long lines of 36 picas: for example 13 picas or 59 to 97 mm or 52 cpl (Tinker, 1963) and 3.5 to 5.5 in. (Burt, Cooper and Martin, 1955).

• As line length is decreased, the angular deviation of the return sweeps is increased, and thus the relative separation between the lines appears to increase and so readability is decreased (Luckiesh and Moss, 1941).

• Line width between 17 to 18 picas for double column printing in magazines, between 21 to 22 picas for a single column printing in scientific journals and between 21 to 22 picas for a single column printing in text books is considered to be optimal (Tinker, 1963).

• For 8- and 10-point font type sizes, line widths can be extended without the loss of legibility with two-point line leading (Tinker, 1963).

• Optimal line length for 10-point type size is about 80 mm, laying somewhere in between 59 and 97 mm (Tinker and Paterson, 1941).

• Line widths for multiple columns, 30 to 40 cpl (Lynch and Horton, 1999).

Design guidelines for on-screen text

• Medium line length of 55 cpl for ease of reading, better comprehension and better reading rates for on-screen text compared to very short lines of 25 cpl and very long lines of 100cpl (Dyson and Kipping, 1998).

• Readers prefer short lines of about 8 to 10 words or 45 to 60 characters long (Grabinger and Osman-Jouchoux, 1996).

• Single spacing produces more fixations per line resulting in fewer words per fixation and so total reading time is longer than with double spacing (Duchnicky and Kolers, 1981).

• Additional line spacing allows longer line length without sacrificing legibility (Lynch and Horton, 1999).

• Lines more than 9 or 10 words on an average must be leaded proportionately to the length of line. Lines with more words should have additional space between them (Hill, 2001).

• Make line spacing at least 1/30 of the line length to give a 2-degree downward angle for finding the next line (HFS, 1988).

• Line length should be around 30 times (between 20 to 40) the size of the font type (Bringhurst, 1992).

• Line widths for single columns, 65 to 75 cpl (Kahn and Lenk, 1998) or 45 to 75 (66 cpl ideal) (Bringhurst, 1992) or 100 cpl (Dyson and Kipping, 1997a).

• Line widths for multiple columns, 30 to 45 cpl (Galitz, 1993) or 30 to 40 cpl (Kahn and Lenk, 1998) or 40 to 50 cpl (Bringhurst, 1992).

• Fully justified text is hard to read because of hyphenation, uneven breaks in phrases and words and variable spacing. So left-justified text is considered easier to read with constant spacing (Grabinger and Osman-Jouchoux, 1996).

Conclusion and directions for future research

Line length of text read from printed and on-screen material affects various aspects of reading such as comprehension, reading rate, method of movement to navigate through the page (scrolling and paging) and eye movements. On the other hand different screen structures with different number of columns and screen size and various typographical factors like font type and size, line spacing, distance from the media, character spacing, length and complexity of passage affect what is considered to be an optimal line length. This is a very complex interaction, occurring on multiple levels. For instance, it seems as though line length affects method of movement, and method of movement affects reading speed.

Studies concluded that moderate line length in between 50 to 70 cpl are the easiest to read and users do not prefer extreme line lengths (very short or very long) while reading from screen. There was no significant effect of line length found on comprehension, though fast readers benefit from narrow columns with short lines due to specific reading patterns (with one contradictory finding). It seems as though good readers can, with columns that are almost exactly as long as their spans of apprehension, simply make downward saccades, down the middle of the column, apprehending all the line, left and right.

And so, let us conclude with the following set of recommendations.

1. Ensure that the width of the column, for running text, is at least as long as the span of apprehension for good readers, to maximize the use of their peripheral vision. The length of this span will vary with the easiness of the material being read, but a reasonable guideline would be 55 to 75 cpl.

2. Keep length of line moderate for ease of reading. It is difficult to make accurate return sweeps in longer lines, and shorter lines decrease the saccade length and don't let the reader make maximum use of their peripheral vision.

3. Longer lines are OK, especially for non familiar or otherwise difficult material and if speedy reading is desired. But again they increase the difficulty in making return sweeps.

4. Use the whole screen, regardless of the screen size. Leaving too much white space allows too much glare and distracts the reader. Consider a grayed background, to reduce glare, while still maintaining high levels of contrast with the text.

5. In using the whole screen, consider multiple (optimal width) columns, rather than one long column.

6. If using multiple columns, ensure that the columns are divided by enough white space to minimize the effects of visual masking.

7. Left-justify the text. If you wish to right-justify the text also, only do so with proportionately-spaced fonts.

8. Have the longer lines spaced (leading) proportionally to the line length to easily identify the beginning of the new line. No less than one-thirtieth of the line length.

9. Increase the line length with increase in distance from which the text is read.

References

Bernard, M.L., M. Fernandez, S. Hull and B.S. Chaparro. 2003. The effects of line length on children's and adults' perceived and actual online reading performance. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting,* 1375-1379.

Bridgeman, B., M.L. Lennon and A. Jackenthal. 2003. Effects of screen size, screen resolution, and display rate on computer-based test performance. *Applied Measurement in Education,* 16.3, 191-205.

Bringhurst, R. 1992. *Elements of Typographic Style.* London: Hartely and Marks Publishers.

Brown, W.S. 1970. Speed-reading made easy. *Journal of Typographical Research,* 4.1, 73-75.

Burt, C., W.F. Cooper and J.L. Martin. 1955. A psychological study of typography. *The British Journal of Statistical Psychology,* 8, 29-57.

Campbell, A.J., F.M. Marchetti and D.J.K. Mewhort. 1981. Reading speed and text production: A note on right-justification techniques. *Ergonomics,* 24, 633-640.

de Bruijn, D., S. de Mul and H. Van Oostendorp. 1992. The influence of screen size and text layout on the study of text. *Behaviour and Information Technology,* 11.2, 71-78.

Dillon, A. 1992. Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics,* 35.10, 1297-1326.

Dillon, A., J. Richardson and C. McKnight. 1990. The effects of display size and text splitting on reading lengthy text from screen. *Behaviour and Information Technology.* 9.3, 215-227.

Duchnicky, J. and P. Kolers. 1983. Readability of text scrolled on visual display terminals as a function of window size. *Human Factors,* 25, 683-692.

Dyson, M. and G. Kipping. 1998. Exploring the effect of layout on reading from screen. *Proceedings of EPi98 and RIDTi98 conferences,* 294-304.

Dyson, M.C. and M. Haselgrove. 2001. The influence of reading speed and line length on the effectiveness of reading from screen. *International Journal of Human-Computer Studies,* 54, 585-612.

Dyson, M.C. and G.J. Kipping. 1997a. The effects of line length and method of movement on patterns of reading from screen. *Visible Language,* 32.2, 150-181.

Dyson, M.C. and G.J. Kipping. 1997b. The legibility of screen formats: Are three columns better than one? *Computers & Graphics,* 21.6, 703-712.

Galitz, W.O. 1993. *User-interface design.* Boston: QED publishing group.

Gough, P.B. 1972. *Language by Ear and by the Eye,* One Second of Reading. Cambridge: MIT Press.

Gould, J. and N. Grischowsky. 1984. Doing the same work with hard copy and with cathode-ray-tube (crt) computer terminals. *Human Factors,* 26, 323-337.

Gould, J. D., L. Alfaro, V. Barnes, R. Finn, N. Grischkowsky and A. Minuto. 1987. Reading is slower from crt displays than from paper: Attempts to isolate a single-variable explanation. *Human Factors,* 29, 269-299.

Grabinger, R.S. and R. Osman-Jouchoux. 1996. *Cognitive Aspects of Electronic Processing,* Chapter, Designing screens for learning. Norwood, NJ: Ablex Publishing Corporation.

Hansen, W.J. and C. Haas. 1988. Reading and writing with computers: A framework for explaining differences in performance. *Communications of the ACM,* 31, 1080-1089.

Hedrick, C. 2002. *Guidelines for Typography in NBCS.* New Brunswick Computing Services (NBCS), Rutgers, The State University of New Jersey.

HFS. 1988. American national standard for human factors engineering of visual display terminal workstations (ANSI/HFS 100-1988). *Human Factors Society.*

Hill, B. 2001. *The Magic of Reading.* Redmond, WA: Microsoft Corporation.

Horton, W. 1994. *Designing and Writing Online Documentation,* Chapter 8: Display. New York: John Wiley & Sons.

Kahn, P. and K. Lenk. 1998. Principles of typography for user interface design. *Interactions,* 5.6, 15-29.

Kolers, P.A., R.L. Duchnicky and D.C. Ferguson. 1981. Eye movement measurement of readability of crt displays. *Human Factors, 23,* 517-527.

Kruk, R.S. and P. Muter. 1984. Reading of continuous text on video screens. *Human Factors,* 26, 339-345.

Luckiesh, M. and F.K. Moss. 1941. The effect of line-length on readability. J*ournal of Applied Psychology,* 25, 67-75.

Lynch, P. and S. Horton. 1999. *Web Style Guide: Basic Design Principles for Creating Web Sites.* New Haven, CT: Yale University Press.

McMullin, J., C.K. Varnhagen, P. Heng and X. Apedoe. 2002. Effects of surrounding information and line length on text comprehension from the web. *Canadian Journal of Learning and Technology,* 28.1, 19-29.

Mills, C.B. and L.J. Weldon. 1987. Reading text from computer screens. *ACM Computing Surveys,* 4, 329-358.

Mourant, R.R., R. Lakshmanan and R. Chantadisai. 1981. Visual fatigue and cathode ray tube display terminals. *Human Factors, 23.5,* 529-540.

Paterson, D.G. and M.A. Tinker. 1942. Influence of line width on eye movements for six-point type. *Journal of Educational Psychology, 33,* 552-555.

Paterson, D.G. and M.A. Tinker. 1943. Eye movements in reading type sizes in optimal line widths. *Journal of Educational Psychology,* 34, 547-551.

Perera, S. 2004 (May). LPfont-An investigation into the legibility of large print typefaces. *Information resource for people working in the field of visual disabilities.* Retrieved October 27, 2004, from http://www.tiresias.org/fonts/lpfont/report/lpfont3.htm

Rayner, K. and A. Pollatsek. 1989. *The Psychology of Reading.* Upper Saddle River, NJ: Prentice-Hall Inc.

Reynolds, L. 1980. *The psychology of written communication,* chapter Teletext and viewdata - a new challenge for the designer, 207-224. New York: Nichols Publishing Company.

Shaikh, A.D. and B.S. Chaparro. 2004. The effects of line length on reading speed and comprehension of short and long passages. Software Usability Lab, Wichita State University. (Unpublished).

Spencer, H. 1968. *The Visible Word.* London: Royal College of Art.

Surber, J.R. 1992. The effect of test expectation, subject matter, and passage length on study tactics and retention. *Reading Research and Instruction,* 31, 32-40.

Tinker, M. and D. Paterson. 1929. Studies of typographical factors influencing speed of reading: Length of line. *The Journal of Applied Psychology,* 13.3, 205-219.

Tinker, M.A. 1963. *Legibility of Print.* Ames, IA: Iowa State University Press.

Turvey, M.T. 1973. Peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review,* 80, 1-52.

Author Notes

Anuj A. Nanavati is a master's student and research assistant in the School of Information at The University of Texas at Austin where he is taking a multidisciplinary approach with focus on usability engineering and user-centered design. He has worked as a system analyst/designer, user interface designer and conducted various user studies and usability evaluations. Anuj's major research interests are ubiquitous user interface design and search interfaces. He received his bachelor's degree in computer science from Nirma University, India.

Randolph G. Bias is an Associate Professor in the School of Information, The University of Texas at Austin. Dr. Bias worked in industry for twenty-five years as a usability engineer, at Bell Labs, IBM and BMC Software and as a consultant. He joined the School of Information in 2003 to research human information processing and human-computer interaction. Randolph has written over fifty technical articles and recently published a 2nd Edition of Cost-Justifying Usability (R.G. Bias and D.J. Mayhew, editors.). He is a vigorous advocate for designing technology to fit the user.

R a t i o n a l i z i n g
D e s i g n Sensitivity

Mandar S. Rane

Abstract

The advent of digital technology has created a radical shift in execution tools within the realm of graphic design. This has turned out to be a blessing and a problem in relation to the context and the user. Working with traditional tools, like the brush, ink, paper or pencil, which were simple to use, fortunately allowed errors while executing a task and indirectly promoted learning and sensitivity. More was understood by doing, sharing and observing each other, in comparison to computers, which nowadays, only permit individual participation from the user. Today's new tools and software offer error-free execution, making a task easier for an individual to create a layout, use a typeface, choose a color or an image with 'utmost insensitivity,' particularly among novice learners of the discipline. Apparently, it leads them to demand more rational approaches to understanding graphic design sensitivity.

Considering the above issue as an impediment to explorations in foundation design courses, this paper focuses on enhancing dual processing modes, i.e., vertical and lateral, in the context of contemporary design education, with the introduction of a tailored course for teaching visual order in two-dimensional graphic design. The reference is to issues in graphic design (typography) dealing with sensitivity; which at times seem difficult to rationalize. The method the experiment adopts is to create a problem for students based on certain predefined criteria, which needs to be fulfilled, plus ensuring them the freedom to generate solutions laterally. Once students transform the given problems into solutions, the instructor unveils the underlying principles of graphic design with simple analogies, finding parallels with elements and principles of design. Instead of adopting a conclusive approach of being right or wrong, that hinders exploration, students engage with the contextual nature of graphic design. The method also uses collaborative learning as a remedy to address the contemporary issue of individual submission to the dominance of the new media, in which beginners in the discipline struggle most of the time with the tool, rather than the task.

Introduction

Design assignments in foundation courses usually have a more artistic bent and at times completely lack a methodological and scientific approach. For this they are criticized. Debatable among different school of thoughts, this leads to dominance of one kind of thinking over another. Most of the time, this discourse is more vocal about which approach is more effective for modern graphic design education, rather than suggest empirical methods to create new paradigms for design education. To balance rational and lateral thinking modes, this paper proposes a method for criteria based design education (making it more comfortable for rational minds) as well as at the same time extending the range for exploration, that in turn encourages lateral thinking.

Before we begin, lets get a brief background of the audience for whom this assignment was created; in this case, first semester communication design students of the Department of Design at the Indian Institute of Technology Guwahati, Assam, India. IIT's are India's technical universities known for their engineering and computer science programs. There are two design courses running at two of the IIT's in the country. One of them is the Industrial Design Centre at IIT Bombay, a two-year post graduate program (M.Des) offering two streams, product design and visual communication, with multidisciplinary entrants, such as commercial artists, engineers and architects. While the other one is the Department of Design at IIT Guwahati, one of the newest design institutions, started in 1994. It is actually the only university level undergraduate design program in the country, again offering the same two streams. This one recruits students who expect to become computer scientists and engineers (Scotford, 2004). Therefore, teaching such rational minds at IIT Guwahati poses an unique challenge; students at IIT Guwahati find it difficult to express ideas without a rational basis. This is especially true of expressive and sensitive issues related to typography, in comparison with their counterparts in other art and design schools in India (Scotford, M. 2003). A clear need for a specially 'tailored' course, especially for students of

design at IIT Guwahati exists in order to enhance their dual processing modes (Kumar, 2000). This paper shares a new framework towards teaching courses in graphic design, one that provides an opportunity to exercise both thinking modes, analytical as well as synthetical, in the context of tradition and modernity.

For this assignment, we start by choosing a few students whom we shoot with predefined objectives (i.e., define a numerical order to the students), to take still pictures and later identify and discuss the relation of the analogy (of the presented solution) to elements of two-dimensional graphic design. The first part of the assignment works with a single student, to understand the principle of emphasis and its relation to surrounding white space. Further, the assignment deals with increasing the number of students to shoot still frames (pictures) as solutions to the posed problems in order to understand visual hierarchies. The whole class participates in the act, but only three subjects (students) are required in the rudimentary levels of the assignment. The rest of the class participates in the thinking process and sharing their thoughts to execute the solution. Casting the three subjects is very critical for the assignment, because that is what makes it challenging for the students as well as easy for the instructor, at a later stage, to relate it analogically to two-dimensional graphic design.

Experiment: (assignment)

For this assignment we need three subjects, it is necessary that the subjects have an incremental difference in relation to their physical features; additionally, they should form an obvious (ascending or descending) visual hierarchy, when they stand, together, as shown in figure 1. Rahul seems to be the most dominant, due to his physique (first in the hierarchy), Kshitij (middle in the hierarchy) and then Navendu (last in the hierarchy), as our third subject.
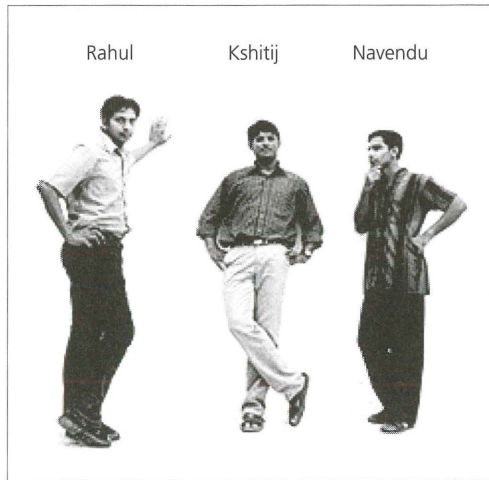

Rahul   Kshitij   Navendu

Figure 1. Participating students. From left, Rahul, Kshitij and Navendu

Task 01: One subject, create emphasis.

In task 01, students are given one subject to shoot in a pre-defined frame. Figure 2 was the outcome for their first frame. Here the subject 'Rahul' settles himself in front of the camera with a very casual gesture. The gesture was not considered important to the conclusion; it is neither right nor wrong in relation to the task. If the students are satisfied with the gesture, as the solution to the problem as posed, then they move on to the next task, which is not known to them.

Figure 2. Task 01

Results: When cross-checked with users (students), they naturally tend to look at the subject because nothing else is found in the space beside the subject. When a figure floats isolated in a sea of empty space, its presence is emphasized. Hence, in comparison to two-dimensional graphic design, layouts with a single element easily create emphasis. It automatically augments the presence of the subject, since it is the only element in that space. Quite simple! Students were told to compare the situation to a word, or a text written in a given format on paper. As the subject has a style, a character and meaning of its own (identity of the person), a font also behaves in the same manner with a style, character and meaning of the written word itself.

Task 02: One subject, create more emphasis in comparison to task 01.

In task 02, students shoot the same subject and are expected to create more interest (emphasis) in the frame compared to their prior task. Figure 3 was the outcome for their second frame. Here the subject 'Rahul' communicates a gesture, which depicts the act of listening, a more active and dynamic composition than the task performed earlier. All the students participate in rendering the gesture for the solution, i.e., students behind the camera and the students acting as subjects. The solution is an outcome of discussion and debate amongst themselves.
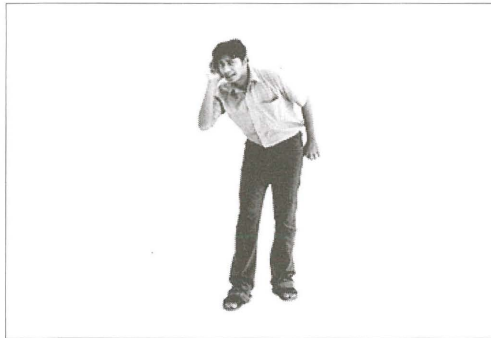

Figure 3. Task 02

Results: In this sequel, the solution is obviously more emphasized than the earlier one. The gesture, including body movements, makes it more dynamic; the subject appears to move in a certain direction. When both frames were compared by a user, task 02 naturally gained over task 01, supporting the assignment objective. Students were told to compare this to the task of modifying the written text, i.e., maybe changing it to italics or altering the orientation of the written word itself to make it more dynamic. The student's efforts were examined before they proceeded to the next task.

Task 03: One Subject, create more emphasis in comparison to task 02.

There is no major difference if task 02 and task 03 are compared. The scale of emphasis is comparatively less when both tasks are compared (see figure 4).
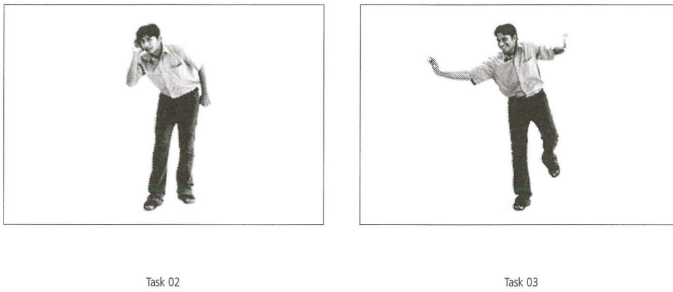

Task 02   Task 03

Figure 4. Task 02 and Task 03

Results: In task 03, students were unable to achieve any desired results for the objective of the given task. This means that when a user compared task 02 and task 03 *(figure 4)* neither frame took prominence over the other. Results varied. Students were not sure whether task 03 would attract all the attention. The situation was ambiguous, it did not fulfill the objective of increasing emphasis (interest levels) throughout the sequence.

Task 04: One Subject, create more emphasis in comparison to task 03.

In task 04 (figure 5), students were told to address the situation by increasing font size, covering the whole format and trying to get more attention with a single element. In an earlier case, we defined it as single font, maybe Italics, or bold, here it serves as an analogy to represent an expressive font, i.e., an expressive gesture.
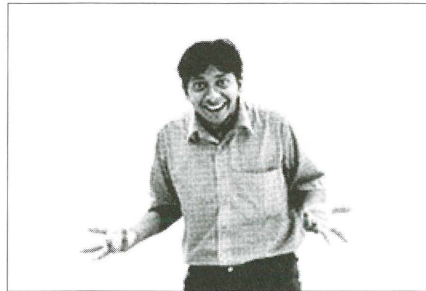

Figure 5. Task 04

A comparative study of all tasks (01-04) at one time, delivers a look through the whole aim of the exercise. In each of the frames in figure 6, the subject is compared to a written word. It tries to depict the relation among incremental gestures that have simultaneously affected the written word. The subject with its character, style and dynamism, symbolizes expression and emotion in a written word. This is easy to relate visually for a rational mind. Students themselves realized the essence of adding life (which is usually a sensitive term) through a font to the communicated message. The student begins to appreciate enhancement of the desired message with typography in the context of a single element; he comprehends the relation and importance of the white space and its discussion with the written text. It justifies for the student, typing a word on a software canvas as a meaningful act, and the need to be conscious while choosing a font. As an instructor, one can now easily comprehend why students were purposefully given a single subject in the preliminary levels of the exercise and the reason to restrict the exercise to black and white, instead of confusing them with a color variable, that could be dominant in catching attention.

Figure 6.

Task 05: Two subjects, create equal emphasis for both or make both of them equally unimportant, in other words create a cohesive whole.

Task 05 introduced an additional subject. Among the three subjects mentioned earlier, the last subject in the hierarchy was chosen for this task. Now students had to deal with two elements in the given space. In task 05 students were required to create equal importance for both subjects, in other words, neither of the two should be prominent. Students were expected to come up with a solution, which distributes almost equal attention to both the subjects, when perceived by the user.
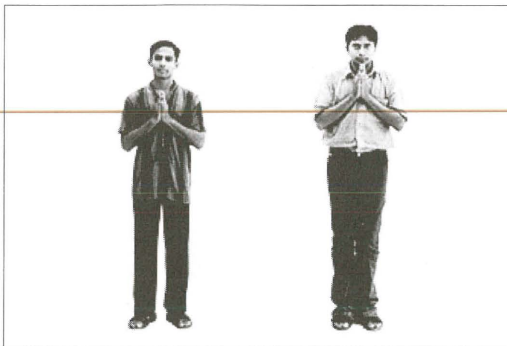

Figure 7. Task 05

Results: In the solution presented *(figure 7),* students tried to achieve uniformity in the gestures, so that, they could achieve equal importance to satisfy the pre-defined objective. Unfortunately, the task failed when tested with users; the right subject had more prominence over the left, with few exceptions. To discuss the analogy with graphic design, the result of one's communication cannot be ambiguous with reference to the defined objective. When one designs there is always a purpose to be achieved. If results happen to vary with users, then one needs to redesign to get closer to the defined objective, not with the sole aim to find the right answer, (vertical thinking), but to find more alternative solutions (lateral thinking) to get closer to the defined objective.

The solution was analyzed further for students. After the task, students were told if they had brought the subjects closer to each other, their individual identities would dissolve. Two subjects standing apart from each other, created distinct spatial differences and focused attention on their most prominent feature, difference in height, students were told to compare the situation with two words written on a white canvas and if the point size of one word had been larger, how could they achieve equal importance? Would spatial distances have mattered (for the two words to lose their identities and form a group as a whole), if the given task is creation of equal or unequal importance for both words? How are they to form a group of two elements as one, when both entities are distinctly different. Analogy applies to graphic design as well, where size, color, orientation, content and other related variables play an important role in creation of visual balance (symmetrical or asymmetrical) between elements. Students thought the use of similar gestures would fulfill the purpose.

(The reason for the choice of the first and the last subject amongst the order of three, was primarily to maintain the necessary challenge required for the given task.)
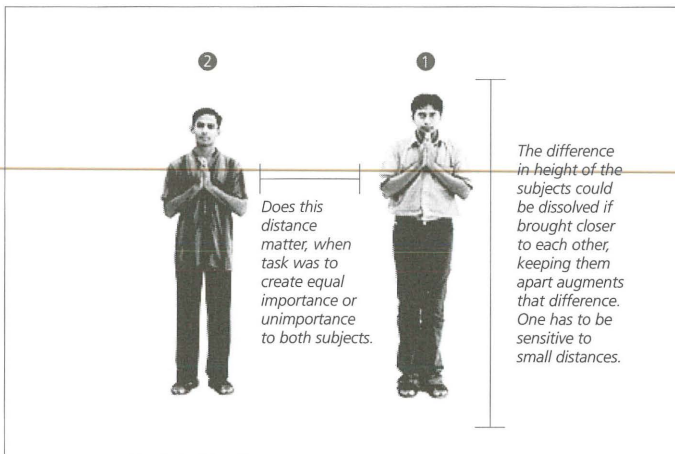

Does this distance matter, when task was to create equal importance or unimportance to both subjects.

The difference in height of the subjects could be dissolved if brought closer to each other, keeping them apart augments that difference. One has to be sensitive to small distances.

Figure 8. Task 05 (Analysis)

Compare the words in figure 9 with the subjects. Do the fonts resemble the properties of the subject, of it being condensed, tall and having an identity of the font face itself (in this case Zurich condensed bold) similar to the visual identities of those individual subjects? Another important point is the meaning of the written word, i.e., the content (in the context of the subject being a particular person and the written word "Rahul" transferring some information to the reader), if the reader personally knows the subject. The pragmatic expression of the task can be seen in figure 9 where the written words resemble the qualities of our subjects.
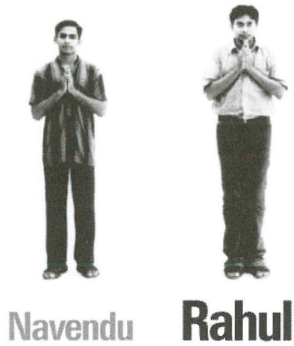

Figure 9. Task 05 Pragmatic Analysis

Task 06: Two subjects, follow the pre-defined order.

Task 06 begins with the concept of visual order. Here students are given a pre-defined visual order; now, in their solutions, adhering to the criteria, Rahul was to be perceived as No.1 and Navendu as No.2.
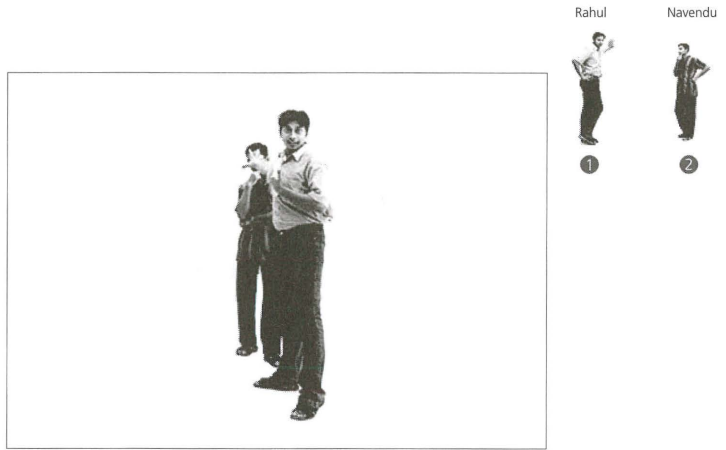

Figure 10. Task 06

Results: Yes, this time students got it right. When the picture was shown to users, Rahul was perceived as No. 1 and Navendu was No. 2. Purpose achieved. Quite an obvious solution, because Rahul is stronger than Navendu and moreover, subject No. 1 overpowers subject No. 2 by covering his face. Students were told, if more options were explored, without hiding the comparatively weaker subject's face, then it would have been considered a creative solution. Students were told to compare the situation with a layout where one word is written in red and the other in gray. Naturally the user would look at the red word first and the grey later (excluding the meaning of the word). Students were introduced here to the concept of doing the 'obvious,' as in, everyone's first solution to a given problem. Thus the instructor can rationalize the term 'obvious' and then refer it to novices, when they execute actual design solutions.

Task 07: Two subjects, order of perception for the solution was reversed in comparison to the previous.

This task was really tough; now students had to reverse the order and make the stronger subject (Rahul) No. 2 in the order of perception. In other words, students had to attract the user's attention towards Navendu, the weaker subject first and then towards Rahul.
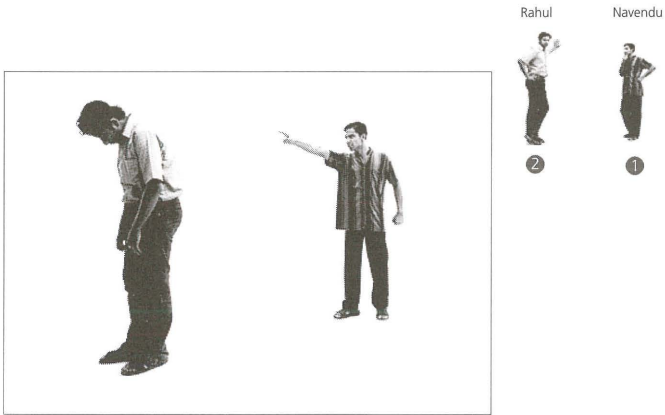

Figure 11. Tas k 07

Results: Excellent solution. When presented to the user, Navendu was perceived as No. 1 and Rahul as No. 2. Purpose achieved. However, there was an exception to the statement, one of the users perceived the order, exactly in reverse. The results varied with different users. Students were questioned — was it due to the emotional personality of the user, or was the reversed order of perception in opposition to the pre-defined objective? All users needed to perceive the information in the same order — this was not a coincidence. Students were expected to design information, to perform a well planned act. Now the communication was ambiguous.

Task 07: (Repeat) Reducing ambiguity with the help of elements of design.

The solution criterion is restructured with help from design elements. Radiating lines emerge from Navendu; this helps him dominate Rahul, creating a visual preference for him. We are trying to demonstrate the analogy that can be applied to graphic design problems, especially with consideration for rational minds.
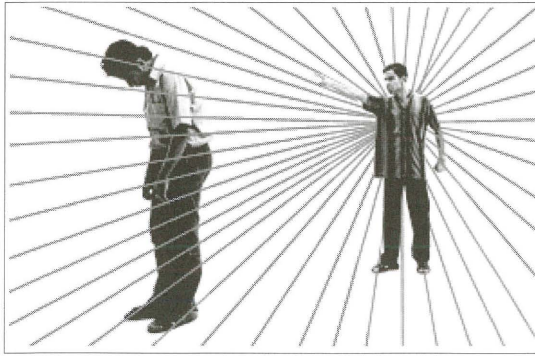

Figure 12. Task 07 (Analysis)

Figure 12 reduces the ambiguity of the desired order or reduces confusion along linguistic boundaries, where words like "maybe" or "unsure" overrule the defined objective. We are now quite definite about Navendu as No. 1 and Rahul as No. 2 in the order of perception. Students were exposed to the pragmatic aspects of design elements to enhance messages in graphic design. Thus the instructor explains how, with the help of line, color, size, shape, etc. (elements of design), one is able to create a structure or order for the visual message in graphic design, making it convenient for the reader to decode information in the order we want the message to be received.

Task 08: Three subjects, create solution to perceive the pre-defined order.

New subject (Kshitij) was introduced and now the task demanded creation of visual order among three subjects to match the visual order as defined to the right of the picture (see figure 13).
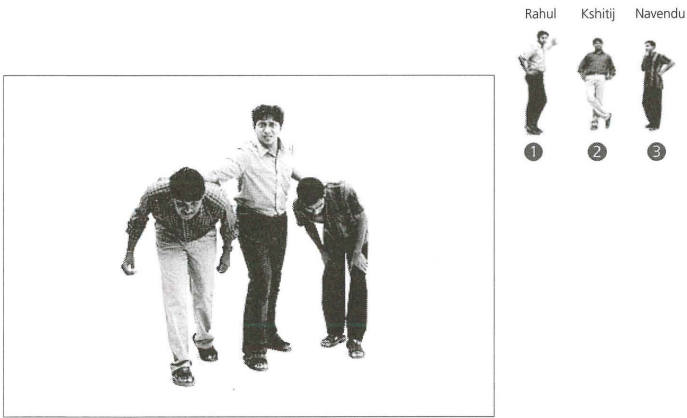

Figure13. Task 08

Results: This is a commendable effort from the students. The solution for task 08 has no uncertainty in perception for the pre-defined visual order. Rahul acts in control of the other two subjects, is perceived as No .1, and Kshitij, who is physically pushed forward, is perceived as No. 2 , whereas Navendu hides his face to step-down to No. 3 in the visual order. The subject in the center actually directs the order of visual perception and guides us to follow the rest. Students were told to build their solutions on the basis of visual flow, which can lead the user through a designed order (in this case, through gestures). The instructor explains the task, showing an analogy of flow in graphic design. He supplements his statements to find similarity with examples of reading direction in graphic design.

Task 08: (Analysis)

Substitution of the subjects with written text in order to rationalize the analogy depicting direction of reading. Chunks of texts are substituted for each subject.

Figure 14. Task 08 (Analysis)

According to the visual hierarchy perceived, the identity of the person acts as a heading of the text. Students were told while designing a layout, brochure or a pamphlet, font size and orientation of the elements, leads a user through a preferred visual order (devoid of the content). If the same 'context' was treated with content of the written words (i.e., the meaning of the written text, itself acting as an ingredient to lead towards a visual order) e.g., a very catchy headline, plus, with other decisive elements, such as a strong color, the hierarchies would have been different, compared to as they are presently perceived. Understanding the subjective nature of the discipline, students internalize the subjectivity of design solutions offered by different contexts and their appropriate solution, rather than consider it an irrational activity. Rational minds always search for absolute values.

Task 09: Three subjects, create solution to perceive a reversed order in comparison to task 08.

This task was more challenging. The pre-defined visual order in task 08 is now exactly reversed. It was interesting to make the weaker subject as No.1 in the order of visual preference.
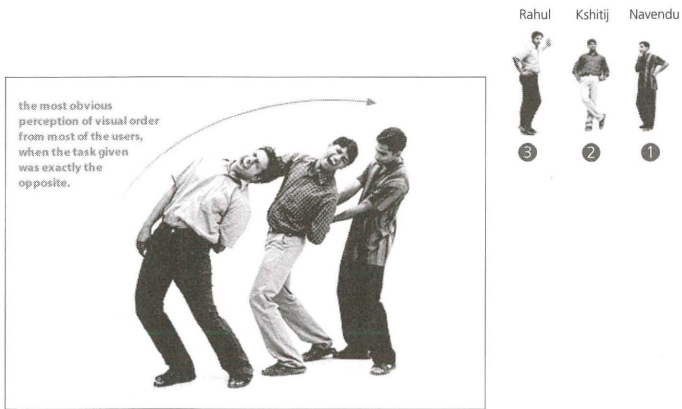
Figure 15. Task 09

Results: The idea was unique and expressive too, but could not meet the requirements of the posed problem. Unfortunately for the students, the exact reverse order was perceived by most of the users. Task 08 left an impression on their rational minds, regarding who is in control. Rahul, who was supposed to be perceived as No. 3, was perceived as No. 1 in the visual order. His expression and posture facilitated his dominance in the visual order, against the rest. Students were told, that if Rahul had taken Navendu's place (the subject who was supposed to be perceived as no. 1) this could have met the demands of the objective they were seeking. Lets understand this with an example.

The example shown in figure 16 relates to the fact, that a more attractive subject's gesture (in this case, Rahul), orientation or posture can be compared to an unusual or dominant element in a typographic layout, such as text written diagonally in the adjacent typographic layout. A bold font face and increase in font size attracts and captures more attention in comparison to the remaining elements. Students understood such miscellaneous factors (in pictures, gestures, layout, font face, size and value) act together towards the desired solution, and how it is necessary for a student to become sensitive to
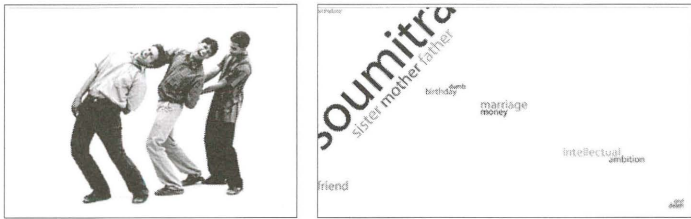
Figure 16. Task 09 Comparative analysis of task 09 with a typographic layout to find parallels, in order to strengthen the analogy

these aspects. The significance of this task can be understood as; Principles — contrast, emphasis, balance, movement, rhythm, etc. and elements of design — color, value, texture, shape, form, etc. act in a similar fashion, as variables to be balanced in different contexts for each graphic design problem. Visual order therefore can be considered as a rational criteria to confirm your design solution, based on the objective of the communication, but may not function as a formula or method. Creation of interest or persuasion in varied contexts encountered while designing requires both vertical and lateral thinking. So, graphic design involves both logic as well as sensitivity, as two sides of a single coin.

Task 10: 10 Subjects. Follow the predefined order and execute the solution.

Quite a difficult problem was posed for the students; the format or the space was the same, but the number of subjects was increased to ten. A new addition to the subjects was a female. Students were given the choice to pre-decide their own visual order for the ten subjects and then execute the solution to achieve that goal. The aim of this task was to challenge students with complex problems and help them internalize order and chaos in graphic design.
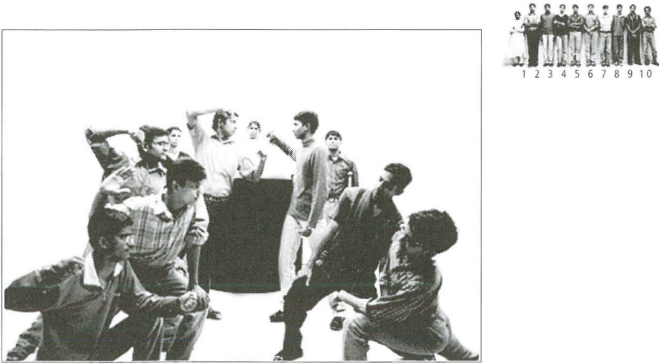
Figure 17. Task10

Results: This time students did not succeed in their task. They experienced the difficulty of handling too many subjects in a given space. They were told to relate to the situation by thinking of subjects as amount of information. To crunch a lot of information into a given space, and still make it comfortable for the user to extract information, was well understood by the students. Simultaneously, the act of designing information was described as a planned activity rather than random.

Task 10: Analysis

"Alcheringa" is the cultural festival of IIT Guwahati held every year, with many events and sponsorships. A billboard announcing the event tries to find its analogy with ten subjects and their visual order.
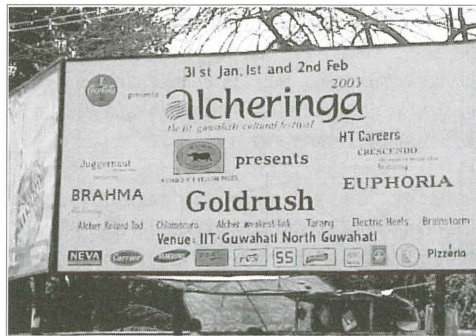
Figure 18. Hoarding displaying Annual college festival, Alcheringa 2003 in Guwahati city, Assam, India

The exercise served its purpose to explain the syntactics of design elements affecting the resulting solution. The example in figure 18 explains the lack of order, too many elements results in chaos, make it difficult for the user to perceive the message. The whole exercise confronts us with the question of whether communication of information can afford to be ambiguous in nature, when designing billboards, brochures, pamphlets or even websites. Do the subjects in the tasks resemble the elements present on the billboard? The instructor encourages contemplation of this to make them sensitive and conscious while viewing other visual communications they might encounter in the future.

Conclusion

Its all information and hierarchies of information we arrange around us. As designers, our sole aim is to facilitate ease in communication for the user. Creating interesting visual order to persuade the user to comprehend sometimes hidden visual order or at times to guide him through a well-defined visual order. In the tasks executed by the students, subject No. 10 can be considered as the base line for some information, subject No. 2 can be considered as a headline. If considered in a layout, subject No. 1 can be substituted as a visual element for a poster. If this information was in an interactive mode, one could imagine these subjects as clicks or hyperlinks of a website. When you click over a subject, the subject talks back to you or performs a gesture, which could resemble a mouse-over in a website. In this case the visual order of preference could be considered dynamic rather than static. Depending on the choice of the user, visual order could be shuffled to animate itself to the preference of the user, imitating interactive menus. For example, if we had to shoot a video with the subjects used in the tasks, and if we click over any subject, others can automatically rearrange themselves by physically moving into places to represent the descending order. This could lead to new paradigms to understand visual order with dynamic data through collaborative learning, like a theatrical performance.

Reason seems to be a prime concern of today's modern world. The link from theory to practical application is difficult for students to grasp, because the nature of the design discipline itself is very subjective and contextual in its solutions, which sometimes demand more lateral thinking when juxtaposed against vertical. This link becomes difficult as the student tries to seek rules (formulas) that can be applied to the problem solving activity, trapping him- or her-self in the dilemma of right and wrong in design. Quantitative results seem more pleasing in such cases compared to qualitative, since they are easily articulated by a rational mind. In this course, instructors face arguments regarding the precision of the analogies, for the reason that most of these tasks are analogous; i.e., understood by doing, seeing and comparing and not based on results, translated numerically. Comparison provides insights, not results, as they are based on learning through perception. Design assignments today, are currently under pressure to rationalize and at times, reason out acts as an incentive to make someone work towards a goal. Exploration finds it difficult to flourish within such environments, as they are based on a foundation of interests, rather than reasons. Therefore new assignments balancing both modes of thinking need to be designed. The assignment presented here was a result of teaching specific students, but may hold true for various facets of the discipline. Considering the current context, where choice of font is available at a mouse click, the exercise becomes extremely important to act as a rudimentary level course in order to familiarize novices with the lateral side of the discipline, yet satisfy students with their rationale queries and the medium of analogy through which instructors can answer them fully.

In most of the earlier traditional practices, "not knowing" was never related to the domain of 'tools,' but always focused on the 'task.' With new tools of execution, maximum effort is spent on understanding the tool rather than involving oneself in the depths within the task. The task concludes as a result of the novice's knowledge about the tool, as a result the communication objective as well as learning becomes distorted. The objective of this exercise is to encourage collaborative environments for learning, where physical participation, dialogue and involvement with a purpose (task) are of prime concern. This experiment is an attempt to balance the future of design education, namely traditional and modern ideas, through recognizing and associating rational thinking with exploration and creativity.

References
Aicher, Olt. 1994. *Analogous and Digital.* Germany: Ernst & Sohn.
Kumar, H. Vyas. 2000. *Design The Indian Context.* India: National Institute of Design, 17-19.
Scotford, M, 2004. Design Education in India, *In Form,* 14.2, 22.
Scotford, M, 2003. Function and Expression. *Visible Language,* 37.2, 128.

Author Note
Mandar S. Rane is an assistant professor at Department of Design, Indian Institute of Technology (IIT)Guwahati, Assam, India. He has a Bachelors in Applied Art from J.J. Institute of Applied Arts, Bombay and M.Des in Visual Communication from Industrial Design Centre (IDC), IIT Bombay. He has worked in the areas of web, multimedia and print design. Prof. Kirti Trivedi inspired him to teach. His recent research focuses on creating new methods to understand graphic design sensitivity by employing a rational approach. His interests include design education and semantic issues related to visual communication. www.mrane.com

Abstract

Though William Blake is a central figure in the academy, there
is one particular area of his work that receives little attention: the marginalia. And
when annotations are incorporated into Blake studies, scholars tend to turn for quota-
tion to typeset Blake editions, which do not communicate the visual complexity of the
annotations. In addition to being visually dynamic, the marginalia provide evidence of
Blake engaging the printed book of the late 18th and early 19th centuries, and are thus
part of his work as a book-maker. Blake's books are radically different technologies for
representation than are the books he was annotating. Further, Blake's experience as a
reader and annotator are reflected in his poetic universe, in which readers, writers and
books figure so prominently.

U n R e a d i n g[1] *W i l l i a m B l a k e ' s M a r g i n a l i a*

Jason Snart

*Despite the degree to which William Blake has become a cen-
tral figure in the academy,* there is one particular area of his work that could
benefit from further attention—or at least a different kind of attention than it currently
receives: his marginalia. Study of the original annotated volumes sheds light on many
of the issues that are central to Blake scholarship, including textuality, authority, syste-
maticity and materiality. The annotations are themselves significant beyond their "con-
tent" (what we might call their referential or semantic value), for they are the traces that
remain of Blake engaging with the printed page of his day; that is, the page produced
from moveable type. The marginalia function as part of Blake's working through the
possibilities for a radically different kind of representational technique and technology.
Further, Blake's ideas about art and about what his art was supposed to do (its episte-
mological and ontological status in the world) stem in part from his experience with the
conventionally printed books that he owned and borrowed from others.

I look at the way in which Blake's "anti-Newtonianism"—his opposition to finality, univocality and fixity—emerges from the marginalia, as it does from his other works, suggesting the degree to which Blake was working through problems of representation, and the role played by textuality and materiality in the production of meaning, as much in the marginalia as in his other artistic productions. Blake's relation to Newton has been of central concern to Blake scholars for some time, thanks in large part to Donald Ault's *Visionary Physics: Blake's Response to Newton,* published in 1974. But Blake's marginalia have not been adequately integrated into this discussion. My approach to the marginalia and to Blake's anti-Newtonianism as two aspects of the larger issue of Blake engaging the problems and possibilities associated with representation (with the disjuncture between "imagination," as Blake called it, and materiality, for example) allows us to see how the marginalia are, indeed, central to the development of Blake's art. *Additionally, I want to uncover the way in which Blake scholars (and scholars in other fields as well) have engaged the marginalia with little attention to their textual variability.* The tendency to use the marginalia for their content alone is a product of there being no reasonably accessible edition (digital or print) that provides high-quality, photographic reproductions of the annotated volumes. Organizations like the William Blake Trust and projects like the William Blake Archive are working to provide scholars access to versions of Blake's work that are closer, in both visual and verbal dimensions, to Blake's originals. The marginalia could certainly benefit from this kind of attention and treatment as well.

Currently, however, Blake specialists, general readers and students alike, are left to read and to experience Blake's annotations as they are printed in typeset editions like those of Sir Geoffrey Keynes, G.E. Bentley, Jr. and David Erdman. As I suggest below, having recourse only to typeset editions has had (and continues to have) serious consequences for Blake scholarship. In general, access to the marginalia in typographic format alone has hindered how the marginalia can be integrated as visually important documents into Blake studies. Yet even as the annotated volumes are scanned and digitized, and photographic reproductions made available, it will be important to approach them in such a way that certain traditional biases (towards "ideas" versus "materiality," or "content" versus "form") do not preclude our seeing some very important

aspects of the marginalia. My objectives here include drawing attention to the way in which typesetting Blake's marginalia makes certain modes of scholarship less or more possible and certain kinds of issues less or more accessible. And I explore the kinds of issues that do arise when annotations can be studied in their original context (or at least in high-quality reproduction).

There is, in fact, a history of attempts to make the margi-nalia available to readers in a form that retains something of their original look. In 1947, a letter arrived to Mr. Geoffrey Keynes from Josiah K. Lilly, Jr. It is a response to an earlier request by Keynes to have the volume of Francis Bacon's *Essays Moral, Economical, and Political* (1798), which contains Blake's annotations, in some way reproduced so as to facilitate research and study. Lilly's response is straightforward: "it is quite out of the ques-tion to do as you request" (Lilly). Lilly continues, "there are something over a thousand entries in Blake's holograph throughout this book, pres-ent on over a hundred pages, and it would be about as expensive to have these reproduced as the book is worth!" The estimation of a thousand entries is off the mark. Even a generous count would suggest that there are no more than two hundred entries by Blake—some as short as "A Lie!" Lilly is correct, however, in asserting that "the book is in very fragile condition." When I studied the volume in 2001 it was on the verge of coming entirely apart due to deterioration of the spine. I do not believe that the book could have been in good condition even in Lilly's time. He is also right that "the Blake entries are in pencil, some of which are very difficult to read except with a mag-nifying glass." "I am just sure," he writes, "that several of the entries would thus not reproduce well at all" (Lilly).

Since in his final paragraph Lilly agrees to sell the volume to Keynes for "$1,000 (U.S. dollars)," it is clear that Lilly stood to profit by asserting that the pages of the book could not be copied. The letters that Keynes and Lilly exchanged stand as testament to early efforts to study Blake's marginalia, and to the difficulties even then presented by the mate-riality, and the economics, of the situation. There really seemed no way around the necessity of having the actual book if one wanted to study the marginalia.

The Bacon volume was sold for a thousand dollars in 1948, as recorded in a letter of 3 August, 1948 from a librarian at the Yale University School of Medicine to Lilly. A handwritten note at the bottom of the page from the librarian to Keynes promises that "The library will dispatch the volume to you as soon as received" (Keynes correspondence). The letter with the note to Keynes is a copy of the one actually sent to Lilly, which of course did not have the note written on it. The Bacon volume is now in the Keynes collection at University Library, Cambridge. The letters are included in the book box which contains Bacon's *Essays* (Keynes U.4.20). The volume was dispatched to Keynes, and by 1957 Keynes published *The Complete Writings of William Blake* with Nonesuch Press; the volume includes the annotations to Bacon, along with other of Blake's annotations, though the annotations are typeset, laid out very much as they would be in later editions like Erdman's 1982 (and 1988) *Complete Poetry and Prose of William Blake.*

Annotations as editorial construction

In these editions, Blake's annotations are accompanied by a piece of the original text to indicate their "position" on the page and thus their "relation" to the original text. Erdman describes the format this way: "Excerpts from the works marked and annotated by Blake are followed by Blake's remarks in larger type" (583); and "excerpts from the [...] works annotated have been trimmed to the bare minimum necessary to show the immediate context of Blake's remarks" (883). It may be that such a format obscures more than it reveals, however. Certainly Erdman's layout does reveal one possible relation between text and annotation, but that relation remains, in many cases, an editorial construction. There are moments, for example, when such a format is quite misleading, since annotations, when they are typeset, can be placed with portions of text to which they may not directly refer, but to which they are visually close in the original. Or, as is often the case, one annotation may make a more general comment than is suggested when the annotation is paired with a short block of the original text. For example, Erdman provides the following from Blake's annotations to Watson's *Apology for the Bible* "(the smaller type is material from Watson; the larger type is Blake's annotation; the page designation, square brackets and material therein, are Erdman's)":

[BISHOP WATSON☐S PREFACE]

PAGE [*iii*]
... the deistical writings of Mr. Paine are circulated ... amongst the unlearned part of the community, especially in large manufacturing towns; ... this Defence of the Revealed Religion might ... be efficacious in stopping that torrent of infidelity which endangers alike the future happiness of individuals, and the present safety of all *christian states*. ...

Paine has not Attacked christianity. Watson has defended Antichrist.

(Erdman, 612)

However, the annotation actually appears at the top of the page, not below Watson's passage. A double horizontal line (which is part of the page layout, not added by Blake) separates Blake's note from the text. Blake has underlined "christian states" but there is no textual mark that links the phrase with the annotation (elsewhere in the marginalia, Blake will use lines or brackets to indicate passages he is annotating). Since the note appears at the top of the page (and on the first full-text page of the volume), it would seem quite likely to be a general statement, not one necessarily tied to a particular passage. We can see quite clearly in this case how the annotations as presented by Erdman are often a purely editorial construction.

The inability of typesetting to convey the position of annotations on the page is problematic, though not the only limitation of the anthological format. As Keynes explains of his typeset annotations to Bacon, "Words underlined by Blake are printed in italic." And "Passages from Bacon are in smaller type" (Keynes 397). There is little sense to be had of the actual layout and look of individual pages, the size or style of the annotations or their legibility.

Legibility and audience

Indeed, legibility, which Lilly alluded to in the 1947 letter cited above, often presents great difficulty when studying the originals. That some of Blake's notes (but certainly not all) are nearly unreadable invites questions as to who, if anybody, Blake imagined would (or could) read his annotations. The annotation on page 1 in Bacon's *Essays,* for example, which surrounds the upper right corner of text, is written in pencil that is unsharpened and light (perhaps from having faded over time, though not all Blake's pencil annotations are so light). Erdman transcribes the passage, "But more Nerve if by Ancients he means Heathen Authors" (621). Keynes offers the same, but adds a period at the end of the statement, although there does not appear to be a period in the original (Keynes, 397). In this case, a word like "Ancients" (and even "Heathen" to some extent) is not decipherable in and of itself. The pencil strokes are simply too close together, and the pencil too dull. Under magnification the pencil marks blur to an even greater degree, making the words less, not more, readable. It is only by context that one could guess at the correct word; that is, the text being annotated provides clues as to the meaning of otherwise illegible words that appear in the margin. This annotation ("But more Nerve if by Ancients he means Heathen Authors"), for example, appears next to part of Bacon's text: "there be not so much blood in them as was in those of the ancients" (Erdman, 621).

This begs the rather intriguing question of just what such an annotation meant to Blake. Who was he writing for, if he was writing for any one person or audience all the time? It is worth noting that elsewhere in the marginalia Blake addresses readers (or a reader) directly. In Sir Joshua Reynold's *Discourses,* for example, Blake writes, "The Reader must Expect to Read in all my Remarks on these Books Nothing but Indignation & Resentment." Where the typeset marginalia present the reader with a sense of relative homogeneity among Blake's annotations, what we actually find in the volumes themselves is a wide diversity in terms of manuscript style, the care with which some annotations are written relative to others, and even choice of annotating medium (sometimes ink, sometimes lead pencil, sometimes both). Do we afford equal weight to all annotations regardless of their material diversity on the page? In the attempt to derive from the marginalia Blake's theories and beliefs, does the illegible annotation mean as much as the note Blake has written carefully and then traced over in permanent ink? My own sense is that definitive answers may not exist. A system for understanding the visual implications of the marginalia relative to their semantic importance is neither attainable nor desirable. Presently, however, the questions themselves, along with the pursuit of many possible answers, is sufficient, since they have not, so far as I know, been posed before.

Can context be ignored?

The implications for Blake scholarship are crucial. When scholars use the annotations at all, they tend to treat them as they would any other of Blake's published work. Northrop Frye, in his *Fearful Symmetry,* goes so far as to assert the safety of quoting marginal material: "it is quite safe to use these quotations [from Blake's annotations] here" (15). It is important to consider what "safe" implies, though Frye's explicit point is that Blake's notations about Reynolds' "theories of painting" are relevant in terms of understanding Blake's theory of knowledge (14-15). Safety in this case suggests that the marginalia can be removed from their context in the figurative sense of transposing discussion of painting to discussion of epistemology. I am interested, however, in reaffirming context (not least on a literal level), particularly when it comes to deriving a "theory" or very general philosophical principle. Attention to context reveals even the degree to which the space of the margin itself is materially restrictive, quite literally constraining the degree to which Blake can offer, and then perhaps qualify, a response. Given the textual variability of the annotations themselves, along with their

materially constrained context, it perhaps becomes more difficult to assert a consistent Blakean "theory of knowledge," derived from quotations deemed "safe" by virtue of the tautology that they seem to fit the general, unified model. Frye's need to reassure that quotation from the marginalia is "safe" stands, perhaps ironically, as fair warning that such safety is not as straightforward as it may seem.

What do marginalia represent?

Indeed the apparent safety of the annotations when type-set—that is, their apparent regularity and legibility—leads to arguments which treat the annotations as a consistent "body" of work, just as Blake's *Jerusalem,* for example, is treated as a single work. The result has been arguments asserting that Blake's "philosophy," or indeed that the "whole of Blake" is available to readers in the marginalia. Since these arguments tend to depend on the regularized, typeset versions of the marginalia as printed in editions like Erdman's, they miss how textuality and materiality inform meaning.

Thomas McFarland's "Synecdochic Structure in Blake's Marginalia" provides a good example of the way in which a certain kind of academic production is made possible by the regularized, typeset form in which the marginalia are currently available. As his title suggests, McFarland sets out to uncover in Blake's annotations the "whole of Blake" (79).[2] McFarland asserts that in the books Blake annotated in his lifetime, "the whole intellectual ethos of Blake stands revealed, if not in comprehensive detail then by implication" (76), and that readers "glimpse the part that represents the whole of Blake" (79).

Paul Valéry, writing about the marginalia of Edgar Allan Poe, asserts that "marginal notes represent part of the notations of pure thought" (177). Subscribing to Valéry's notion of the marginal note as pure thought, McFarland argues that "the marginal notations [...] reveal Blake's intellectual essence with peculiar directness" (76-77). Quoting Poe, McFarland writes that "in marginalia an author speaks 'freshly–boldly–originally'" (77). Working from Erdman's 1982 edition of Blake, McFarland cannot know that in Reynolds *Discourses,* for example, Blake made marginal notes in pencil and later traced some over with ink (an issue I take up below). How "fresh" is a traced over marginal note? How "bold" is it to leave some untraced? And it is hard to know what "original" might mean exactly. Was Blake writing things he hadn't thought of before?

Things no one else had thought of before? Or just things he hadn't written down before? In any case, even McFarland admits that "marginalia [...] invade their host text" and "the marginal notation forces open the text" (78). It seems problematic to argue that a marginal note is both original but also deeply dependent on a host text. It also seems clear, based on the material variety of the notes themselves, that Blake is inconsistent in terms of the degree to which he thinks through an "independent" response to the text at hand, versus the degree to which an annotation is an immediate reaction entirely dependent on what might actually be a limited portion of the text he is reading.

Alternately, Blake's annotations to Swedenborg's Divine Providence *reveal that he must have read some volumes through before annotating.* For example, in his note to aphorism #185, Blake directs attention to #69, but also to "329 at the End," "& 277," "& 203." In #69 he directs attention ahead to #185, thus creating an interesting, almost hyper-textual, set of cross-references.[3] It is thus impossible that Blake annotated as he read for the first time only, since he is able to call attention to numbers later than the one he is annotating. Blake does the same in his annotations to Lavater's *Aphorisms on Man;* he writes on the first page, "for the reason of these remarks see the last *aphorism*" (Lavater; my italics). It is thus clear that Blake must have returned to certain passages (or reread entire volumes), not necessarily making marginal comments upon first reading. Indeed, to cross reference, as Blake has done for numbers 69 and 185 in *Divine Providence,* suggests a good deal of thoughtful rereading. Thus, his reactions might not have been entirely bold or fresh or immediate.

McFarland asserts that "Blake's marginalia throughout are united by the common feature of synecdoche for his entire position" (86) and that "Blake's marginalia present, again and again, a synopsis of his entire meaning" (87). I believe that it is in large measure the regularity of typeset that invites the reading of one annotation as "synecdochic" (to use McFarland's term) of a larger body of thought, since all annotations are made to appear materially equivalent on the page. It is for me the questions surrounding how annotations exist on the page—their ontological status as textual and material marks—that seem more pressing than how or if annotations represent pieces of a stable whole.

Are marginalia ephemeral?

Attention to the materiality of Blake's marginalia is pressing if for no other reason than because some of the marginal notes Blake made, especially those in pencil, are literally disappearing from the pages on which they were written. Without this coming to public attention, even more of these documents will be lost than have already disappeared. Of particular concern is the volume of Swedenborg's *Wisdom of Angels* concerning *Divine Love* and *Divine Wisdom* in the British Library (shelf mark c45e1). In this volume, as in most others, Blake filled blank pages with his own writing; however, the blank flyleaf upon which Blake wrote is almost entirely faded and is now completely unreadable, even under 7x magnification. Infrared or X-ray photography, techniques which have been used on *The Four Zoas,* may be able to recover some of what has been lost. But Erdman transcribes marginal text in his *Complete Poetry and Prose of William Blake* that is now no longer visible; I assume he was studying the annotated volumes at some time at least prior to 1965 when his *Poetry and Prose of William Blake* was first issued. Erdman notes that "the pencilled paragraphs on the flyleaf have been badly rubbed or erased, possibly not intentionally; the words supplied within brackets [in the transcription] are conjectural" (884). Erdman's transcription offers two full paragraphs; only a few words are bracketed as conjectural. To my eye, however, those paragraphs have all but disappeared, save for at the very edges of the page where only portions of words remain visible. It seems unlikely that the words were or have been erased intentionally, but rather that the pencil marks have faded away. *In addition to pencil annotations which seem to be fading entirely away,* there are pages in the volume of Reynold's *Discourses* (British Library shelf mark c45e18) whose edges are deteriorating to such an extent that annotations themselves are beginning to literally fall away. The outer edge of the annotations in the right margin of page 67 (a recto page), for example, has deteriorated badly. What Erdman transcribes as "Generalizing in Every thing the Man would soon be a Fool but a Cunning Fool" (649), now reads more like, "Generaliz / in Every th / the Man w / soon be a / Fool bu / a Cunning / Fool." Page 64 has suffered during the process of rebinding. The bottom of the page has been cut so as to destroy some of the annotation.

Erdman's transcription reads "Age & Youth are not Classes but...<Properties> of each Class so are Leanness & Fatness" (E 648). However, after "Properties," all that remains of the last line is "Classes so are Leanness & Fatness." Page 71 offers an even more striking example: what Erdman transcribes as "let them look at Gothic Figures & Gothic Buildings . & not talk of Dark Ages or of Any Age : Ages are all Equal But Genius is Always Above the Age" (E 649) now reads, "let them look at Gothic Figures / & Gothic Buildings . & not talk of Dark Ages or of / Any Age : Ages are all Equal . But Genius."

Erdman also asserts that "Blake's notes [in Reynold's Discourses] were written first in pencil and later, with erasures and additions, in ink. Differences [...] between pencil and ink versions are treated as deletions and additions" (E 886). However, in a number of cases, the words that Erdman asserts to have been in pencil no longer appear, either to my naked eye or under 7x magnification. Curiously though, there are some pages on which the marginal notes have been quite clearly written over with ink yet the pencil remains visible beneath. Perhaps Blake used a different pencil for various annotations, and thus may have annotated at many different times throughout the course of his owning the book (perhaps over the span of a few hours, a few days or a few years). Erdman calls attention to the general assumption that "these marginalia are all of one kind written all at one time," though the variously faded pencil annotations suggest otherwise (E 886).

Further puzzling is the disparity between the care with which Blake inked some annotations but not others (assuming that it was Blake writing with the pen, and not someone later). Often, the original pencil is only barely visible beneath the pen; Blake seems to have carefully traced the shape of the pencilled note. Yet in other places, there appear to be double words (faint lines or ghost words), where Blake has inked a word into place, but has taken no care to follow the original's shape. For example, on page 74 Blake has inked over his original pencil note: "Here he is for Determinate & yet for Indeterminate" and "Distinct General Form Cannot Exist Distinctiveness is Particular Not General." However, the pencil is still clearly visible (written larger and more irregularly). Does allowing for this doubleness suggest that Blake was, consciously or otherwise, responding materially to Reynolds' argument at this point in the text for the generality of form? The "same" annotations exist, after all, yet are still visible as two distinctive sets of marks: one in pencil, the other in pen. Can it just be coincidence that this occurs for the annotation that reads,

"Distinct General Form Cannot Exist Distinctiveness is Particular Not General"? Further, did Blake have plans to erase the visible pencil lines (or did he know that some would, indeed, fade over time)? Or did he not care that they would show (but still cared enough to ink some of them over carefully)? I am certain that he was writing in pen so as to be more legible, not just more permanent, which suggests that Blake imagined an audience for at least some of his marginal notes.

There are a number of further questions that this re-writing raises which are worth pursuing, though they are perhaps no more definitively answerable than any of the others I've posed so far. First, to what degree might Blake have known that his pencil annotations would fade over time. If fading was not a concern, then why take the trouble to ink some over? In another situation, it appears that Henry Crabb Robinson, whose encounters with Blake are recorded in Robinson's *Reminiscences,* inked Blake's pencil annotations to a volume of Wordsworth, almost as if he (Robinson) too feared, or knew, that pencil would not be permanent. Further, what would persuade Blake to ink some annotations but leave others in pencil within a given volume? And what reasons might have caused him to not ink over annotations in other volumes at all (as in Berkeley's *Siris,* for example). I realize that there are more questions than answers here; it remains a considerable research project to discover some of these answers (if they are available at all).

Does annotation open discourse?

What remains clear is that Blake's annotations indicate a voice and hand trying to make itself heard and seen on the otherwise univocal page, though the force with which Blake wanted his voice(s) to be heard seems to vary greatly. That Blake took care to ink some of his annotations suggests he was consciously constructing a text of his own which would respond to the original, at points with as much force as the original itself asserted, and which would present to the reader a "text" (or better, multiple possible textual configurations) which was as viable and as authoritative as the original. It is in this sense of multi-vocalizing the otherwise univocal page that Blake's marginalia resonate with books, writing and reading as symbols he used throughout his poetry and thus why the marginalia deserve attention as, paradoxically, central documents in the field of Blake studies. Annotating should not be treated as an activity separate from any of Blake's other activities as an artist and engraver, particularly because Blake's own bookmaking techniques likely grew in part from his experience reading and annotating the books that circulated around him on a day to day basis.

The "book," for example, represents throughout Blake's work a multivalent symbol, one that is very often connected with authority and the assertion (usually auto-rhetorical) of that authority. Consider Urizen's command to Orc in *The Four Zoas* to "Read my books" (Night VII: line 90), or Urizen's statement in *The [First] Book of Urizen:* "Here alone I in books formd of me- / -tals / Have written the secrets of wisdom" (Erdman 72, lines 24-25). Blake's assertion (and insertion) of another perspective onto the seemingly finished page—the act of annotation— represents a serious challenge to authorial control which is vested in material control of "the book" and thus of meaning. It is in this respect, this disturbance of interpretive authority, that Blake's marginalia function as part of the anti-Newtonian element that pervades his work. If Donald Ault is right that the "Newtonian voice equates 'multiplicity' with 'confusion' and therefore needs to ground [its] direction of the reader's responses in a [...] need for reduction of multiplicity to univocality" ("Incommensurability," 162), then to annotate any text—to multi-vocalize the univocal text—represents a direct challenge to the drive toward Newtonian univocality. In this sense, the marginalia are extremely important in Blake's *ouevre,* since they offer material evidence of Blake confronting what is clearly one of the most important and complex symbols in his poetry: the printed book. The book is a symbolic site in Blake's work where the "Newtonian" relationship between readers and authors is enacted and contested, both by literal readers and authors, as well as by the fictional readers and authors that appear throughout Blake's poetic universe.

To say that the act of annotation is anti-Newtonian is not necessarily to argue that Lavater, or Reynolds, or Thornton, or Wordsworth or any author whose work Blake annotated, was necessarily "Newtonian" in the sense that they espoused Newton's ideas directly. What Blake seems to have sensed as a fundamental condition of communication in general, and of communication through art specifically, was the degree to which it participated, like it or not, in the Newtonian tendency toward fixity, stability and univocality. If there is something "Newtonian" about all communication, it is the inescapable necessity of fixing into place, of stabilizing, that which at some point was fluid and as yet unshaped.

What gives much of Blake's art its intensity is its awareness of its own participation even in that which it tries to work against, including finitude, fixity and stability. Yet the sense of always trying to work against the text as univocal authority (that is, text as speaker of its one "correct" meaning) pervades Blake's work, especially the marginalia. For as the links between reading, writing, books and authority throughout his poetry suggest, Blake was particularly aware of how the printed page could operate as a site of asserted, but also contested, authority. To disrupt the fixity of the finished page is to work against the drive toward univocality that, as Ault asserts, characterizes the Newtonian narrative.

To annotate a text is also to expose the limitations imposed by conventional printing, for there are blank spaces that conventionally typeset print does not enter. Especially when we consider the unconventionality of many of Blake's own pages, it seems likely that part of the critique operating as Blake annotated was implicitly of the growing system of "mass production" which took all kinds of texts and fitted them into relatively uniform moulds. In his *Five Hundred Years of Printing,* S.H. Steinberg notes that printing fonts in the "era of consolidation," as he calls it, from 1550-1800, gained "greater consistency" thanks to improvements in "the mathematical precision of design and in the technical manufacture of punches and matrices" (75). Not to say that all conventionally printed texts in Blake's time looked exactly alike, but there are, undeniably, blank margins and pages in all of the books Blake annotated, which have resulted from their being printed using the same general process. It was not authors themselves who required blank spaces; it was the exigencies of print economy and technology.

The Santa Cruz Blake Study group has identified the role played by the "editorial line of interpretation," in addition to "the exigencies of typographic economics," in the academic production of Blake (*"What Type of Blake?"* 305). This is, I think, usefully extended to include Blake's marginalia. As Blake makes abundantly clear throughout his work, he was particularly sensitive to the connections between industry, commerce and art. To fill a margin was to take advantage of that space left blank in the process of creating a text that could be mass re-producible. To some degree, this is true of almost all acts of annotation; however, it is particularly in the context of Blake's art itself that annotation-as-textual-critique emerges in this way, for his art demands such constant attention to the visual aspect of the page, and to the material processes which bring that page about.

In addition to asserting the importance of the marginalia as "works" that participate in the anti-Newtonian aspect which emerges in his art and poetry, I also want to pursue the intersection between annotating as Blake practised it and the activities carried out by certain characters in his poetry. Consider, for example, Blake's tracing over his own marginalia relative to the figure of Urizen in *The Four Zoas* whom Blake depicts as compulsively tracing his books. These books tend to represent self-defence mechanisms for Urizen, often against feelings of oppression (or attraction) he would rather not admit to. Blake writes, "For Urizen fixd in Envy sat brooding & coverd with snow / His book of iron on his knees he tracd the dreadful letters" (VII: 1-3). And The Book of Urizen is, as Paul Mann writes, "a book about books" (49), suggesting that Blake's struggles with book-writing and -producing became subject matter for those books themselves. It seems that Blake's awareness of his habits as reader and annotator found their way into his poetic work. While I am not suggesting a direct, or conscious, cause-and-effect relation between Blake's annotating and motifs which appear in his poetry, there are certainly particular intersections which illuminate some of the implications of annotating as Blake practiced it, and which suggest the degree to which Blake's activities as an annotator informed the content of his poetry and his work as a bookmaker himself.

Annotating and/as tracing

In Night the Seventh of The Foar Zoas, *Orc rages against Urizen.* Orc is nailed to a burning rock, but Urizen is not similarly bound. "Why shouldst thou sit," Orc asks of Urizen, "cold grovelling demon of / woe [...] thou dost fixd obdurate brooding sit / Writing thy books [...] thy pen obdurate / Traces the wonders of Futurity in horrible fear of the future" (VII: 3-16). (The extension of Urizen's obdurate-ness into an obdurate pen suggests that one writes what one is.) And in the midst of Orc becoming a "Serpent form," and Los sitting in the "showers of Urizen," "Urizen tracd his Verses / In the dark deep the dark tree grew" (VII: 4-11). In these examples, Urizen is both writing and explicitly "tracing," though it is unclear if, each time, he is tracing that which he has already written, that which he's already imagined (though this possibility is complicated since writing often makes imagining possible), or if tracing is a form of displacement: writing what one has already written, but in a new medium.

It is grammatically possible that Urizen traced his verses in the dark deep in the sense that he was in the dark deep; that he traced onto the dark deep; and/or that "In the dark deep" describes where "the dark tree grew." The grammatical instability created by the absence of punctuation plays out relentlessly in the narrative interconstitution in Night the Seventh of Urizen's envy for Los and for Orc, his writing to control Orc, the eruption of the root of Mystery, Urizen's own entanglement in the labyrinth of roots and his compulsion to arrange his books around him. The book he does not arrange around him is the book of iron, which seems to become the rock of iron to which Orc is nailed:

Los felt the Envy in his limbs like to a blighted tree
For Urizen fixd in Envy sat brooding & coverd with snow
His book of iron on his knees he tracd the dreadful letters
While his snows fell & his storms beat to cool the flames of Orc
Age after Age till underneath his heel a deadly root
Struck thro the rock the root of Mystery accursed shooting up
Branches into the heaven of Los they pipe formd bending down
Take root again wherever they touch again branching forth
In intricate labyrinths oerspreading many a grizly deep

Amazd started Urizen when he found himself compassd round
And high roofed over with trees. he arose but the stems
Stood so thick he with difficulty & great pain brought
His books out of the dismal shade. all but the book of iron
Again he took his seat & rangd his Books around
On a rock of iron frowning over the foaming fires of Orc
(VII page 77 line 27 - page 78 line 14)

Again, it is unclear just what "letters" are being traced by Urizen, but the sense of a passage *like* this one seems to be that causation is a textual device. That is, Los feels Urizen's envy (for himself or for Orc) "*like* to a blighted tree" [my italics] following from which an actual root and branches erupt. One of Urizen's books is iron (a characteristic textually "caused" by the iron monsters that chase him) though he leaves the book behind, thus "causing" the rock he sits on to be a rock "of iron." The "logic" of textual causation means that the appearance of a word in the poem (a descriptive adverb like "iron," for example) allows for, or causes, the possibility for that word to appear again, sometimes in a very different context. It becomes difficult to say what specific effect Urizen's tracing has on narrative events, when tracing in fact creates the conditions necessary for certain events to occur at all.

Donald Ault, in Narrative Unbound, *explicates some of the issues surrounding Urizen's books in* The Four Zoas. They are books of iron and brass, which Ault asserts are "direct responses to the monsters [whose scales and fins are made of iron and brass] who devour his [Urizen's] path" (220). The books are, as Ault writes, "one solution to the problem of making a linear path in a path-less space" (220). In the context of *The Four Zoas,* Urizen's books are "a defense mechanism," a "remnant of his need to survive" (225). The compulsion with which Urizen traces (into) his books reflects his own attempts to deny, among other things, what Ault describes as Urizen's "subjugation to Orc" (245). Urizen's books, as they emerge from and submerge into, the narrative field, along with his compulsive tracing, tend to mark those moments where, as Ault writes, "Urizen's control is severely threatened" (221).

Questioning sequence—Questioning authority

Blake is careful to maintain "tracing" as the compulsive action as opposed to "writing," and so it seems reasonable to infer that Blake thought of the two activities differently. His practice of annotating as one that often involved both writing and tracing must have seemed relevant. Ault's analysis of Urizen and tracing makes it possible for me to draw a number of parallels between Urizen's activity and Blake's. Urizen tracing his books as part of his attempt to form a path through otherwise pathless space resonates with Blake himself trying to organize the "pathless" or unused space of another's text, to create for himself a kind of space in which to move, to express his own thoughts or, more often, to arrange his own thoughts in relation to anothers'. Almost any text privileges one correct "path" through its pages; that is, there is only the one printed text to read. Marking in the margins is a way of creating an alternate path through the text by using the otherwise unused spaces. A substantial number of Blake's annotations are meant to oppose the primary text, inviting the reader's awareness, not just of alternate arguments, but of the way in which printed texts suppress those alternatives.

Blake's annotations open provide a different way of moving through the existent text; in this case, annotations act as directional devices as opposed to (or in addition to) functioning as direct commentary. It is obvious to note that books are paginated (or sections numbered) such that readers can move through the argument "properly" by following consecutive numbers. It is perhaps so obvious that it has become an aspect of the control that texts exert over readers that goes largely unnoticed. Jerome McGann, in *Towards a Literature of Knowledge,* asserts that "the printed book is one of the most illusionistic of human works, imputing as it does an aura of permanence to the discourses we manipulate" (12). Philip Cohen, introducing *Texts and Textuality,* writes "the layout, typography, binding, paper, and ornamentation of a book work in concert with or in opposition to the linguistic text they convey" (xvi). Just as Blake's annotations remind us that the original text suppresses certain arguments and perspectives in favor of its own, so too do his annotations remind us that a text relies heavily on maintaining typographic control over the reader. Often, by challenging the implicit authority a text has over how it is to be read one also challenges the authority a text exerts over its own meaning.

As I had mentioned earlier, Blake annotates the numbered entries in Swedenborg's Divine Providence *by harkening to other numbers,* sometimes earlier, sometimes later in the volume. The annotation to aphorism 198, for example, reads, "Mark this it explains N 238." More striking is perhaps the annotation to 330: "Swedenborg contradicts himself & N 69 See also 277 & 203 where he says that a Place for Each Man is foreseen & at the same time provided." Blake seems at once to be formulating his own path through the text, not only by inserting his own text in the space available, but also by suggesting that the reader can move among the numbered sections in ways other than by proceeding in "order." Indeed, Blake asserts that his path (to move from 330 back to 69, or 277, or 203) will reveal a contradiction in the existent text. *Thus issues of textuality are important when it comes to the marginalia,* since a book's appearance can have such a profound effect on the experience of reading. Seeing the way Blake has marked the printed pages of a book opens the possibility for considering how certain formal/textual arrangements control reading. Typeset editions of Blake's marginalia do not give readers access to the textual or material aspect of the printed pages as Blake marked and re-marked them. The degree to which the typographic features he encountered in his reading and annotating prompted him to develop a radically different kind of book technology for his own work has thus never been considered.

Meaning is authorized not just by the author (though the Newtonian text will assume as much) but equally by the material conditions of the book itself. The Santa Cruz Blake Study Group suggests that "The effects generated by the emblematic characteristics of the book will constitute a significant part of the terms on which the contents of the book are offered and received" (311). And further, "Our ability to read has been conditioned by our familiarity with traditional linear text forms and *"the consistent and powerful appearance they present"* (310; my italics). It is particularly the "powerful appearance" of the book which plays such a crucial, yet paradoxically subtle, role in reading. Blake's marginalia function to call attention to the material arrangement of books, even the regularity of their typeset, by occupying spaces that should not be occupied and by presenting an irregular script. Marginalia can remind us that our willingness to accept as authoritative that which we read is deeply conditioned by material appearances; in the case of Blake's marginalia, this occurs when material appearances are disrupted.

For Blake to ink his path, to mark dead or suppressed space and to assert a new text in the margins, are all particularly significant if we pursue the relation to Urizen's tracing in *The Four Zoas.* Such tracing may have been, for Blake just as much as for Urizen, a kind of defense mechanism, a kind of path making and finally a kind of assertion of control over that which he, as any reader, is ultimately subjugated to: the book itself as he/it confronts it/him. That Urizen is compulsive about his tracing and writing may speak to compulsions that Blake himself felt to oppose certain texts (and certain pathways through those texts) with a marginal text/pathway of his own. The annotations to Reynolds' *Discourses* suggest the anxiety and pressure Blake may have felt to respond materially to a figure he aligned himself so completely against. Given Blake's dislike for Reynolds, it is not surprising that in this volume particularly Blake took great care to ink into place those annotations which would go to form an authoritative voice (or set of voices) to rival that of the author.

In his annotations to Reynolds, Blake is also careful to reveal his own contextual position relative to the work he's reading and responding to. On the blank verso of the title page, Blake explains his position at length:

> Having spent the Vigour or my Youth & Genius
> under the Oppression of Sir Joshua & his Gang of
> Cunning Hired Knaves Without Employment & as much
> as could possibly be Without Bread . The Reader must
> Expect to Read in all my Remarks on these Books
> Nothing but Indignation & Resentment....Reynolds and
> Gainsborough Blotted & Blurred One against the other &
> Divided all the English World between them Fuseli
> Indignant almost hid himself I am hid
>
>   (Reynolds)

Ironically, the last word, "hid," is itself almost obscured by the edge of the page. (Blake was running out of room, and the page is now deteriorating.) However hidden he felt though, it is clear that Blake is looking to set a text up that opposes Reynolds'. Even more revealing is the direct address to "The Reader"; here Blake seems explicitly confident that somebody will be reading the volume that he has annotated.

[He also addresses an imagined reader in an annotation on the last page of Lavater's *Aphorisms on Man:* "I hope no one will call what I have written cavilling because he may think my remarks of small consequence" (Lavater 224).] Whether they will be reading for Blake's work or for Reynolds' seems irrelevant to Blake's sense of his own words standing irremovably next to, above, below and around those of Reynolds. The care with which Blake inked over his original pencil notes makes it clear that he was envisioning a text that would stand in forthright opposition to, and with as much material permanence as the text that occasioned it.

Blank space in the text became for Blake the occasion to offer comment. The instance of the word "hid" becoming almost hidden (lost, actually) due to lack of room on the page is only one example wherein Blake had to arrange his marginal commentary as space would allow, under the conditions permitted by the material arrangement and dimensions of the book itself. In some cases, Blake's annotations completely surround the printed text (as on certain pages of Richard Watson's *Apology for the Bible,* for example); it is hard to imagine that in such circumstances Blake would not have developed responses further if the margin size allowed. Just as his notes are often deeply embedded conceptually in the text that they surround, and are thus in some ways controlled by that text, so do the material layout and dimensions of the original control the possibilities open for Blake to establish other voices on the page.

However much Blake's annotations contest the printed text to which they respond, however, it is perhaps surprising that Blake leaves the original text itself for the most part untouched. In other words, he rarely "(if ever)" defaces the text he is annotating. One notable exception occurs in Lavater's *Aphorisms on Man,* though even here while Blake has crossed out certain of Lavater's words, they remain readable. There are occasional underlinings and brackets in the margins. For example, in the annotations to Swedenborg's *Divine Love and Divine Wisdom,* Blake uses large } shaped brackets to "collect" lines, next to which he writes "Note this" (sections 410 and 411) or "Mark this" (section 421). But Blake does not deface the original text. It is almost as though he was careful to preserve the original: the more accessible it remained, the more forceful his annotations would seem as oppositional possibilities. By multiplying texts on the page, rather than displacing one for another, the more dialogic and multivocal the page could become. And there remain, of course, those instances in which Blake's annotation, especially where that notation involved bracketing and underlining, required the original for its meaning. As in Blake's poetic universe, the writer/annotator establishes a complex relationship to books; one that involves, by turns, conflict, opposition, dependence and agreement.

While he does not deface the text by rendering it illegible, there are pages on which Blake's extensive use of the blank spaces makes it almost impossible for the reader's eye to go to the original first and not to the annotations. At best the reading eye is pulled between the two texts, thus altering what would have been the original path of reading. On facing pages 126 and 127 in Blake's volume of Reynolds' *Discourses,* annotation encompasses the entirety of the margins surrounding the two pages. The entire left margin of 126 is filled, as is the footer, in addition to the entire right margin of 127 (and its footer). The annotation in Reynolds which runs the entire length of the left column of page 126 and into the bottom margin reads:

> According to Reynolds Mich Angelo was worse still &
> knew Nothing at all about Art as an Object of Imitation
> Can any Man be such a fool as to believe that Rafael &
> Michael Angelo were Incapable of the meer Language of
> Art & That Such Idiots as Rubens : Correggio & Titian
> Knew how to Execute what they could not Think or
> Invent.
>
> (Reynolds)

The actual line breaks as they occur because of the limited space Blake is working in are quite abrupt, following something like: "According / to Reynolds / Mich Angelo / was worse / still." Many of the semantically tricky spots in the typeset version, wherein one idea runs into the next, almost to the point that they are hard to separate, are much less severe when the line breaks are restored. For example, if we read "According to Reynolds Mich Angelo was worse still & knew Nothing at all about Art as an Object of Imitation," the next line "Can any Man" could read like a question: as in, "can any man know about Art as an object of imitation?" Reading on, however, we see that Man is actually the subject of "be such a Fool," yet again the possibility presents itself that Blake is asking, "Can any man be such a Fool as to know nothing of Art as an Object of Imitation?" Reading yet further reveals that the most likely sense of the lines is to break the first semantic unit after "Imitation," and that the next phrase actually reads, "Can any Man be such a fool as to believe that Rafael & Michael Angelo..." Restoring the line break—"Art as / an Object of Imitation / Can any / Man be"—makes Blake's notation seem much less like the collection of run-on sentences suggested by my typeset version above.

It remains clear that the margins provoked Blake to develop a textual "voice" that was going to be deeply connected to the text that occasioned it (just as Blake warns that the reader is to expect a direct response to Reynolds), yet which would vie with that text for equal importance on the page. In the Reynolds volume especially, there is the sense that Blake is constructing an identifiable, alternate voice (or, more properly, voices) to control as much of the page as it can, not just to comment on Reynolds' text, but to develop a position of its own.

Who is reading?

The question regarding who Blake might have imagined himself to be writing for is crucial here. In Blake's marginalia to Reynolds, we have a direct address to the "Reader." In addition, despite the limitations of space imposed by the margins, Blake's annotations contest for visual dominance of the pages. Blake writes on the Reynolds title page, in script much larger than the printed original, "This Man was Hired to Depress Art This is the opinion of Will Blake my Proofs of this Opinion are given in the following Notes" *(figure 1).* Blake is careful to name himself as author, just as he might do in any of his illuminated works. And the positioning of these lines in relation to those printed on the title page is uncanny. While he does not render the original unreadable, Blake has used the blank spaces between the printed lines to provide his own title page of sorts: what amounts to a centered title, "This Man Was Hired / to Depress Art," attribution, "This is the opinion of Will Blake," and subtitle, "My Proofs / of this Opinion / are given in the / following Notes." Further, Blake has underlined "Reynolds," perhaps to highlight that while the *Discourses* purport to represent the tastes of a larger community (one from which Blake felt particularly excluded) they are nonetheless the opinion of just one man. Blake even stylizes his lettering on this page: the "H" of "Hired" is complete with flourishes, as are the "D" of "Depress," and the "A" of "Art." The stylized lettering certainly challenges the uniform regularity of the typeset, and it does seem apparent that Blake took care in composing this annotation on the page.

Figure 1. Title Page to *The Works of Sir Joshua Reynolds* with William Blake's Annotations (By permission of the British Library; shelf mark c45e18)

SIR JOSHUA REYNOLDS.

Tum demum sana mentis oculus acute cernere incipit: ubi corporis oculus incipit hebescere.

Seneca.

Publish'd according to Act of Parliament, March 1795 by T. Cadell Strand.

This Man was Hired to Depress Art

This in the Opinion of Will Blake: my Proofs of this Opinion are given in the following Notes.

THE
WORKS
OF
SIR JOSHUA REYNOLDS, KNIGHT;
LATE PRESIDENT OF THE ROYAL ACADEMY:

CONTAINING

HIS DISCOURSES, IDLERS,
A JOURNEY TO FLANDERS AND HOLLAND,
AND HIS COMMENTARY ON DU FRESNOY'S ART OF
PAINTING;

PRINTED FROM HIS REVISED COPIES,
(WITH HIS LAST CORRECTIONS AND ADDITIONS,)

IN THREE VOLUMES.

TO WHICH IS PREFIXED
AN ACCOUNT OF THE LIFE AND WRITINGS OF THE
AUTHOR,
BY EDMOND MALONE, ESQ.
ONE OF HIS EXECUTORS.

THE SECOND EDITION CORRECTED.

———QUALI NON EA PRÆCIPIAM ALIIS, QUÆ MIHI IPSI DESUNT. CICERO.

VOLUME THE FIRST.

LONDON:
PRINTED FOR T. CADELL, JUN. AND W. DAVIES, IN THE STRAND.
1798.

Advice of the Popes who succeeded the Age of Raphael
Degrade first the Arts if you'd Mankind Degrade,
Hire Idiots to Paint with cold light & hot shade:
Give high Price for the worst, leave the best in disgrace,
And with Labours of Ignorance fill every place.

It is perhaps no surprise that while Blake did leave some of his annotations to other volumes in pencil, it was in Reynolds' *Discourses* that Blake took care to ink some of his annotations, in most cases making them more legible, and in all cases making them more permanent. As Robert Wark points out in his edition of Reynolds' *Discourses,* the discourses "were prepared as formal lectures to the students and members of the Royal Academy [...] They were delivered [...] on the occasion of the annual prize giving." And significantly, "the *Discourses* were tantamount to a statement of policy for the young institution" (xiv). It was no doubt this sense of the discourses as communally instructional and as statements of policy that provoked Blake into mounting his own counter-statements of policy in the spaces available to him. *Books throughout Blake's work are, as I've suggested above,* representative of the power to authorize and stabilize law, rules, measurement and ideas. Blake may have found Reynolds' *Discourses* a kind of document whose "institutional" purpose was little different from that of Urizen's books. The *Discourses* were "statements of policy" delivered to students in order that they might learn to abide by that policy. Urizen's various books contained, for example, the "secrets of wisdom," the "Laws of peace, of love, of unity: / Of pity, compassion, forgiveness [...] One command, one joy, one desire, / One curse, one weight, one measure / One Kind, one God, one Law" (*Book of Urizen* Copy C, Plate 4, lines 25-40).[4] Against the kind of insularity and singularity implied by the *Discourses,* Blake took care to present an alternate set of opinions, some of which he clearly took time to consider (presumably at least those he traced over in ink). By asserting explicitly that the notes represent "the opinion of Will Blake," Blake is able not only to foreground his own authorship, but also to remind the reader that the *Discourses* themselves are of singular authorship and that their authority to represent a communal taste is potentially suspect. Blake's explicit self-representation may also be a reaction to the particular edition of the *Discourses* he had, which begins with "Some Account of Sir Joshua Reynolds," a biographical essay, some hundred pages long, celebrating Reynolds and his role in developing the artistic wealth and prosperity of England. *Blake reminds, however, that Reynolds was "Hired,"* and might therefore be under compulsion to represent a position that is not entirely his own, just as students themselves will be expected to represent the Royal Academy's ideals. In fact, many of Blake's annotations refer to Reynolds' relationship to the aristocracy on the one hand (those who've hired him) and artists on the other (those who must defer to his authority). For example, Blake refers to "Sr" Reynolds and "his Gang of Cunning Hired Knaves" on the blank verso of the title page. On the page containing Reynolds' dedication "TO THE KING," Blake tellingly writes,

"O Society for Encouragement of Art! O King & Nobility of England! Where have you hid Fuseli's Milton." On page civ of "Some Account of Sir Joshua Reynolds," Blake writes, "This Whole Book Was Written to Serve Political Purposes." Erdman suggests that unerased pencil on this page also reads, "[*?First to Serve Nobility & Fashionable Taste & Sr. Joshua*]" (E 641; brackets and italics are Erdman's). However, I was unable to see these words; they may have faded entirely. The annotations to Reynolds represent one of the best examples of all the volumes Blake annotated of Blake working to alter without physically destroying the original text, and to mount a text of his own in the spaces provided. His attention to detail, his address to the reader, his explicit self-presentation and the time it must have taken to ink some marginal notes into permanence show the degree to which Blake wanted to oppose Reynolds' text by producing an alternate text which demanded its share of any reader's attention.

I have tried to show here some of the important issues that I think surround William Blake's marginalia. In addition, I've tried to give a sense of how these kinds of issues remain inaccessible so long as readers do not have access to anything but typeset editions. Ultimately, Blake's marginalia need to become central to Blake studies. The issues at stake in the marginalia are inextricably linked to those at stake in the rest of Blake's work, and his activity as a reader and annotator certainly informed his work as a bookmaker. Finally, the treatment of Blake's marginalia in scholarship to this point may provide a keen lesson for other fields of study in the way that editorial and economic decisions can have profound effects on academic production by enabling or constraining certain interpretive possibilities.

---

[1] Saree Makdisi has suggested that "the very way we have learned to read is precisely what prevents us from reading Blake properly" (111).

[2] The methodology at work in Molly Anne Rothenberg's *Rethinking Blake's Textuality* provides a useful counterpoint to McFarland; Rothenberg asserts that "one of the principles guiding my work [...] is that Blake's corpus ought not to be regarded as a unity, as a product of a single intentional activity, nor ought it to be read as furnishing philosophical propositions" (2).

[3] See my "Recentering Blake's Marginalia" for an alternate discussion of these issues (*Huntington Library Quarterly* Vol.66, 2003).

[4] There are important issues surrounding the various plate orderings, inclusions and exclusions for this particular book. I address them more fully in other work, including "Blake's Awareness of Blake in a Newtonian World" (History of European Ideas) and my forthcoming *The Torn Book: UnReading William Blake's Marginalia* (Susquehanna University Press). Numerous other scholars provide thorough discussion of Blake's plate ordering; a key resource is Joseph Viscomi's *Blake and the Idea of the Book*.

Works cited

Ault, Donald. 1987. *Narrative Unbound*. Barrytown, NY: Station Hill Press.

_____. 1986. "Incommensurability and Interconnection in Blake's Anti-Newtonian Text." In Hilton, Nelson, editor. *Essential Articles for the Study of William Blake, 1970-1984*. Hamden, CT Archon Books, 141-174.

_____. 1974. *Visionary Physics*. Chicago: University of Chicago Press.

Cohen, Philip, editor. 1997. *Texts and Textuality*. New York: Garland Publishing.

Erdman, David. 1988. *The Complete Poetry & Prose of William Blake*. New York: Anchor.

Frye, Northrop. 1947l *Fearful Symmetry*. New Jersey: UP Press.

Keynes, Sir Geoffrey. 1966. *Complete Writings of William Blake*. London: Oxford University Press.

Lilly, Josiah K. Jr. 1948. Correspondence to Lilly, Jr. from a librarian at the Yale University School of Medicine, 3 August.

Makdisi, Saree. 2003. "The Political Aesthetic of Blake's Images." In Eaves, Morris, editor. *The Cambridge Companion to William Blake*. Cambridge: Cambridge University Press, 110-132.

McFarland, Thomas. 1990. "Synecdochic Structure in Blake's Marginalia." *European Romantic Review*, I, 75-90.

McGann, Jerome. 1989. *Towards a Literature of Knowledge*. Chicago: University of Chicago Press.

Reynolds, Sir Joshua. 1798. [Discourses] *The Works of Sir Joshua Reynolds*. Volume I. The second edition corrected. 3 vols. London, British Library shelf mark, c45e18.

Rothenberg, Molly Anne. 1993. *Rethinking Blake's Textuality*. Columbia, Mo: University of Missouri Press.

Santa Cruz Blake Study Group. 1986. "What Type of Blake?" In Hilton, Nelson, editor. *Essential Articles for the Study of William Blake, 1970-1984*. Hamden, CT: Archon Book, 301-333.

Steinberg, S.H. 1996. *Five Hundred Years of Printing*. New Edition. London: The British Library and Oak Knoll Press.

Viscomi, Joseph. 1993. *Blake and the Idea of the Book*. New Jersey: Princeton University Press.

Watson, Richard. 1797. *An Apology for the Bible, in a series of Letters, Addressed to Thomas Paine*. Eighth edition. London: T. Evans. Huntington library call number, 110260.

Author note

Jason Snart is an Assistant Professor of English at the College of DuPage in Glen Ellyn, Illinois. He has published on Blake in European Romantic Review, Huntington Library Quarterly, *and* History of European Ideas. *His study of Blake's marginalia,* The Torn Book, *is forthcoming from Susquehanna University Press. Research for this article was made possible by the Social Sciences and Humanities Research Council of Canada, the Yale Center for British Art, the Huntington Library and the College of DuPage.*

A reply to *"The Mutamathil Type Style"*
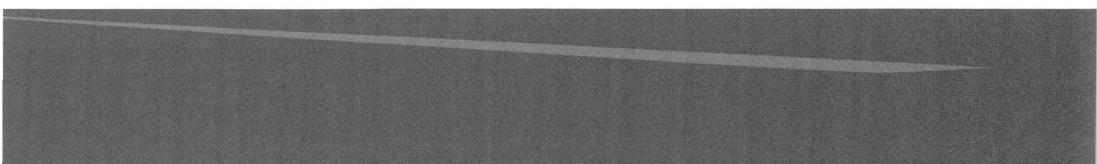
Nadine Chahine

Arabic typography has been the subject of much interest lately in international confer-
ences and publications.[1] The last few years have been especially interesting because of
the new developments in technology, such as Opentype font extension and the develop-
ment of the Unicode standard (Hudson, 2000). This has made easier a large number
of the complexities of dealing with the Arabic script, such as the large character sets,
context sensitivity and the application of vocalization marks. Here, an important question
presents itself. If technology is no longer a barrier to the representation of the script, then
what form should a type designer give to this work? Should one strictly adhere to the
calligraphic and ornamental origin or is there a different way to look at things? Given
that Arabic typography has had a long struggle with technology, today's situation is very
interesting.

Calls for reform: Past and present

In the middle of the last century, the Academy of Arabic Language in Cairo sent out a
call for proposals for script reform. Hundreds of replies came in, and all were rejected
(Yacoub, 1986).

It is quite understandable that no major changes to the script were accepted. As Wellisch
(1978) explains, "once a writing system has been successfully established for a language
community, its underlying conventions become essentially indivisible, all-embracing and
intolerant of any other convention." This is so much so that it would be "impossible to
replace an even unsuitable system with a better one geared to the character of the lan-
guage." He also explains that change could come in the form of the addition or removal
of a few characters, but that the writing system, in general, can be changed to another
"only by force." Such forces could be religious or political. Greek, Latin and Arabic are
examples of scripts that benefited from the powers of religion in order to spread geo-
graphically. The fact that the Latin and Arabic scripts are as widespread as Christianity
and Islam says a lot about the connections between script and religion, and to a large
extent, national identities (Wellisch, 1978). Therefore, when attempting to answer the
question of why no solution was accepted, one should look at the proposed solutions,
while keeping in mind that it is external factors that governed the final decision making.
As such, it is impossible to discuss large-scale script reform without admitting to the
influence of politics, religion and culture.

Some of the solutions were very impractical so it is no shock that they were rejected (Yacoub, 1986). As to others, each has its own advantages and disadvantages. It is very understandable that the call for the adoption of Latin would be rejected because it meant a complete dissociation with the past, especially in the domain of calligraphy. That alone would have been enough to discredit those proposals, no matter how earnest they were. The fact that the Arabic script is highly tied to Muslim religion makes it quite sacred, a fact that almost makes it untouchable for Arabs. Because the Koran was relayed in Arabic, any change to the Arabic script means a change to the way the Koran is written. This close association between script and religion could be the reason why no solution was chosen at all. Those that could have made sense (especially Nasri Khattar and Ahmad Zaki Mawlawi) were just too radical. It is highly probable that a less controversial solution was desired, a solution that would not change "the look" of the script, but would keep its essence. One can never know what really went on in the minds of the committee members. Still, it is probably no coincidence that Linotype would arrive with Simplified Naskh just a few years later, especially that the Linotype Matrix acknowledged various current attempts at script reform and the fact that the solutions looked strange or unattractive (Linotype, 1960).

The widespread acceptance of Simplified Naskh, a cheaper and faster way to print Arabic by using two forms per letter rather than the usual four, as compared to the failure of the Academy of the Arabic language can only confirm the conclusion that the Arab linguistic authorities were ready for a new solution and that the safe and familiar one eventually won out. Linotype had collaborated with a Lebanese, Kamel Mrouwa, on the project so it is again probable that he had been aware of what was going on for the past few years (Linotype, 1959).

Today, again, one reads about the simplification of the script through the "Mutamathil Type Style" by Saad Abulhab. It is a noble idea to undertake a project that aims to make a script easier to deal with. However, there are basic essentials that need to be addressed such as adequate knowledge of the technology of the day, the educational ramifications of any adjustment to the script and a basic understanding of typography and type design, whether Latin or Arabic.

Abulhab starts his article with an introduction of his understanding of the Latin script. It is not the purpose of this reply to address that. However, some errors of judgment need to be pointed out.

"Typography is the art of automated calligraphy." p.306

"Current printed or visual Latin forms vaguely resemble the old ones. It is a challenge to read an old English or German book from a few centuries ago!" p.307

"Producing a font for the extended Arabic set today minimally requires the design of 500-600 glyphs, depending on type or calligraphy style, compared to no less than 200 glyphs to cover all Latin scripts." p.311

One might need 200 glyphs to represent the English language but that's not the only language that uses the Latin script. A font that supports Central and East European languages quickly exceeds that number, especially when one looks at all the accented characters that are needed to properly represent those languages. For example, Adobe's Minion Pro has 1246 glyphs, more that 700 of which are for the Latin script.

Technology

As mentioned in the introduction, there is very little that technology can't do for the Arabic script.

"It would be a mistake for non-Latin typography to settle forever on all the rules imposed by the current Unicode standards. The machine must not be forced to duplicate all and every detail of the old calligraphy." p.310

This seems to be a misunderstanding of what Unicode is. It is a plain text encoding (The Unicode Consortium, 2000). The Unicode standard gives one the freedom to fully represent calligraphy but certainly does not force one to do that. At the end of the day, it is the designer who draws the outlines, whether they are richly calligraphic or very simple in design. Unicode provides a way to deal with the various glyphs needed to fully represent a language; one can fill the boxes with whichever design one sees fit. A designer can even add more glyphs than what Unicode provides, as long as he/she does the proper substitution lookups.

"Ligatures that belong to the basic character set (e.g. Waw with Hamza)…"
p.311

Lam-Alef is a required ligature, yes, but Waw with Hamza is not. It is not a typographic nicety that can be removed without resulting in a spelling error. It is a grammatical neces-sity and it is not a combination of two characters (though it looks like that) but a visual representation of the Hamza in a specific grammatical situation. A better name might be accented character, but not really. This is why it has a separate key.

"Mixing right-to-left texts with left-to-right texts can be a nightmare." p.313
This is true, but the situation is getting better thanks to the Unicode bi-directional algo-rithm (The Unicode Consortium, 2000).

"But the typewriter failed to move the Arabetic scripts into the typographic age. With the emergence of computers, the few positive typewriter-based attempts at simplifying the Arabetic written forms, quickly evaporated. Arabetic typog-raphers were again busy duplicating calligraphy to its fullest detail in their type designs." p.315

There are several problems with the above statements. One, the typewriter was a very successful attempt to simplify the script and to deliver a working model that is cheap, reliable and preserves the look of the script. Two, it paved the way to Linotype's Simpli-fied Naskh which is, again, a good working solution to print Arabic cheaply and with good quality of design (Linotype, 1959) and it is still very popular today as in the example of the widespread popularity of Yakout (Ross, 2002). Three, Arabic type designers might have wanted to be obsessive about full calligraphic details in the 80's and 90's but were not always so in reality, if only because of technical limitations. Now, they can be. Type design should not follow calligraphy blindly, but develop as a separate field of study.

At the same time, it would be rash to just ignore calligraphy completely for the sake of "new" designs. One needs to understand how the script is formed, its structure, the ductus, why the forms look the way they do. This can be learned by observing cal-ligraphic models. When one comes to design one's own character set, it does not have to "look" like a replica of a calligraphic piece, but it should maintain the structure and the internal proportion of the script, and that is influenced by the tool that created it.

"The mutamathil type style proposes a technology-oriented, computer friendly, minimal type style." p.323

This is a bit vague. How can one be computer friendly? Is it like dolphin friendly tuna? The computer does what the user wants it to do. Also, technology-oriented seems to mean that it fully utilizes the capabilities of modern technology, yet the system Abulhab proposes goes out of its way to stay away from anything that does that. He seem to be designing for a technology that has long gone by.

"Incorporating this type eliminates all major and unique obstacles faced by articles of manufacture utilizing the traditional Arabetic alphabets." p.319

Glyph positioning lookups to position the vocalization marks are still needed. That is even more complicated than substitution. The way they are now is too floating. Their position needs to be locked to the proper character, at different heights depending on the design. For example, the "yalid" word he has in his specimen is confusing. The two marks are very close to the lam.

"The required right-to-left ordering, which is technically the main challenge facing Arabic typography…" p.318

That has not been such a challenge since the first printed Arabic book. Context sensitivity, the large character set and the vocalization were worse to deal with.

Education: The issue of bi-directionality [2]

It would be common sense to assume that teaching people to read Arabic text from left-to-right might be misleading when these same people will have to read real right-to-left Arabic. The Greeks had a similar model a long time ago[3] but they soon switched to a single direction. They certainly had a point there.

If people find Arabic strange because it is read from right-to-left, maybe they should put in some effort and get used to it. Knowledge is not a piece of cake that one eats in two minutes. One is expected to work hard to learn any new language or script. In a world being overtaken by globalization and the widespread example of fast food chains and large international corporations and vendors, it is a pleasure to note that people are different after all. It's a big world and people should celebrate the uniqueness of every culture, script or language, not try to modify its essence to suit global trends. It's a matter of mutual respect and tolerance.

It is true that the script is complex, but the solution to have it bi-directional adds to that complexity. Also, the many different forms are something that the eye can get used to, as one can see in every single literate Arab and non Arab (the number is increasing thankfully). The existence of too many dots is worse than the different forms because the dots can be erased, misunderstood, fall off…

It would be ludicrous to imagine an Arab designer proposing a symmetrically designed Latin alphabet that can be read from right-to-left so that Arab readers can learn English faster.

Though the script is complex, literacy rates are improving greatly (UNESCO, 1999). There is a lot that one can do before we blame the script. If people are interested in learning it, then they would put in the extra effort to get used to it. Typographers can help by studying legibility and by experimenting within the boundaries of accepted aesthetics, perhaps pushing those boundaries farther than expected, trying new things, but always remaining faithful to the essence of the script, its structure and its internal logic.

The concern to keep the natural feel of the script is understandable. Still, one of the "natural" qualities of a script is its inherent direction so how can reading Arabic from left to right with symmetrical letters preserve that quality? It changes the way one reads and even the location of the text in relation to the book. Also, how will people know if a text set in Mutamathil is left to right or right to left? The slight details to differentiate the two options are not visible enough except if one goes looking for them, and text should be transparent. Experienced readers will probably be able to guess because of the vocabulary, but new readers might get confused. It gets worse in the case of justified text.

The concept

If one tries to summarize the characteristics of the Mutamathil type, it has two main ones: 1) it is unattached and 2) it is bidirectional.

The issue of bi-directional has already been discussed. The issue of unattached is not really debatable. It's been done before and it is not a script reform as much as a typographic simplification. As such, there is no problem there and it really is an extra option on the menu. It would probably work more in display typography rather than in text settings though.

The problem, though, is that the Mutamathil type still has some of the problems that Abulhab stated were the reason why the simplification attempts were not successful. For example:

"Some of these designs truly violated the spirit of Arabic writing and ignored legibility."

"Many ignored addressing the vowel diacritics completely."

"… all glyphs failing to include the important visual effects of the traditional letters joining/non-joining process."

Also, his assessment of the reasons why these failed were unfair to Nasri Khattar's work:

He *did* advocate an open design system and designed his system in several styles.
He did *not* approach type design calligraphically.
He did *not* insist on the ligatures (Khattar, 1947).

The reasons why his solution was not accepted are more complex as explained earlier.

The design

The design is changing some of the fundamental structures of the characters (like the dal) to the extent that they are almost unrecognizable, let alone readable. The fact that they are symmetrical has taken the direction out of them so they seem not to form words, just a string of static forms. The eye doesn't know where to go.

Abulhab might not claim to be a type designer, but in reality, he is offering a typographic product, and as such one cannot ignore the look of it and the way it functions.

There are a few design problems:
• the rhythm is abrupt and truncated
• the internal proportions need some changes
• the lam doesn't fit the typeface
• the hamza is not recognizable
• there are too many little details that create tension, especially the truncated descenders
• some characters are very strange: qaf, dal, waw, lam, sin

The proposal that less symmetrical design is possible is a better idea but what happens to the people who are introduced to Arabic as a left to right script? They will have to flip things around in their head when they start reading real Arabic. On a personal note, I found his example of left to right text hard to read but fun to look at and try to guess what's the text, but he chose the most common sentence to Arabs (the bismillah) and I'm a native reader.

It is very easy to list the many problems that Arab readers face and say that yes, the script needs to be reformed. However, such musings are still based on what linguists or laymen say. If one were to agree to that, then what should one reform? The problems are many and solving all of them would completely change the script. Should one, then, recognize which problem is the most serious and try to solve only that one? Would that be the

diacritics or the context sensitive forms? How can one judge? Technical complexities are not problematic anymore but the Arabic public, though literate, still does not read many books per year (UNESCO,1999).

It is very important that people address the various problems that Arab readers face. Whether such questioning leads to fruitful results or not, the journey is still worth the effort, if only to make way for further questioning and investigation. Arabic typography is in dire need of research and study and it is crucial that such endeavors are encouraged and supported through institutions of higher education, research centers and publications.

[1] One example would be the long afternoon session dedicated to Arabic typography in the annual Atypl conference in Vancouver in 2003.

[2] As a matter of clarification, on a technical level, the Arabic script is defined as bi-directional because the numerals are written from left to right while the text is written right to left. However, that is an established tradition of the script and is accepted as the norm.

[3] Known as boustrophedon reading in which the reading direction alternates from right-to-left to left-to-right. This is different from the Mutamathil system in which the text is set in only one direction throughout the whole paragraph. Still, it is cited here as an example of a system that supports changing directions.

Bibliography

Khattar, Nasri. 1947. *The unified Arabic type.* New York: privately published.

Linotype and Machinery Limited. 1959. "Simplified Arabic." *Linotype Matrix,* No. 32.
——, 1960. "New faces of the '50s." *Linotype Matrix,* No. 33.

Mawlawi, Ahmad Zaki. 1957. *Al-Arabiya al-jadida.* (Translator, the new Arabic). Oman: New Printing Co.

The Unicode Consortium. 2000. *The Unicode standard version 3.0.* Reading, Massachusetts: Addison Wesley Longman, Inc.

Wellisch, Hans H. 1978. *The conversion of scripts: its nature, history and utilization.* New York: John Wiley and Sons.

Yacoub, Emile. 1986. *Al-Khat Al-Arabi: Nash'atuhu, tetawurahu, mushkilatuhu, da'wat islahihi.* Tripoli, Lebanon; Jrous Press. (Translation: The Arabic script: its origin, its development, its problems, its reform).

Tawfik, M., editor. 1999. *The World Communication and Information Report 1999-2000.* Paris: UNESCO. Also http://www.unesco.org/webworld/wcir/en/index.html

Hallouda, A. Mokhtar. 1999. Chapter 14. In Tawfik, M., editor. *The World Communication and Information Report 1999-2000.* Paris: UNESCO. See also www.unesco.org/webworld/wcir/en/pdf_report/chap14.pdf

Hudson, John. 2000. "Windows glyph processing" available from http://microsoft.com/typography/developers/opentype.htm

Ross, Fiona. 2002. "Non-Latin type design at Linotype" available from http://stbride.org/conderence2002/NonLatin.html

Response from Saad Abulhab

The key concept of my article missed by the reply writer was that my proposed Mutamathil type is not an attempt to "reform" Arabic, replace the traditional forms, or force a certain way to display it. It is rather a call to make widely available types that are not tied completely with the old rules of the beautiful Arabic calligraphy and the habits of the Arabic writing hands, so that we can let users decide, not the few influential elite. This is why I likened creating non-conventional Arabic types to the composing of free non-conventional Arabic poetry from the last century that has proven very popular after it became available to readers. I do not agree with the notion that the rules of writing Arabic by hand should be referred to as the "rules of the Arabic." As a matter of fact, scholars are not even sure if the earliest Arabic (Kufi), which was either derived from the northern Aramaic Nabtic or southern al-Musnad writing systems, was based on attached letters. Logically and most likely, it was written in isolated forms until people found it more practical to connect the letters. After all, Arabic writing is derived from all isolated-form writing systems so it must have started with isolated forms.

If technology is no longer a barrier to the representation of the script, then what form should a type designer give to his work?

How can one be computer friendly? Is it like dolphin friendly tuna? The computer does what the user wants it to do. Also, technology-oriented seems to mean that it fully utilize the capabilities of modern technology yet the system Abulhab proposes goes out of its way to stay away from anything that does that. He seem to be designing for a technology that has long gone by.

Technical complexities are not problematic anymore but the Arabic public, though literate, still does not read that many books per year.

"Technical complexities" are always going to be with us and will always be problematic. One must be careful not to claim that Arabic technological problems or any other technological problems end forever because technology is constantly changing and can always present a barrier. To create technology-friendly Arabic types is to make sure that such types are not dependent on any complex technology and to ensure that such types must "stay away from" it! Arabic must not be again a victim of any future unknown technology and now is a good time to act. BTW The Mutamathil design is promoting the abandonment of the unnatural way today's technology handles Arabic input. As I

have pointed out in one discussion forum, what would English users feel if they must see capital letters as always shown and are changed automatically by the computer to lower case only when keying the next letter? I am designing for a simpler technology that is not available yet and therefore is **not** "long gone by!" Furthermore I have submitted two proposals to the concerned international technology organization to implement a new technology that is not a barrier to the Arabic natural input system.

"The required right-to-left ordering, which is technically the main challenge facing Arabic typography…" p.318. That has not been such a challenge since the first printed Arabic book. Context sensitivity, the large character set and the vocalization were worse to deal with.

This is a crucial issue. *Being involved in computing, Arabic computing and programming since the early eighties, I can say with confidence that the bidi requirements of Arabic was and still are the major challenge to integrating it **economically and in a timely fashion** in software and on the device level. Major sites spoke left- to-right simplified Chinese before Arabic. Context sensitivity and larger sets are far less complicated tasks from the programmer side. Yes, it is better now after the Unicode bidi Algorithm but my point was that Arabic directionality requirements was and still are the major challenge of the Arabic computing field and **hence** Arabic typography today. In the age of computers, typography advances are closely related to advances in computing. In my statement, I was not referring to Arabic typography before the emergence of computers that is why I used technically and is to refer to its **current computer-related typography challenges.***

As Wellisch [1978] explains, "once a writing system has been successfully established for a language community, its underlying conventions become essentially indivisible, all-embracing, and intolerant of any other convention." This is so much so that it would be "impossible to replace an even unsuitable system with a better one geared to the character of the language." He also explains that change could come in the form of the addition or removal of a few characters, but that the writing system, in general, can be changed to another "only by force."

Simply put *I disagree with most of Wellisch's assertions above. History proved otherwise.*

The fact that the Arabic script is highly tied with Muslim religion makes it quite sacred, a fact that almost deems it untouchable for the Arabs. That of course would be because the Koran was relayed in Arabic and any change to the Arabic script means a change to the way the Koran was written.

This is mainly a stereotype! *The Arabic script is a very flexible and progressive script. It is very "touchable" and is not sacred at all. No Muslim religious authority ever objected to changes in Arabic writing forms as long as they are not along the line Kamal Attaturk chose. My Mutamathil type is very Quraan-aware. Incidentally, as long ago as a thousand years past, and to a less extent today, Arabic religious calligraphers competed in producing writings of Arabic both in right-to-left and left-to-right directions. This mirrored calligraphy style was referred to as al-Muthanna, or al-Maraya calligraphy style. Sometimes it was called al-Muzdawaj, al-Mutanadhir or al-Muta'akis calligraphy. How much more flexible and open minded could they be?*

The widespread acceptance of Simplified Naskh, a cheaper and faster way to print Arabic by using 2 forms per letter rather than the usual 4, as compared to the failure of the Academy of the Arabic language can only confirm the conclusion that the Arab linguistic authorities were ready for a new solution and that is was the safe and familiar one that eventually won out. Linotype had collaborated with a Lebanese, Kamel Mrouwa, on the project so it is again probable that he had been aware of what was going on for the past few years (Linotype, 1959).

Simplified Naskh used *two forms for most letters, not all, and it is very progressive and beautiful indeed.*

One might need 200 glyphs to represent the English language but that's not the only language that uses the Latin script. A font that supports Central and East European languages quickly exceeds that number, especially when one looks at all the accented characters that are needed to properly represent those languages. For example, Adobe's Minion Pro has 1246 glyphs, more that 700 of which are for the Latin script.

Yes one might need *200 glyphs to represent English but the minimum needed is 52 glyphs. I was comparing the absolute minimum glyphs needed. Today, a typical traditional Arabic (not extended Arabic) type would need minimally around 130 glyphs.*

This seems to be a misunderstanding of what Unicode is. It is a plain text encoding (The Unicode Consortium, 2000). The Unicode standard gives one the freedom to fully represent calligraphy but certainly does not force one to do that. At the end of the day, it is the designer who draws the outlines, whether they are richly calligraphic or very simple in design. Unicode provides a way to deal with the various glyphs needed to fully represent a language; one can fill the boxes with whichever design one sees fit. A designer can even add more glyphs than what Unicode provides, as long as he/she does the proper substitution lookups.

Especially in Arabic, Unicode standards go hand-in-hand with other standards like OpenType to present its final look and feel. I agree that Unicode is not responsible for the glyph substitution process but it definitely complemented it. In my article I was referring to the right-to-left directionality that Unicode has imposed as a requirement on Arabic. But I could have been more accurate and mention "current standards" rather that "Unicode standards."

One can't define typography as automated calligraphy and then complain that Arab designers are following calligraphy as a model.

I do not see any contradiction at all. When automating calligraphy and hand written models, one must stay away from copying their full details, rules and habits. The concept here is "the art of automated calligraphy" and not "the art of calligraphy."

Lam-Alef is a required ligature, yes, but Waw with Hamza is not. It is not a typographic nicety that can be removed without resulting in a spelling error. It is a grammatical necessity and it is not a combination of 2 characters (though it looks like that) but a visual representation of the Hamza in a specific grammatical situation. A better name might be accented character, but not really. This is why it has a separate key

Lam-Alif is a combination of the Lam and Alif glyphs. Waw-Hamzah is a combination of the Waw and small Hamzah glyphs. They both have separate keys. They are both required and are both "ligature like" glyphs typographically speaking since they are not part of the official Arabic alphabet. My IBM Arabic typewriter, which was a top-notch tool until the computers arrived, did not have a Waw-Hamza key. The user had to type Waw followed by small hamzah separately.

There are several problems with the above statements. One, the typewriter was a very successful attempt to simplify the script and to deliver a working model which is cheap, reliable and preserves the look of the script. Two, it paved the way to Linotype's simplified Naskh which is, again, a good working solution to print Arabic cheaply and with good quality of design (Linotype, 1959).

In my statements, I too was trying to praise the progress that the Arabic typewriter had made in simplifying Arabic even further than what Linotype's simplified Naskh did. This progress the computer era has wasted in favor of creating its complex technology to duplicate and force the old calligraphy models. The reply comments seem to agree so I do not see the problem.

He still needs to use glyph positioning lookups to position the vocalization marks. That is even more complicated than substitution. The way they are now is too floating. Their position needs to be locked to the proper character, at different heights depending on the design. For example, the "yalid" word he has in his specimen is confusing. The 2 marks are very close to the lam.

In my designs, *I am proposing to get rid of positioning the marks (Harakat) on 3-5 height levels depending on the letter as it is done today and to place them simply after the letters on two levels only taking advantage of the new letter separation gaps created. I am proposing this precisely because placing marks is "even more complicated than substitution" as the writer replies. My prototype samples are further updated on my arabetics.com site. Take a look at: http://arabetics.com/jawahiri.htm or http://arabetics.com/quraan.htm. If one examines how marks are positioned with many of today's traditional Arabic types, one would find many worse examples than the mentioned word "yalid." I am proposing a concept not final product!*

It would be common sense to assume that teaching people to read Arabic text from left-to-right might be misleading when these same people will have to read real right-to-left Arabic.

It is true that the script is complex, but the solution to have it bi-directional adds to that complexity. Also, the many different forms are something that the eye can get used to, as one can see in every single literate Arab and non Arab (the number is increasing thankfully). The existence of too many dots is worse than the different forms because the dots can be erased, misunderstood, fall off…

In the Mutamathil style, *I wanted to lift all obstacles to Arabic without violating its spirit. As I mentioned above, bi-directional Arabic is a very old practice and was not created by my style. There are some applications that even require learning Arabic in both directions. I proposed a symmetrical letter that can look the same from both directions. Users can order it anyway they like. Chinese and Japanese are read in both directions without any problem. Options can only do well and enrich a writing system. English is written both connected and isolated and no one is confused!*

It would be ludicrous to imagine an Arab designer proposing a symmetrically designed Latin alphabet that can be read from right-to-left so that Arab readers can learn English faster.

Arabic is an international language *like English. It is not ludicrous for a non-native of any of these languages to propose anything.*

If one tries to summarize the characteristics of the Mutamathel type, it's 2 main ones: 1) it is unattached, 2) it is bidirectional.

Another main concept is: 3) It represents each Unicode letter by one glyph.

The problem, though, is that the Mutamathil type still has some of the problems that Abulhab stated were the reason why the simplification attempts were not successful. For example: "Some of these designs truly violated the spirit of Arabic writing and ignored legibility." "Many ignored addressing the vowel diacritics completely." And "… all glyphs failing to include the important visual effects of the traditional letters joining/non-joining process."

A good example of violating the spirit of Arabic script is the Boutemene design style, which is available today as DecoType ArabicLatin font. The Mutamathil type style glyphs are clearly Arabic glyphs that are being used today or were used historically. Also, the style did address the vowel diacritics adequately. As for the joining/ non-joining visual effects, our take on previous separate glyphs designs (including Khattar's) was that, unlike our Mutamathil style, they imposed equal spaces even with letters like "waw," "dal" and "alif" which need additional spacing after. Still I think all options should be available widely for users to judge.

Also, his assessment of the reasons why these failed were unfair to Nasri Khattar's work:
He *did* advocate an open design system and designed his system in several styles.
He did *not* approach type design calligraphically.
He did *not* insist on the ligatures (Khattar, 1947).

The reasons I stated were general and did not primarily apply to Khattar's attempt and I do not want to detail my personal opinion about his design here. Incidentally, I singled out Khattar's attempt as one of the two best attempts to simplify Arabic. His designs were open indeed and were not very calligraphic. But all samples I have seen of his work kept the lam-Alif ligature. It was the mentality of absolute or forced reform that probably prevented his work from spreading.