AN EVALUATION OF TEST IMAGES FOR MULTIPLE-CHOICE
COMPREHENSION ASSESSMENT IN APHASIA

A thesis presented to
the faculty of
the College of Health and Human Services
Ohio University

In partial fulfillment
of the requirement of the degree
Master of Arts

Sabine Heuer
June 2004
This thesis entitled

AN EVALUATION OF TEST IMAGES FOR MULTIPLE-CHOICE COMPREHENSION ASSESSMENT IN APHASIA

BY

SABINE HEUER

has been approved for

the School of Hearing, Speech and Language Sciences

and the College of Health and Human Services by

Brooke Hallowell

Associate Professor of Hearing, Speech and Language Sciences

Gary S. Neiman

Dean, College of Health and Human Services
The use of images in multiple-choice formats for comprehension testing is common. The validity of responses of those tests may be confounded by (a) physical stimulus features, such as size, luminance or clarity of images or (b) semantic content conveyed by multiple-choice images such as image familiarity and frequency. An instrument, based on the literature and validated by experts, was developed to assess such influences within multiple-choice images. Twenty-three adults rated 20 sets of images from five published aphasia batteries. Eye movements were recorded for a separate group of 20 adults viewing the same image sets. All sets prompted disproportionate visual attention, suggesting a need for consideration of these influences when using current tests and better control in designing test images.

Approved:

Brooke Hallowell

Associate Professor of Hearing, Speech and Language Sciences
Acknowledgments

I want to thank, foremost, my advisor Dr. Brooke Hallowell, for her steady support and exceptional guidance.

I also would like to acknowledge the following people for their helpful insights: Dr. Norman Garber and Dr. Sally Marinellie, members of my thesis committee, Don Adleta and Ryan Douglas who reviewed the Multiple-Choice Visual Stimulus Control Instrument. I am thankful for the support of Natalie Douglas, who collected the data with me and who shared her thoughts about eye movement research, a topic that she also has explored. I want to thank the members of the Neurolinguistics Research Laboratory who supported this project, especially Anshula Odekar who helped with the data collection. I want to thank Hans Kruse for his technical support with the eye movement recording system. Finally, I want to thank Dr. Sunny Kim, the statistical consultant of the College of Human Services, who developed the equation to calculate the eye movement pop-out score, an important tool in my data analysis.
Table of Contents

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
</tr>
</tbody>
</table>

**Abstract**................................................................................................................................3

**Acknowledgments**................................................................................................................4

**List of Tables** .......................................................................................................................7

**List of Figures** ......................................................................................................................8

**Chapter 1** Introduction .......................................................................................................9

**Purpose of Study** .................................................................................................................10

**Physical Stimulus Features and Semantic Content**

**Conveyance**..........................................................................................................................11

**Physical Stimulus Features** .................................................................................................11

**Semantic Content Conveyance** .............................................................................................19

**Research Questions, Hypotheses and Assumption** ...............................................................25

**Research Questions and Hypotheses** ....................................................................................25

**Assumption** ............................................................................................................................27

**Chapter 2** Method ...............................................................................................................28

**Phase 1: Assessment of Multiple-Choice Images** .................................................................28

**Phase 2: Eye Movement Assessment** .....................................................................................30

**Analysis** .................................................................................................................................32
List of Tables

Tables                                                                                           Page

1. Physical Stimulus Properties and Semantic Content Conveyance Factors..................................26

2. Correlations Between Ratings of the Stimulus Properties and the Pop-out Eye Movement Score ..........................................................................................................................34

3. Correlations Between Ratings of Physical and Semantic Scores and the Pop-out Eye Movement Score ..................................................................................................................35
List of Figures

Figures                                                                                                                 Page
1. Illustration of the pop-out effect evoked by a different color.........................12
2. Illustration of the pop-out effect due to a difference in orientation.............14
3. Illustration of the pop-out effect due to a difference in image size. ..........14
4. Illustration of the pop-out effect due to a difference in depth cues. ...........15
5. Illustration of the pop-out effect due to contrasts in luminance. ...............16
6. Illustration of the differences in complexity. .............................................17
7. Illustration of the pop-out effect due to a difference in shape. ....................18
8. Illustration of the pop-out effect due to a difference in symmetry. ...............18
9. Illustration of the pop-out effect due to a difference in clarity. .................19
10. Illustration of differences in scene context that may influence
     disproportionate visual attention. ...............................................................22
11. Illustration of difference in actions versus objects. ..................................23
12. Illustration of differences in perspective.....................................................23
B1. Eye movement pop-out score of each image set. .........................................46
B2. The mean eye-movement pop-out score of individual aphasia tests...............47
Chapter 1

Introduction

The use of images in multiple-choice formats for linguistic comprehension testing is common. Almost all published tests in current use, such as the Boston Diagnostic Aphasia Exam (BDAE) (Goodglass, 2001), the Western Aphasia Battery (WAB) (Kertesz, 1982), the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA) (Kay, Lesser & Coltheart, 1997), and the Reading Comprehension Battery for Aphasia (RCBA-2) (LaPointe & Horner, 1998), include multiple-choice tasks involving images. It is assumed that the patient with aphasia perceives what is represented by the images represented in such tasks. There are many reasons why such an assumption may be invalid. The reason for a participant’s failure in a multiple-choice task might be the assumed lack of linguistic understanding. However, the participant might respond inappropriately (by selecting an incorrect foil) because of the influence of a) the physical stimulus features and for b) confounds in the semantic content conveyed by the visual stimuli within an array of multiple-choice images. The role of such stimulus-driven aspects is rarely considered carefully in the design and evaluation of visual stimuli used in multiple-choice components of aphasia test batteries.

There is ample evidence that physical stimulus features and aspects of content conveyed by images influence the visual attention that even language-normal adults allocate to various components of visual displays (Hallowell, Wertz, & Kruse, 2002; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). Most patients with aphasia and related neurogenic language disorders demonstrate generalized deficits in attention
(Helm-Estabrooks, 2002; Murray, 2002), and many also present with specific visual attention deficits (Denes, Semenza, Stoppa, & Lis, 1982; Edmans & Lincoln, 1987; Harrington & Drake, 1990; Liesegang & McPhee, 1992; Myers, 1998; Tompkins, 1995; Wilson & Wyper, 1992; Ylvisaker, 1992). Attentional impairments may affect the accuracy and efficiency of an individuals’ language comprehension skills (Murray, 2002), as well as the accuracy and efficiency of their selection of an appropriate multiple-choice response. Thus, consideration of the influences of stimulus features and image content on test responses of patients with neurogenic communication disorders is especially important.

It is impossible to predict the degree of influence of physical stimulus features and semantic content conveyance factors and whether this influence may override the selection of an image of a particular individual. The only way to ensure that these factors do not affect any patient is to control for them in the design of images. An important first step in learning to control for such influences is to study their influence with existing test stimuli. The current study addresses means of evaluating physical stimulus features and semantic content conveyance in multiple-choice images for comprehension testing in aphasia.

Purpose of Study

Both issues described above, the physical features of the images and the semantic content conveyed by images, are difficult to control. The first aim of this study is to develop an instrument to assess visual stimulus properties and semantic content conveyance in multiple-choice images, based on an extensive review of empirical literature, and validated by experts in graphic design. The second aim is to validate
specific aspects to be controlled for on the design of multiple-choice images through empirical testing of viewers’ eye movement patterns as they look at a sample of images drawn from published batteries for the assessment of aphasia.

*Physical Stimulus Features and Semantic Content Conveyance*

*Physical Stimulus Features*

There is plentiful evidence that the physical stimulus features of images, such as size, color, texture, shape, symmetry, and luminance, have an impact on a viewer’s visual attention as indicated by eye movement and fixation patterns. (e.g., Barbur, Forsyth, & Wooding, 1980; Locher et al., 1993; Wolfe, 2000). Such features tend to make certain images within complex stimulus arrays more attractive or distracting than others to participants. Additionally, some patients with aphasia have concomitant cognitive deficits that may make them more sensitive to the influences of such visual stimulus features.

Cognitive processes performed to solve tasks are called top-down processes. Reactions that are evoked by the task material and which do not underlie volitional control, such as the physical stimulus features and semantic content conveyance factors of test images, are called bottom-up processes. Recognizing the multiple-choice visual, corresponding to a linguistic stimulus and selecting the appropriate stimuli, are top-down processes performed when solving a task for linguistic comprehension in multiple-choice format. The viewer may be “distracted” by bottom-up processes that he or she cannot control. For example, it is possible that a participant understands a verbal stimulus and identifies the correct corresponding visual stimulus but becomes distracted by the physical features of another stimulus (a non-target foil) so that she or he selects the distracting stimulus instead of the correctly identified one. In such a case, accurate
comprehension of the verbal stimulus would not be captured in the “incorrect” score the participant would receive for that item. Methodical control of specific physical stimulus features in multiple-choice images is thus essential to the validity of patient responses to those images during testing. Items that do not share the same basic visual characteristics as all other items in the perceptive field may be detected automatically (Wolfe, 2000). For instance, a single red item between blue items would be automatically detected. This phenomenon is referred to as the “pop-out” effect. Features of multiple-choice images in a display evoke the pop-out effect when the viewer’s visual attention to specific images becomes disproportionately allocated among presented images. For instance, colored images mixed with black and white images may attract greater visual attention, as illustrated in Figure 1.

![Figure 1. Illustration of the pop-out effect evoked by a different color.](image)

Visual attention can be influenced by a set of basic features: color, orientation, size, depth cues, luminance, shape, and clarity. Thus, each of these features, if not controlled for across sets of items in a multiple-choice display, may evoke the pop-out effect. Their influence will be described in detail below.
*Color.* Color functions as a distractor in image based tasks. Colored items are automatically selected between black and white items or significant different colored items during early visual attention processing (Wolfe, 2000). Generally, if black and white images and colored images are presented together in the same display, more attention is given automatically to the colored images than the black-and white images, as shown in Figure 1. Deffner (1995) conducted a study concerning image characteristics considered critical in image evaluation. Participants were shown a series of images and had to express their preferences regarding the image quality. Color saturation, color brightness, and color fidelity were all items shown to influence how participants viewed images.

*Orientation.* If a single item within an image has a different orientation than other items within that image, the item with the differing orientation becomes salient. The same effect occurs if one object presented in a multiple-choice display has a different orientation. The viewer’s visual attention might be influenced by the different orientation of the object and he or she looks more often and longer at this image than at the other images with the homogenous orientation. For example, viewers are more likely to look longer at the third image in the set at three images shown in Figure 2. The pop-out effect decreases if there is more than one item with a different orientation (Wolfe, 2000).
Size. When viewing multiple images within one display, relative size is a physical property of images that influences scanning patterns (Hallowell, 1999). The size of a stimulus refers to the spatial extent of the item. The disproportionate size of an object is likely to attract disproportionate attention to images within a multiple-choice display. The viewer is more likely to focus on the biggest or the smallest object in a display of several images. When viewing the images within Figure 3, for example, the smaller size of one object is likely to become salient when shown with the others.

Depth cues and shading. Individuals give more attention to visual stimuli cued in depth, for instance, through shadows than to two-dimensional stimuli without depth cues (Wolfe, 2000). Shading, highlight details, and shadow contrast influence eye movement.
patterns when viewing images (Deffner, 1995). Disproportionate looking at a multiple-choice image display occurs when two-dimensional images and images with depth cues are displayed together as illustrated in Figure 4.

Figure 4. Illustration of the pop-out effect due to a difference in depth cues.

**Luminance.** Barbur et al. (1993) found that background color and luminance have an impact on viewers’ scan patterns. In their study numbers were recalled better using a middle-grey background instead of a black one. The correct performance of tasks also increased when the luminance of the background was greater than one-third of that of the target. Additionally, contrasts in luminance have been demonstrated to be recognized faster and also with higher frequency than changes in motion and complexity (Nothdurft, 2001). The different degrees of luminance of images cause a disproportional distribution of eye movements in multiple-choice displays as illustrated in Figure 5. Likewise, differences of luminance between the backgrounds of the images can influence the viewer’s visual attention as well.
Complexity. Eye movement patterns differ depending upon the complexity of images one is viewing (Andrews & Coppola, 1999). Having different levels of complexity across individual items within a display may result in differential visual attention, as measured by eye movement patterns allocated to those images, regardless of the content of an accompanying verbal stimulus.

When viewing complex shapes, viewers tend to focus on familiar geometrical shapes within them (Lass, Lueer & Schemat, 1993). For example, viewers are likely to look longer at the images on the left compared to the images on the right in Figure 6. Bozkov, Bohdanecky, Radil-Weiss, Mitrani and Yakimoff (1982) showed that viewers fixate the angles of figures (polygons) when they have to identify their shapes. According to the authors, “the existence of converging line segments of the polygonal contour shapes are an important cue for form perception and eye movement guidance” (p. 725). Disproportional looking in a multiple-choice display can occur if the shapes of the presented polygons include different amounts of converging line segments and angles. Recognition time varies, depending upon how easily the complex shapes of the objects can be recognized as familiar geometrical concepts. The complexity of the shape of an object has an impact on a viewer’s eye movements. For instance, a circle is more easily
recognized than a shape combined of a polygon and a circle. Viewers tend to focus longer on a complex shape than on simple ones, as illustrated in Figure 7.

Figure 6. Illustration of differences in complexity.

The frequency and size of saccadic eye movements and corresponding fixations differ according to the content of images viewed (Andrews & Coppola, 1999). While viewing a complex natural scene, the duration of individual fixations decreases while the number of eye movements increases. The opposite effect is observed while watching simple patterns, such as dots, squares and lines (Andrews & Coppola), as illustrated in Figure 7. If the complexity of the images in a multiple-choice format differs among the images within a display, the proportion of fixation duration allocated to images within the display may be imbalanced. More fixations and greater total fixation durations may be observed on complex images than on simple ones.
Symmetry and asymmetry. Symmetry is identified during the first few saccades within a visual scan, and has a strong influence on the subsequent scan patterns. Because redundancy is the nature of symmetry, eye movements are often recorded primarily on one side of the symmetric shape. In contrast, while scanning asymmetric stimuli, the eye movement fixations are distributed widely over the stimulus (Locher, et al., 1993). Thus, symmetric figures are recognized more easily and faster than asymmetric ones. Disproportional distribution of eye movements may result when a different degree of symmetric and asymmetric forms are displayed in images in a multiple-choice format. For instance, viewers are more likely to fixate for longer periods of time on the asymmetric form compared to the symmetric forms in Figure 8.
Clarity. The time a viewer spends fixating on images tends to be greater when the image is blurred than when it has clear boundaries (Mackworth & Bruner, 1970). If images in a multiple-choice display have different grades of clarity, the viewer is likely to fixate longer on the most blurred image compared to other images. The distribution of eye movements would not be balanced amongst images within the display. For example, viewers are likely to fixate longer on the blurred image illustrated in Figure 9.

Figure 9. Illustration of the pop-out effect due to a difference in clarity.

Semantic Content Conveyance

A second important issue when using stimuli for language comprehension tasks is to ensure that the form of the visual stimulus corresponds with the semantic content of the associated verbal stimulus. To recognize the stimulus, the brain compares the image features to an ideal mental representation, or a prototype, of this image. Based on this comparison, the viewer assesses the image and makes a decision whether the image is recognized (Locher & Nodine, 1974; Wolfe, 2000). Other issues of content conveyance are how much information and what kind of information is given. The scene context provides information and has an impact on semantic content conveyance. The category of
the stimulus, whether an object or an activity is represented, is also important. The complexity of the content, the viewer’s viewpoint, the familiarity and frequency of use of the presented object, as well as the cultural background of the viewer influence the semantic recognition and processing of the image. These aspects of semantic content conveyance will be described below.

*Scene context.* According to Boyce and Pollatsek (1992) “scene context contains geometric as well as semantic information that operates on the identification of objects” (p. 229). These authors observed the influence of scene context on object identification. They found that participants fixate longer on objects that do not belong in a specific context, for instance a bathtub in a living room. Additionally, objects in inappropriate contexts are identified less accurately than objects in ecologically valid contexts, where the viewer expected them. The “positional violation” (right object in wrong context) “produces costs” in identification accuracy according to the authors (p. 229). The background has an impact on identification accuracy of objects. Boyce and Pollatsek attribute this phenomenon to the “spreading activation model” (p. 231): Objects are represented in the memory through “nodes.” If one of these nodes is activated by a picture, that node sends signals to all possible nodes in its environment. The viewer builds up expectations for what he or she might identify next. The identification of an object is, according to Boyce and Pollatsek, facilitated through such a network if the context is appropriate for the objects depicted within it. If participants are shown images with targets in a typical context, then it is easier to identify them, compared to when they are presented without context as illustrated in Figure 10. The strength of this mental network is defined by semantic relationships between the nodes. “The more properties
two concepts have in common, the more links there are between the two nodes ... and the more closely related are the concepts” (Collins & Loftus, 1975 p. 411). Semantic congruence between context and object depends on the interconnections between the two. According to the authors, activation can only start from one node but can spread from there in different directions simultaneously. The more often one path in the network is used, the more stable this connection will become (Collins & Loftus, 1975). According to Collins and Loftus the names of concepts are stored phonemically. Each phonemic node is connected to concept nodes.

Disproportional looking may be evoked when the context of images within a display is not controlled. For example, if some objects in multiple-choice displays are shown in isolation while others are shown within a scene context, the distribution of fixations is not likely to be balanced among the isolated objects and the images with scene contexts. Likewise, if one object is displayed in an unusual or inappropriate context, the viewer might need more time to identify the object accurately and a disproportionate distribution of eye fixation might occur as well.
Figure 10. Illustration of differences in scene context that may influence disproportionate visual attention.

*Imageability.* Imageability refers to the ease and accuracy by which a semantic idea is conveyed by a visual stimulus. It corresponds to the notion of abstraction versus concreteness of a depicted concept. If one or more of the target images within a display are not “imageable,” this may influence where a person looks within a display. For example, it is harder to represent the abstract concept of “angry” than to represent the concept “flower” and “ball.” The image for “angry” may disproportionately attract a viewer’s attention when shown with images of a flower and a ball. The imageability of concepts underlies the finding of Loftus, Kaufman, Nishimoto and Ruthruff (1992), who demonstrated that objects are recognized faster and at higher rates than actions when controlling for physical stimulus features. The authors’ interpretation for these results is that stationary objects, such as a chair or lamp, are easier to distinguish from one another, whereas actions look similar. For example, the baseball and soccer players in Figure 11 are likely to take longer to recognize than the glass and bottle depicted. If varying stimulus categories are presented in a multiple-choice format, disproportional looking can
be evoked because one category can be more easily recognized than the other, and eye movements are therefore likely to be distributed differently.

*Figure 11.* Illustration of difference in actions versus objects.

*Viewpoint.* The recognition of an object might depend on the viewer’s point of view (Biederman, 1995). For instance, a familiar object, such as a cup might be harder to recognize within a display when the viewpoint is from above the cup, as opposed to the front of the cup as illustrated in *Figure 12.* Disproportionate allocation of eye movements may occur if one object in a multiple-choice display is represented in an unusual perspective in contrast to the other objects.

*Figure 12.* Illustration of differences in perspective.
Image familiarity. According to Mackworth and Bruner (1970), the fixation time on the visual stimulus increases by 40% when the stimulus is not familiar to the viewer. Rayner, Sereno, Morris, Schmauder, & Clifton (1989) found similar results when observing the reading of unfamiliar words. The amount of time a reader looks at a word increases when the word is not familiar to the reader. This phenomenon must be considered especially with regard to the viewer’s social and cultural background. His or her semantic lexicon might include a different vocabulary than the one corresponding to the concepts represented by the images in a multiple-choice display. More visual attention is generally allocated to the unfamiliar objects compared to familiar ones if they are presented together in a multiple-choice display.

Concept frequency. Concept frequency is a construct representing the frequency with which an individual encounters a particular concept in everyday life. The vocabulary of each person contains words that are used often and ones used rarely. Word frequency is traditionally determined by how often words occur in printed material (Carroll, Davies & Richmond, 1971). Although not a construct previously used in the research literature, it may be a potentially important aspect of semantic content conveyed by images. The meaning of the construct “concept frequency” parallels in the cognitive domain what word frequency means in the linguistic domain. Rayner, Sereno, Morris, Schmauder, & Clifton (1989) observed effects of word frequency on eye movements during reading. The ease or difficulty in processing a word is reflected in the fixation time on this word while reading. The fixation time depends not only on the number of syllables in a word but also on the word’s predictability. Compared to high-frequency words, low-frequency words tend to be fixated longer while reading. Rayner showed that word frequency has an
effect on eye movements. Although word frequency and concept frequency are not identical, objects representing concepts that correspond to low- and high-frequency words displayed together within a display might cause disproportional viewing patterns.

*Social and cultural influences.* For the purpose of directly complementing linguistic stimuli, each visual stimulus in comprehension assessment represents a prototypical image of the semantic content it is intended to convey. This presents several social and cultural challenges due to individual life experiences. For instance, a *sari* is not a prototypical image of clothing in western European and Northern American cultures, but it does represent prototypical clothing in eastern Indian cultures. Furthermore, if a linguistic assessment task contains a picture of a *computer mouse* and the viewer has never heard of a *computer mouse*, but the other images within the display represent familiar objects, the viewer may fixate longer on the unknown object as compared to the other items within that display. Also, if a viewer has a mental image of *car* as one that dates from the 1950s, an image of a modern car will not fit the viewer’s mental image of that concept. Thus, if the other images within the display better fit that individual’s ideal representation of the intended concept, the viewer may fixate longer on that particular image within a display.

**Research Questions, Hypotheses and Assumption**

*Research Questions and Hypotheses*

1. Hypothesis. If the features that have been found to have an impact on the viewer’s visual attention are not well controlled in a set of images in a multiple-choice display, then the number of fixations and the total duration of fixations will be unequally distributed across the images presented in the display. The degree to which visual
stimulus properties and semantic content conveyed by images within multiple-choice displays are not controlled will correspond to the degree of disproportionate visual attention allocated to those images. The properties are summarized in Table 1.

Research question. Do independent ratings indicating distraction or lack of control within sets of multiple-choice images correspond with viewers’ disproportionate looking at individual images within multiple-choice displays, as indexed by eye movement patterns?

Table 1

*Physical Stimulus Properties and Semantic Content Conveyance Factors*

<table>
<thead>
<tr>
<th>Physical Stimulus Properties</th>
<th>Semantic Content Conveyance Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>Scene context</td>
</tr>
<tr>
<td>Orientation</td>
<td>Imageability</td>
</tr>
<tr>
<td>Size</td>
<td>Viewpoint</td>
</tr>
<tr>
<td>Depth cues and shading</td>
<td>Image familiarity</td>
</tr>
<tr>
<td>Luminance</td>
<td>Concept frequency</td>
</tr>
<tr>
<td>Complexity</td>
<td>Social and cultural influences</td>
</tr>
<tr>
<td>Symmetry and asymmetry</td>
<td></td>
</tr>
<tr>
<td>Clarity</td>
<td></td>
</tr>
</tbody>
</table>
2. *Hypothesis.* Some stimulus properties such as size and familiarity influence disproportionate viewing of multiple-choice images more than others, such as symmetry.

*Research question.* Which visual stimulus characteristics and semantic content conveyance factors most influence disproportionate viewing of multiple-choice images?

3. *Hypothesis.* Naïve viewers allocate disproportionate visual attention to pictures within multiple-choice image sets, as indexed by eye movement measures.

*Research Question.* Do naïve viewers allocate disproportionate visual attention to pictures within multiple-choice image sets, as indexed by eye movement measures?

4. *Hypothesis.* The ratings of experts indicating distraction or lack of control within sets of multiple-choice images will correspond more accurate with the eye movement patterns than the ratings of naïve raters.

*Research Question.* Do independent ratings of experienced raters correspond different with viewers’ disproportionate looking at individual images within multiple-choice displays than the ratings of naïve raters?

*Assumption*

These hypotheses are based on the assumption that the physical stimulus properties and semantic content conveyance factors that have an impact on the viewer’s visual attention will be detected through disproportionate eye fixations in multiple-choice displays of images. The degree to which a feature influences the viewer’s visual attention will be indexed by the degree of disproportionate looking.
Chapter 2

Method

Phase 1: Assessment of Multiple-Choice Images

Participants. Twenty adults, ages 18 to 25 were chosen from the Ohio University community. Students, taking undergraduate coursework, were asked to participate. They did not have any knowledge about the purpose of the experiment. They were not majors in Linguistics, Psychology, or Hearing, Speech and Language Sciences. Students in these majors may have experience with the aphasia tests from which the samples were drawn. All were native speakers of English. Prior to enrollment in the study, potential participants were asked if they have any history of neurological impairments. Those reporting neurological impairment would have been excluded. However, none of them reported any impairment. An additional group of three individuals familiar with the purpose of the experiment served as experts regarding the stimuli properties. All participants took part in a vision screening which includes the assessment of near vision and documentation of whether participants use glasses. All participants passed a hearing screening at 25dB for the frequencies of 500, 1000 and 2000 Hz.

Stimuli. The images were selected from a sample of different multiple-choice components of published aphasia tests including the Boston Diagnostic Aphasia Exam (BDAE) (Goodglass, 2001), the Western Aphasia Battery (WAB) (Kertesz, 1982) the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA) (Kay, Lesser & Coltheart, 1997) the Test of Communicative Abilities in Daily Living (CADL) (Holland, 1980) and the Reading Comprehension Battery for Aphasia (RCBA-2)
(LaPointe & Horner, 1998). The images were arranged in 20 sets of three without any accompanying verbal stimulus. All stimuli used in this study were black and white images. Colored multiple-choice images were excluded, since there were not enough multiple-choice image sets across the five tests to allow sufficient representation of image sets in which colored images would be included. The image sets were chosen from single items within single tests. The WAB, the BDAE and the CADL include multiple-choice sets of more than three images. Paper copies were made of those image sets. The images were turned over so that the image was not visible, and were presented to a graduate student, naïve to the purpose of the study. The student chose three pieces of paper randomly. This procedure was repeated with the four image sets of the three tests mentioned above. For assigning an order for the presentation of image sets, the volunteer numbered the sets from one to 20 without seeing the images, but by writing a number on the blank side of the paper copies. To assign an order of presentation within each image set, the naïve person assigned a number from one to four on the backside of the paper copy of each image. One represented the upper left corner in the display, two represented the upper right corner, three was assigned to the lower left corner and four to the lower right corner.

Procedure. An instrument for assessing viewers’ perceptions of pop-out effects or distraction by any stimulus features or semantic conveyance factors was implemented. The instrument is shown in Appendix A. The items in this instrument are considered to be valid because they are drawn from a careful review of the relevant empirical literature. Prior to its administration, two professional graphic artists were told the nature of the current project and were asked to evaluate and comment on the validity of items in this
instrument. Refinements were made based on their suggestions. The physical stimulus characteristic color was not considered in this study. A group of 20 naïve raters and three experienced raters then evaluated the images in each display using the rating instrument.

A short training was given at the beginning of the experiment with a sample of stimuli to ensure that participants understood what is meant by the pop-out effect. The multiple-choice image displays were then presented on a computer screen. While the participant viewed the display, the experimenter read each rating question to the participant. The participant answered each question with “yes” or “no” and the instructor documented the answers on an answer sheet. The answers were recorded for each display. When copying the answers into a data base, “yes” was coded as 1 and “no” was coded as 0. The participants were encouraged to ask before and during the experiment if they were not sure about the meaning of any term.

Phase 2: Eye Movement Assessment

Participants. Twenty students who did not participate during the first phase of the experiment were chosen from the Ohio University community through academic courses. Students were asked to participate in the experiment. They were not majors in Linguistics, Psychology or Hearing, Speech and Language Sciences since students in these majors might have experience with the aphasia tests in which the samples were drawn. They were native speakers of English. Prior to enrollment in the study, potential participants were asked if they have any history of neurological impairments. Those reporting neurological impairment were excluded. All participants took part in a vision screening which included the assessment of near vision and documentation on use of
All participants passed a hearing screening at 25dB for the frequencies of 500, 1000 and 2000 Hz.

Procedure. Each participant sat in a comfortable high back chair and looked at a computer screen from a distance of 34 inches. A set of 20 multiple-choice displays was presented. The participant saw three images at a time. The image sets were chosen only from multiple-choice reading and comprehension tasks and were identical to the image sets used during the second phase of the experiment. Within each experimental display were images shown simultaneously in multiple-choice comprehension tasks similar to the traditional administration of the corresponding test.

To control for the potential influence of peripheral vision on the processing of the images, the stimuli were presented at least at 20 degrees of visual angle apart in the four corners of the screen. The participants were not asked to look at a specific target but rather to simply look naturally at the images to be displayed. The purpose of the experiment was explained to the participants after the experiment was conducted in order to prevent “purposeful looking” at certain images. Each participant’s eye movements were monitored and reported with an ISCAN RK 426 pupil center/corneal reflection system at 60 samples per second. According to Hallowell, Wertz & Kruse (2002), near-infrared light shone onto one eye creates two points of reflection, one on the pupil and one on the cornea. The reflections are then digitally recorded. Subsequent vector calculations are used to determine the eye position in relation to the visual display.
**Analysis**

Pop-out effects were detected in the form of disproportionate visual attention allocated to at least one image within an image set. For each set of the eye movement recordings of viewers’ scans of the test images an average pop-out score was calculated. This score lies between 0 and 1. A score with a value close to 0 indicates equally distributed eye movements over all three images. A value close to 1 indicates a high degree of disproportionate looking. The pop-out equation is as follows:

\[
\text{Pop-out eye movement score} = \frac{\text{Highest} - (1/ \#\text{images})}{1 - (1/ \#\text{images})}
\]

In this equation “Highest” refers to the highest proportion of fixation duration within the display which lies between 0 and 1. “\# images” refers to the number of images within the display. For this study three images were displayed in one set.

The relation between degree of disproportionate visual attention allocated within a display and ratings of the degree of visual stimulus control for each set of images was assessed through correlation analyses of the image ratings and the image recordings. A single-sample \(t\)-test statistic was conducted to demonstrate the statistical significance of the pop-out effect within the eye-movement pop-out scores.
Chapter 3

Results

Ratings of Experts and Naïve Raters

The ratings of the 20 experts correlated significantly with the ratings of the three naïve raters \((N = 20, r = .73, p < .001)\).

Ratings of the Pop-out Effect (Phase 1) and Pop-out Eye Movement Score (Phase 2)

There were no significant correlations among 12 of the 13 stimulus properties with the mean pop-out eye movement score for the 20 image sets \((p < .05)\). A significant correlation was found for one of the 13 categories between the expert rating and the pop-out eye movement score \((r = .47, p = .039)\). The experts’ rating for imageability correlated significantly with the pop-out eye movement score as shown in Table 2. No significant correlations exist between the imageability ratings of the naïve raters and the pop-out eye movement score \((p < .05)\). Correlations for depth and frequency could not be computed for the experts because all raters rated these categories with 0.
Table 2

*Correlation Between Ratings of the Stimulus Properties and the Pop-out Eye Movement Score (N = 20)*

<table>
<thead>
<tr>
<th>Category</th>
<th>Expert Raters</th>
<th>Naïve Raters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p (2-tailed)</td>
</tr>
<tr>
<td>Orientation</td>
<td>.160</td>
<td>.501</td>
</tr>
<tr>
<td>Size</td>
<td>.020</td>
<td>.933</td>
</tr>
<tr>
<td>Depth</td>
<td></td>
<td>-.419</td>
</tr>
<tr>
<td>Luminance</td>
<td>-.001</td>
<td>.997</td>
</tr>
<tr>
<td>Complexity</td>
<td>-.409</td>
<td>.073</td>
</tr>
<tr>
<td>Symmetry</td>
<td>-.190</td>
<td>.422</td>
</tr>
<tr>
<td>Clarity</td>
<td>-.235</td>
<td>.319</td>
</tr>
<tr>
<td>Context</td>
<td>-.212</td>
<td>.369</td>
</tr>
<tr>
<td>Imageability</td>
<td>.465(*)</td>
<td>.039</td>
</tr>
<tr>
<td>Viewpoint</td>
<td>.064</td>
<td>.789</td>
</tr>
<tr>
<td>Familiarity</td>
<td>-.015</td>
<td>.949</td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td>-.132</td>
</tr>
<tr>
<td>Social/cultural infl.</td>
<td>-.241</td>
<td>.307</td>
</tr>
</tbody>
</table>

*Note. (*) = significant at .05 level*
To investigate whether image ratings of the physical stimulus properties and semantic content conveyance factors were related to eye movement pop-out results, separate “physical” and “semantic” scores were calculated based on the ratings for each image set. The physical score includes the following categories: orientation, size, depth, luminance, complexity, symmetry and clarity. The semantic score includes the categories: context, imageability, perspective, familiarity, frequency, and social/cultural influences. The ratings of experts and naïve raters for 20 sets of images were not significantly correlated with the score obtained from the eye-movement recordings ($p < .05$), as presented in Table 3.

Ratings of the pop-out effect obtained with the image evaluation instrument do not correlate significantly with the recordings of viewers’ eye movements looking at the same sets of multiple-choice images ($p < 0.01$). Only 14 of 20 individuals’ eye movement recordings were used. The recordings of six participants were lost due to technical difficulties with the eye movement recordings.

Table 3

*Correlations between Ratings of Stimuli Properties and Measures of the Pop-out Eye Movement Score (N = 20)*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>-.317</td>
<td>.042</td>
<td>-.279</td>
<td>.042</td>
</tr>
<tr>
<td>$p$ (2-tailed)</td>
<td>.173</td>
<td>.861</td>
<td>.233</td>
<td>.860</td>
</tr>
</tbody>
</table>

*Note.* Phys. = physical; Sem. = semantic; N = naïve raters; E = expert raters
Analysis by Specific Tests

Analysis by specific tests precluded valid statistical analysis because of the small sample size (four images) per test. Eye movement pop-out score of the five tests are shown in Appendix B.

Analysis of the Pop-out Eye Movement Scores Obtained from the Eye-Movement Recordings

An average pop-out score was calculated for each set of the eye movement recordings of viewers’ scans of the test images. This score lies between 0 and 1. A score with a value close to 0 indicates an equally distribution of eye movements over all three images. A value close to 1 indicates a high degree of disproportionate looking. The smallest and largest pop-out scores for the image sets were .27 and .54, respectively. A single sample t-test statistic was calculated to investigate the eye-movement pop-out scores were different from 0. The pop-out scores were statistically significant different from 0, \((M = .43, SD = .09)\) \(t (19) = 21.66, p = .05\) (two tailed). The pop-out scores indicate disproportionate looking for all image sets.
Chapter 4

Discussion

The first research question examined whether independent ratings indicating distraction corresponded with viewers’ disproportionate looking at individual images within multiple-choice displays. With one exception, no significant correlations were found between the ratings of the images and the pop-out eye movement scores. The second research question, asking for specific visual stimulus characteristics and semantic content conveyance factors that influence disproportionate viewing of multiple-choice images remains unanswered.

There are three important reasons for which correlations between the individual or grouped semantic content conveyance factors and physical stimulus characteristics and the eye movement recordings were generally not significant. First, the nature of the pop-out effect is that it is detected early when looking at images. When tracking eye movements, the image sets were shown to viewers for 6 seconds. The pop-out effect might be lost after a shorter viewing duration. Second, although each rater received training and an explanation of the term “pop-out effect,” raters might have been overwhelmed by the task of detecting the pop-out effect in terms of 13 different categories. Even the expert raters had difficulty distinguishing between certain categories, such as complexity and luminance, and confused categories occasionally. Third, this study used subjective measurements for the stimuli properties. Objective measurements can be obtained for some of the properties such as orientation, luminance and frequency and might have higher validity than subjective measurements.
The research question, whether naïve viewers allocate disproportionate visual attention to pictures within multiple-choice image sets, was answered positive. Across all participants in eye movement recordings, eye movement pop-out scores indicate disproportional looking at individual images within all 20 image sets. This result highlights the importance of a better controlled image design in multiple-choice test images to provide a valid assessment of an individual’s language comprehension abilities.

The fourth research question asked if independent ratings of experienced raters corresponded differently with viewers’ disproportionate looking at individual images within multiple-choice displays, than the ratings of naïve raters. The ratings of the experts correlated significantly with the ratings of the naïve raters. Experience or training is not a criterion that influenced the individuals in their ratings.

Future Investigations

More research is needed to develop methods to assess and control for the pop-out effect in multiple-choice images. First, instead of using subjective measurements for stimuli properties, objective measurements can be obtained for categories such as size, orientation, luminance, light-dark contrast, and image frequency. Those measurements might have higher validity than subjective measurements and can be analyzed for the degree of relationship they have with eye movement derived pop-out scores. Second, control of physical stimulus characteristics and semantic content conveyance factors in image sets will lead to an equal distribution of total duration of fixation across the image set. The proportional fixation duration on each item in the display would be equal. This ideal distribution of fixations may be unobtainable even with careful control. However, a reasonable number to expect for an eye-movement pop-out score in controlled image sets
may be obtained from sets of images that are designed carefully regarding the semantic and physical stimulus characteristics. The pop-out eye movement scores for these image sets may represent reasonable numbers for well controlled design of images in multiple-choice image sets. A third research project may focus on the time span in which the pop-out effect is detected and disappears. It is possible to modify the collected data for this study. The eye movement recordings have a length of 6 seconds per image set. Systematic analyzing of the eye-movement data of this study in .5 second segments (.5 seconds, 1 second, 1.5 seconds and up to six seconds) may yield important findings regarding ideal viewing durations to capture significant pop-out effects.
References


Appendix A

Multiple-Choice Visual Stimulus Control Instrument

Many factors influence the way individuals look at sets of images. For example, viewers may be distracted by a specific detail of one image in a group of images. The bright color of one object in a relatively dark picture or the unusual size of an object compared to other objects depicted are examples of these potential distracters that make one image in a set of images “pop-out.” A viewer’s internal representation of a particular image or presentation of an image in an unfamiliar perspective or context may also distract the viewer from the intended purpose of the image. You are being asked to consider the factors that might influence the way individuals look at images.

Look at each set of images. Rate each set of images by answering the following questions with “yes” or “no”. If one or more images in a display seem to pop out because of a factor indicated in the question please answer “yes”.

Physical Stimulus Properties

1. **Color**: Did color make one or more images within the display “pop-out”?

2. **Orientation**: Did the direction to which an item pointed make one or more images within the display “pop-out”?

3. **Size**: Did size make one or more images within the display “pop-out”?

4. **Depth Cues and Shading**: Did depth cues or shading make one or more images within the display “pop-out”?

5. **Complexity**: Did the complexity of lines, shapes or items in the image make one or more images within the display “pop-out”?
6. **Luminance**: Did the light-dark contrast make one or more images within the display “pop-out?”

7. **Asymmetry**: Did asymmetry make one or more images within the display “pop-out?”

8. **Clarity**: Did a lack of clarity make one or more images within the display “pop-out?”

Semantic Content Conveyance

1. **Scene context**: Did the scene context make one or more images within the display “pop-out?”

2. **Imageability**: Did the degree of conceptual abstraction make one or more images within the display “pop-out?”

3. **Viewpoint**: Did the perspective in which the images are presented make one or more images within the display “pop-out?”

4. **Image familiarity**: Did the degree of unfamiliarity of the concepts conveyed by these images make one or more images within the display “pop-out?”

5. **Concept frequency**: Did the frequency with which you tend to see concepts corresponding to these images make one or more image within the display “pop-out?”

6. **Social and Cultural Influences**: Might a person’s specific social or cultural experience make one or more images within the display “pop-out?”
Appendix B

Analysis by Specific Tests

The eye-movement pop-out score of each image set is represented in figure B1. The higher the score, the higher is the degree of unequally distributed durations of fixations across the images in the display. The lower the score, the more equally are the numbers of fixations distributed across the images.

Figure B1. Eye movement pop-out score of each image set.

Note. 1 = Psycholinguistic Assessments of Language Processing in Aphasia (PALPA)
2 = Test of Communicative Abilities in Daily Living (CADL)
3 = Boston Diagnostic Aphasia Exam (BDAE)
4 = Reading Comprehension Battery for Aphasia (RCBA-2)
5 = Western Aphasia Battery (WAB)
Four sets of images were derived from each of the 5 tests. The mean pop-out score was calculated regarding each test. Differences in the mean eye-movement pop-out score among the tests are presented in figure B2.

![Chart showing mean eye-movement pop-out score for each test]

**Figure B2.** The mean eye-movement pop-out score of individual aphasia tests.

*Note.* 1 = Psycholinguistic Assessments of Language Processing in Aphasia (PALPA)  
2 = Test of Communicative Abilities in Daily Living (CADL)  
3 = Boston Diagnostic Aphasia Exam (BDAE)  
4 = Reading Comprehension Battery for Aphasia (RCBA-2)  
5 = Western Aphasia Battery (WAB)