Putting Words to Pictures: Defining and Categorizing the Range of Static Visuals in Multimodal Research

Jannah Fusenig[1], Lauren M. Singer Trakhman[2], & Patricia A. Alexander[2]

[1]College of Behavioral Sciences, California State University Chico [2]College of Education, University of Maryland

Aim

- In the current study, we conducted a two-phase investigation of static visuals in text and their effects on comprehension.
- Phase 1 entailed a systematic review of recent research involving static visuals in extended discourse (i.e., a page or more) that measured comprehension.
- In Phase 2, we drew on the findings from Phase 1, along with other available resources (e.g., theoretical and review pieces), to proffer a corpus of terms describing the range of visuals encountered within the multimodal literature
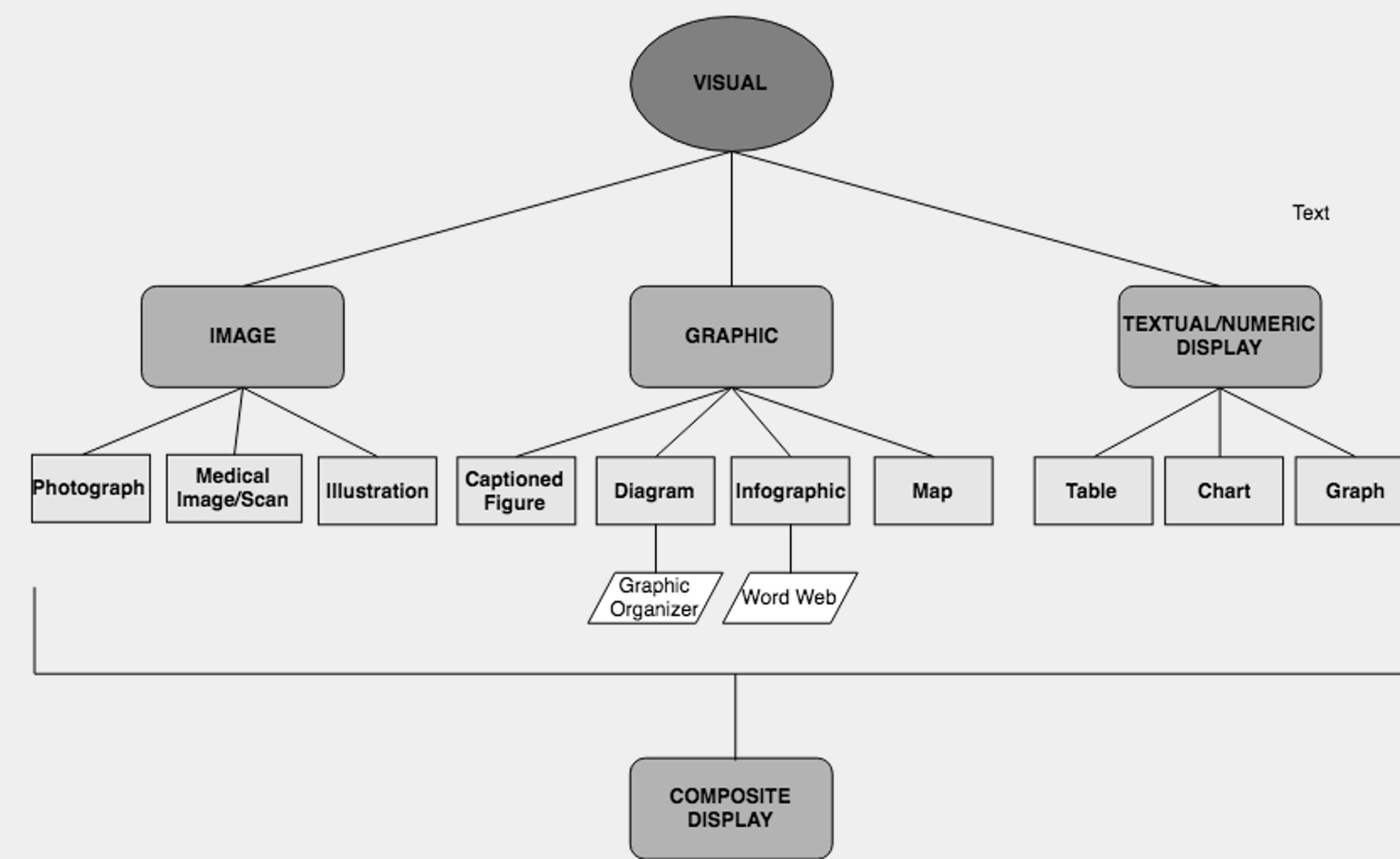
Background

- Within the multimodal literature, interest has been directed toward the function that visuals serve in conveying meaning vis-à-vis the written document, and characteristics of the visual-text pairing are presumed to facilitate or frustrate student learning (Clark, Mayer, & Thalheimer, 2003).
- There are concerns over the terms researchers use for visuals, the ecological validity of the visual and textual materials used, the semantic relations presumed to exist between visual and text, and characteristics of learners.
- We have intentionally chosen to refer to the range of non-textual elements appearing in written documents as "visuals."

Research Questions

1. How do researchers identify and describe the visuals they use in their multimodal studies and are those visuals researcher-generated or naturally-occurring?
2. What is the nature of the texts used in this literature in terms of their length, genre, and subject-matter and was the semantic congruency of the visuals and text considered?
3. Who are the participants in these studies and what information about their topic knowledge, reading proficiency, or visual literacy is provided?
4. How is comprehension measured and what effects are reported or conclusions about the visual-text pairing are reached?

Proposed Corpus of Terms

Proposed Definitions

TYPE	DEFINITION
Visual	A static external representation.
Image	A visual display that solely includes pictorial information.
Graphic	A visual display that combines an array of pictorial information with any form of symbolic representation.
Textual / Numeric Display	A visual display using text and/or numeric values to convey information in a concise and organized fashion-- excluding the use of images.
Composite Display	A visual display made up of two or more visual types.
Photograph	A visual display produced by photography using a camera.
Medical Image/Scan	Image generated from medical procedure such as an X Ray, EKG, EEG, or fMRI.
Illustration	A visual display that can convey an abstract design or rendering of persons, objects, places, or events using artistic mediums.
Captioned Figure	A visual display that may include text and/or numeric values in accordance with images to convey information in a concise and organized fashion.
Diagram	A visual display designed to show conceptual relation and depict the organization and structure of key concepts.
Infographic	Visual representations created principally to inform through imagery that may include both text and numbers, infographics aim to convey information in a clear, concise and often pleasing way.
Map	A visual display showing spatial relations.
Chart	A visual display used to represent data normally using bars, lines, or slices.
Table	A visual display using columns to display figures and information.
Graph	A visual display showing relations among variable quantities.
Graphic Organizer	A visual display that demonstrates relationships between facts, concepts or ideas.
Word Web	A visual display that is used to show how words and ideas are related.

Example of Using Terms Interchangeably

"When glancing at a magazine, or browsing the Internet, we are continuously being exposed to **photographs**. Despite of this overflow of visual information, humans are extremely good at remembering thousands of **pictures** along with some of their visual details. But not all **images** are equal in memory."

- Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011, June). What makes an image memorable?. In CVPR 2011 (pp. 145-152).

Conclusions

Visuals
- The majority of studies involved a single visual representation that was only cursorily described by researchers. In fact, researchers only rarely defined or delineated the salient attributes of visuals (20.7%, n=6).

Texts
- For the most part, researchers relied on exposition addressing a range of topics (e.g., anatomy, biology, concrete construction, a pulley system, and geography). The exact length of the texts could only be discerned for 19 articles (65.5%) and that length ranged from 168 to 2,165 words.

Participants
- Due to the inclusion criteria, study participants were 12 to 25 years of age. This age group was selected to ensure the participant pool consisted of advanced readers that were familiar with educational materials and settings. Of the analyzed studies, 80.8% (n=21) used university students from Germany, US, China, Japan, Italy, Israel, Taiwan, and the Netherlands.

Comprehension Measures and Outcomes
- Comprehension was measured with multiple-choice items in 17 articles (58.6%), while 12 (41.4%) employed open-ended questions. We identified no studies in our search that used standardized reading measures, with all opting for a researcher-developed comprehension measure.