GLOBAL/LOCAL PROCESSING IN INCIDENTAL PERCEPTION OF HIERARCHICAL STRUCTURE

by

Mark Mills

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Psychology

Under the Supervision of Professor Michael D. Dodd

Lincoln, Nebraska

August, 2016

GLOBAL/LOCAL PROCESSING IN INCIDENTAL PERCEPTION OF HIERARCHICAL STRUCTURE

Mark Mills, Ph.D.

University of Nebraska, 2016

Advisor: Michael D. Dodd

The goal of the current thesis is to provide a framework for investigating and understanding visual processing of hierarchical structure (i.e., local parts nested in global wholes, such as trees nested in forests) under incidental processing conditions—that is, where processing of information at global and local levels is both uninformative (cannot aid task performance) and task-irrelevant (need not be processed to perform the task). To do so, a novel method combining two widely-used paradigms (spatial cueing and compound stimulus paradigms) is used for implicitly probing observers' perceptual representations over the course of processing. This compound arrow cueing paradigm was used in five experiments to address a series of objectives. First, which level (global or local) is more dominant in the evolution of a percept? Relatedly, is the temporal structuring of global and local processing fixed or flexible? And what is the time course of level-specific advantages—do they occur earlier or later in the course of processing, and do they follow a transient or protracted time course? Finally, what controls level-specific selection (sensory, perceptual, and/or attentional factors)? The results of the five experiments addressing these issues contribute to a greater understanding of visual perception by elucidating the nature of global and local processing under incidental processing conditions.

Acknowledgements

I would like to thank the University of Nebraska for the opportunity, space, and financial resources to pursue my graduate studies, without which this thesis could not have been undertaken. I would also like to thank my lab mates, Gerald McDonnell, Monica Rosen, and Brett Bahle, for their insight and support throughout the development of this thesis. In addition, I would like to thank my brother, Eric Mills, for sage advice when it was critically needed, as well as my parents, Charles Mills, Gerda Mah, Cheryl Goodgine, and Bill Goodgine, for their steady support—I am indebted. I would also like to acknowledge the training and mentorship of Lesa Hoffman, without whom many of my research pursuits would not have been possible. Finally, none of this would have been possible at all without my advisor, Mike Dodd—I am eternally grateful for his guidance, support, and encouragement during the development of this thesis and throughout the course of my PhD studies.

Table of Contents

List of Tables

List of Figures

CHAPTER 1 PREFACE: THESIS SUMMARY AND OUTLINE

Visual environments (scenes, objects, faces) can be conceptualized as containing global and local information, where global information corresponds to overall form and local information to finer-grain detail. An important question concerns how information across levels contributes to scene understanding. To investigate this, a common approach is to present observers with compound stimuli (i.e., stimuli containing hierarchical structure such as larger figures constructed of smaller figures) and measure responses to a target presented at either the global or local level. The preponderant findings are global advantage (speeded responses to targets presented at the global versus local level) and global interference (greater difficulty ignoring global versus local distractors). On the basis of these findings, Navon (1977) proposed a perceptual global precedence hypothesis which states that identification of global aspects is a faster and obligatory process completed before identification of local aspects, which is a slower and optional process. As a result, global aspects are dominant in the percept (giving rise to global advantage) and their availability for selection is constant (together giving rise to global interference). Subsequent work has corroborated the empirical findings of global advantage and interference but ultimately rejected Navon's claim that the nature of these effects is obligatory and their source perceptual.

The current thesis reviews issues in global/local research (summarized in Table P.1), highlights some limitations of how previous work has addressed these issues, and then reports a series of experiments aimed at circumventing these limitations. It is argued that rejection of perceptual global precedence is premature—namely, because in all

previous studies observers have been required to process the imperative compound stimulus in a task- or stimulus-driven manner. Accordingly, the most important aspect of the current experiments is that visual processing of global and local levels was measured under incidental processing conditions—that is, where processing of global and local levels is uninformative (does not aid task performance) and task-irrelevant (need not be processed to perform the task). In these experiments, participants made simple detection responses to a peripheral target that was preceded by a centrally presented compound stimulus (big, global arrow composed of smaller, local arrows). Critically, the compound stimulus was irrelevant to the detection task (i.e., it did not need to be processed in order to respond to the peripheral target) and also spatially uninformative (i.e., the direction of global and local arrows did not predict target location). Thus, processing of the compound arrow was not necessary to perform the task and could not be used for making reliable inferences regarding the location of the target, meaning any evidence that it was processed can be taken as incidental.

Table P.1. Summary of the main issues in global/local research addressed in the current experiments.

Main Issues Addressed in the Current Experiments
Issue #1: Temporal organization in the percept (global-to-local or local-to-global)?
Issue #2: Temporal availability in percept (stationary or variable)?
Issue #3: Nature of global/local precedence effects (obligatory or voluntary)?
Issue #4: Source of global/local precedence effects (perceptual or attentional)?

Experiment 1 was a multi-experiment study that examined global and local processing under incidental processing conditions. The critical manipulation was processing duration (operationalized as stimulus-onset-asynchrony; SOA). The rationale

was that the priming potency of a level (i.e., global or local) should vary with SOA in correspondence with how early that level was processed (Kimchi, 1998; Navon, 1991). Accordingly, the finding of a level-specific spatial cueing effect would reflect a processing advantage for that level, and a non-significant interaction with SOA would reflect a stationary advantage of that level over the course of processing. Experiment 2 replicated Experiment 1 using a more widely-tuned and finer-grained manipulation of processing duration in order to characterize precisely the functional form of change in level-specific advantage over time. The finding of significant "jumps" in the size of level-specific advantage in the course of processing would reflect discrete change in the availability of that percept, whereas non-significant "jumps" and smooth or no change would reflect gradual change in the availability of that percept. Experiments 3-5 modified the compound arrow cueing paradigm in order to examine control of level-specific selection in perception of hierarchical structure. To examine the role of attention in level selectivity, Experiment 3 manipulated the task-relevance of a level via task instructions that called for processing of either a global or local aspect of a compound arrow. The finding of a task-compatible, level-specific advantage (i.e., a local advantage in the local orienting task and a global advantage in the global orienting task) that was independent of processing duration would reflect an influence of top-down attentional control. Experiment 4 examined the role of bottom-up salience in level selectivity. To do so, the local level was modified to include a color singleton. The finding of a larger and more rapidly arising local advantage with cues containing a local color singleton compared with those that are homogenously colored would reflect a bottom-up attentional source of

level-specific advantage. Finally, Experiment 5 examined the role of attentional focusing and adjustments in level selectivity. To do so, the compound arrow cueing paradigm was modified to include a go/no-go component, which served as a manipulation of attentional focus. The finding of a window-compatible, level-specific advantage regardless of processing duration may suggest an alternative account of global/local processing phenomena, one predicated on the focusing an attentional window rather than availability in the percept.

Table P.2. Summary of the primary research question in each experiment and the main finding addressing it.

Exp.	Primary Question Addressed	Main Finding
1.1-1.4	Is global precedence observed in incidental perception of hierarchical structure?	Global-to-local shift in advantage over time.
2	Does change in availability of global/local features proceed gradually, or discretely?	Gradual, as opposed to discrete, shift in advantage in over time.
3	Does top-down task set control level selectivity? In particular, will a global orienting task cause global precedence and a local orienting local precedence?	Task-compatible, level-specific advantage with conflict cues; global advantage with task-incompatible, level-specific neutral cues.
4	Does bottom-up salience control level selectivity?	Local salience reduces, but does not reverse, global advantage.
5	Do top-down and bottom-up factors interact to control level selectivity--e.g., if advantage depends on salience, is such bottom-up influence contingent on the focal scope of attention (focused or diffuse)?	Focal scope of attention modulates global advantage, though, not independent of physical salience.

The main finding in each of the five experiments, and the research question it addressed, are summarized in Table P.2. The emerging picture is that something like global precedence guides perceptual microgenesis (i.e., the time course of development in the percept) in incidental perception of hierarchical structure: processing of the global level was obligatory, whereas processing of the local level was somewhat optional. In short, the experiments show that interactions between task demands and the structure of the input information selectively modulate the relative needs of visual information at different levels of stimulus structure. It seemed that a global-to-local shift in advantage was obligatory when there was conflict between levels and processing was incidental, as well as that processing could be restricted to a task-relevant global level but not to a task-relevant local level. These findings suggest that level-specific advantages may not have been due to top-down task set per se but rather to conflict between levels, with the effect of task set serving to modulate the availability of level-specific information, effectively prolonging (in the case of task-relevant global information) or expediting (in the case of task-relevant local information) the advantage at a given level. To select symbols at different levels of structure, therefore, the current thesis concludes that at least two factors need to be considered: the observer's current task, which specifies the demands of visual information from the input, and the globality of this visual information across the percept, which specifies the availability of the percept for selection.

In sum, a novel compound arrow cueing paradigm was used in five experiments to examine fundamental issues of broad theoretical interest (entry point of visual perception and its flexibility; selection of competing symbols at different levels of

stimulus structure). The results of these experiments contribute to a greater understanding of visual perception by elucidating the nature of global/local processing under incidental processing conditions, which closely reflect the manner in which a great deal of information is processed in the real world. Though attention, perception, and action can be strongly influenced by goals and intentions, processing of stimuli routinely occurs in a passive manner given the overwhelming number of inputs available at any given time. As such, the results reported have broad implications for the study of attentional control in general and perception of hierarchical structure in particular.

Thesis Outline

Chapter 1 (Overview of Global and Local Processing) reviews fundamental issues in the study of visual perception and how contemporary theories have addressed these issues. Chapter 2 (Thesis Objectives, Methods, and Hypotheses) describes the primary objectives of the current thesis, the methods used to achieve these objectives, and the hypotheses underlying each experiment. Chapters 3-7 are empirical chapters, with chapters 3-4 reporting experiments that investigate global/local processing in incidental perception, and chapters 5-7 reporting experiments that investigate control of level-specific selection. Chapter 3 (Experiments 1.1 through 1.4: Global and Local Processing in Incidental Perception) introduces the compound arrow cueing paradigm and incidental viewing procedure designed to investigate the role of globality (i.e., an item's relative position—more global or local—in hierarchical space) in perception of hierarchical structure under incidental processing conditions. Chapter 4 (Experiment 2: Capturing Time Sensitive Effects) investigates the time course of the availability of global and local

information in the microgenesis of the percept. Chapter 5 (Experiment 3: Role of Top-Down Task Set) investigates top-down mechanisms of level selection. Chapter 6 (Experiment 4: Role of Bottom-Up Salience) investigates bottom-up mechanisms of level selection. Chapter 7 (Experiment 5: Role of Attentional Focusing and Adjustments) investigates hybrid mechanisms of level selection. Finally, Chapter 8 (General Discussion) summarizes and discusses the experiments reported in the preceding chapters, as well as identifies current issues in the study of global/local processing and makes recommendations for future research.

CHAPTER 1: OVERVIEW OF GLOBAL AND LOCAL PROCESSING

1.0 Chapter Overview

Chapter 1 provides an overview of fundamental issues in global/local research (focusing primarily on the visual domain), highlights a number of limitations in how previous work has addressed these issues, and concludes with a discussion of how the experiments reported in the current thesis circumvent these limitations. To introduce the concepts involved in the issues at hand, Chapter 1 begins with an example illustrating some of these.

1.1 Challenges to Visual Perception

1.1.1 An Illustrative Example

Figure 1.1 shows two common visual symbols with hierarchical structure, a STOP sign and an EXIT sign. An important question is how does the visual system know that the word (“STOP” or “EXIT”) and form (octagon or rectangle) should be paired to represent a STOP sign and an EXIT sign, respectively? In cases of a busy intersection, a red sign with big, bold, white block letters is critical to driver safety. Similarly, in cases of a fire, a red sign with big, bold, white block letters will guide one to safety. Of course, the effect of these two symbols on behavior are polar opposite, with the latter activating a behavior (to move toward the exit) and the former suppressing a behavior (to stop moving toward the intersection). Thus, with the exception of form, a “STOP” sign is perceptually equivalent to an “EXIT” sign, yet cues an opposing behavior (to ‘stop’ rather than to ‘go’). Moreover, the cueing properties of each are also similar, as each sign likely has relatively automatic or effortless effects on behavior. How, then, does the same

basic symbol elicit two opposing, automatic responses? What are the control processes involved and how is this flexibility in behavior achieved? The current thesis aims to understand the control processes involved in coordinating the time course of visual processing in support of flexible, goal-directed behavior.

Continuing with the example in Figure 1.1, it is proposed that the binding of word and form occurs because each is trying to communicate a single, coherent message (to STOP or to EXIT)—that is, they group by common function (i.e., communication of symbolic value). In this view, then, the reason STOP and EXIT signs cue rapid and effortless, but opposing, responses in spite of reasonably equivalent perceptual features is because their messages differ, meaning that the function on which their respective constituents group differs, which in turn leads to a different percept and a different response. That is, as different functions or tasks are likely to tap different stages of processing or to evoke different optional strategies available to the processing system, this leads to different percepts and thus different responses.

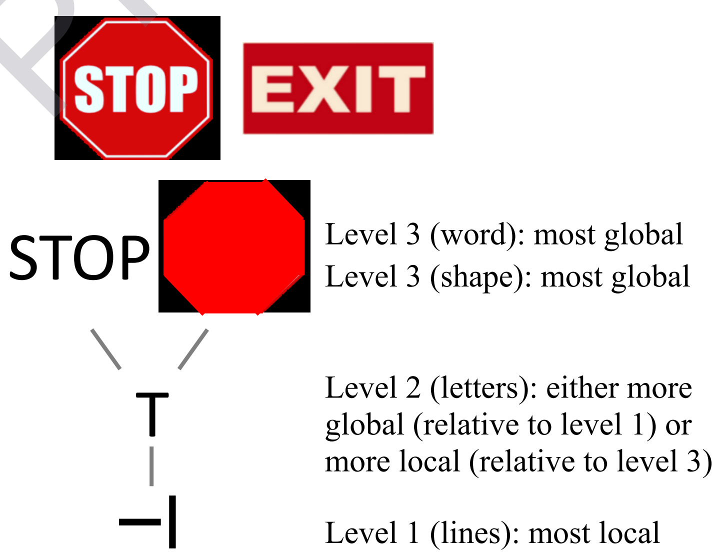

Figure 1.1. Visual symbols with hierarchical structure (STOP and EXIT signs), in which lines are nested within letters, which in turn are nested within a word and a form and where word and form are crossed at the highest level of globality (i.e., relative position in hierarchical space) given that each line-letter combination appeared with both word and form in a given sign.

1.1.2 Selection Problem in Perception and Action

A control process is any process that helps to overcome rigid responding. Cognitive control processes are particularly required in situations of task uncertainty, in which no cue is given for the selection of the relevant action, or in situations that produce problems in selecting which action or task to perform at any given moment. Many common behaviors and visual scenes constitute such situations (Monsell, 2003). This is because everyday visual environments contain more information than can be processed within a glance, which can be appreciated by considering that humans make 3-4 eye movements each second (Henderson, 2003; Henderson & Hollingworth, 1998) or several billion (10^9) over the lifespan. Despite this seemingly impressive rate of orienting, real-world visual environments are often too complex to allow for visually guided behavior. Even under the obtuse assumption that eye movements are random, at this sampling rate, localizing a target in scene with only 24 distractor locations would take 3-4 seconds. Humans and other animals, therefore, have to cope with an enormous amount of perceptual data.

Similarly, everyday visual environments also continuously bombard humans and other animals with various opportunities for action. Some stimuli automatically elicit the

same behavioral response from an observer each time that particular stimulus is encountered (e.g., visual transients cause reflexive oculomotor orienting), whereas other stimuli afford a variety of alternative responses depending on situational constraints, learned associations, and the goals of the observer (e.g., when standing in front of a sink there are numerous behaviors that could be activated such as washing one's hands, brushing one's teeth, washing dishes, getting a drink of water, or fixing one's hair amongst a multitude of other options). Because one cannot perform all possible actions at the same time, there must be mechanisms to reduce the number of potential actions. To overcome the computational limitations imposed by the environment and enable efficient, flexible behavior amenable to the current goals of the observer, goal-driven perception and action depend on selective attention for allocating limited resources towards a subset of relevant items—that is, effective and efficient goal-directed behavior depends on a combination of processes that suppress irrelevant responses from interfering with one's goals and that select and initiate responses relevant to the completion of that goal (collectively regarded as executive control processes). But what gets selected, a whole scene or just certain parts? And how is efficient selection achieved? The current thesis aims to understand the factors that control orienting and enable efficient selectivity in sampling behavior.

Approaches to selectivity. One way to deal with this selection problem is through extensive filtering, which should minimize the use of resources and processing effort, as both of these are limited. Another approach to the selection problem is to make the scene virtually larger—by zooming out the focus of attention—so that the information

can be held in low-cost storage. Yet another approach is to offload work with automated visual routines (e.g., obligatory grouping and parsing processes involved in perceptual organization), which aid object constancy and thereby effectively shift cognitive processing load to the perceptual system. Each of these approaches to the selection problem is intended to decrease the costs for performing information-intensive tasks (resource accounts), or, alternatively, to increase the scope of the information that can be utilized for the same cost (adaptation accounts). The current thesis describes these approaches in greater detail below and then reports a series of empirical studies addressing these possibilities.

1.1.3 Hierarchical Structure

An important property of real-world objects is that they tend to be structured hierarchically. The human body, for example, is composed of a number of parts that cannot exist independently. A hand, for instance, cannot exist without an arm, which cannot exist without a torso, and so on. Accordingly, visual scenes (scenes, objects, faces) are generally conceptualized as containing global and local information, where global information corresponds to overall form and local information to finer-grain detail (Neisser, 1967). Hierarchical levels of structure (fingers within hands within arms within person), therefore, may be ordered from global (body) to local (fingers), where more global levels contain or encompass local constituents. Such structure poses a challenge to the visual system because the identity of an item with hierarchical structure is potentially ambiguous in that its interpretation depends on which level of globality (i.e., the relative position of an item in hierarchical space) is dominant. For example, a herd of sheep

arranged in a circular pattern is a potentially ambiguous stimulus: given strict interpretation of the global level, it would be interpreted as a plain circle; given strict interpretation of a more local level, it would be interpreted as a herd of sheep. Visual perception, therefore, requires the coordinated processing of objects on both global and local levels. During global processing local elements are grouped into perceptual wholes (Koffka, 1935; Kohler, 1947; Wertheimer, 1955), whereas during local processing scene elements are analyzed as individual objects (Titchener, 1909; see also Han, Weaver, Murray, Kang, Yund, & Woods, 2002).

The visual world is cluttered and complex, typically containing a multitude of objects that can appear in an infinitude of orientations, sizes, shapes, colors, and so on and which are often degraded by atmospheric conditions and occluded by other objects. Despite such vagaries of sensory experience, the phenomenological experience of perception is one of instantaneous completeness and coherence imbued with meaning. How is this meaning apprehended? A challenge to understanding the temporal dynamics of visual perception is that even when the physical properties of a stimulus are constant over time, one's perception of that stimulus may still change as a function of processing duration. For instance, a brief glimpse of an unfamiliar object or scene may be sufficient to categorize it at a basic level (e.g., table or indoor room), but to perceive the table as a 'coffee table' or the indoor room as a 'kitchen' tends to require additional processing time (Potter, 1975, 1976). Captured in this example is the notion that there is a temporal order to visual processing, wherein the skeletal-structure of a stimulus tends to be processed more rapidly than finer grain details. That is, visual processing is held to

progress in global-to-local fashion (Navon, 1977). In terms of metaphor, the one most commonly applied in describing the typical course of visual processing is "perception of the forest before the trees". Whether or not the speeded processing of global properties has meaningful consequences for cognitive and behavioral activity is still under investigation (see Hegdé, 2008; Kimchi, 1992, 2003a,b, 2012, 2014; Navon, 2003; Morrison & Schyns, 2001; for reviews). Can people actually use global information to classify and act on objects and scenes? Or is this information processed simply to constrain local processing?

1.2 Global and Local Processing

1.2.1 Compound Stimulus Paradigm

To investigate how the availability and use of information at different levels of hierarchical structure contributes to scene understanding, researchers have made extensive use of compound stimuli (i.e., stimuli containing hierarchical structure such as larger figures constructed of smaller figures), introduced first by Asch (1962) and later by Kinchla (1974, 1977), and popularized by Navon (1977). In Navon's compound stimulus paradigm (Figure 1.2), observers are presented with a compound letter (e.g., a large, global F constructed from smaller, local Fs) and instructed to respond to either the larger (global) or the smaller (local) letter in separate blocks of trials.[1] Importantly, the relation

[1] There are actually two versions of Navon's (1977) compound letter paradigm, which differ in terms of whether attention is directed to a particular level via task instructions or stimulus manipulations that directly call for the selection of one level while ignoring the other or is divided between levels. As the directed attention version is the more widely used one, it is this version to which the current proposal refers where discussed, unless otherwise noted.

between levels is either consistent (global F, local Fs) or inconsistent (global F, local Hs). The preponderant finding is that the global letter is responded to faster (global advantage) and is more difficult to ignore (global interference). The allure of the paradigm, apart from its elegant simplicity, is that it serves to equate the complexity, identifiability, familiarity, codability, and relative diagnosticity of the levels, such that they differ only in level of globality (i.e., relative position in hierarchical space). Additionally, the two structures can be independent such that one cannot necessarily be predicted from the other. Compound letters, therefore, satisfy two conditions considered by Navon (1977, 1981, 2003) to be critical for testing the global precedence hypothesis.

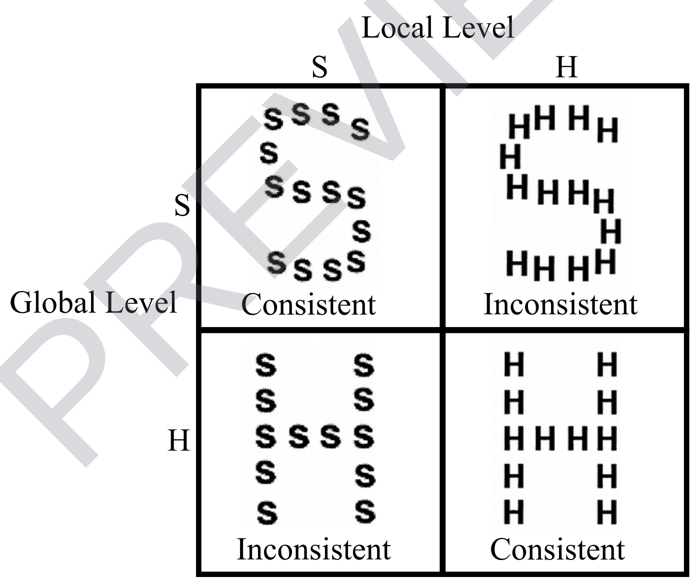

Figure 1.2. Example hierarchical stimuli in Navon's (1977) compound stimulus paradigm, in which a compound letter is presented (e.g., large, global F constructed from small, local Hs) and observers are instructed to identify the letter at either the global or local level in separate blocks of trials. Importantly, the relation between levels is either consistent (global F, local Fs) or inconsistent (global F, local Hs).

1.2.2 Global Precedence