Title
Visual Search in Microscopy Implies High-Level Cognition

Permalink
https://escholarship.org/uc/item/2j6799bs

Journal

ISSN
1069-7977

Authors
Chariker, Julia H.
Claudio, Natalie M.
Fell, Ronald D.
et al.

Publication Date
2006

Peer reviewed
Visual Search in Microscopy Implies High-Level Cognition

John R. Pani, Ron D. Fell, Natalie M. Claudio, and Julia H. Chariker

Abstract

Analysis of visual search data suggests that the use of the microscope in the study of microanatomy is a high-level visuo-cognitive skill. Although the process of interpreting a microscope slide begins with perception and recognition, it quite often incorporates recall of anatomical knowledge and reasoning about the slide in the context of anatomy. Due to the system of cognitive processes in which slide interpretation takes place, patterns of visual search do not strongly differentiate better from poorer performers in general or for particular slide images.

Keywords: medical imaging; visual cognition; visual search; microscopy; perception; reasoning; scientific reasoning; education.

Visual Search in a Real World Domain

Creation of visual artifacts is characteristic of human cognition, giving us pictures, diagrams (including maps), instrument readouts (e.g., radar screens), and a variety of technological displays that are characteristic of modern science and technology (e.g., microscopes, telescopes, and radiological images). The use of some visual displays is so deeply ingrained in the practice of formal and professional disciplines that these displays may be considered to be part of visual information systems. Study of cognition in relation to visual information systems provides an opportunity to study visual cognition at the highest level of cognitive organization. In addition, it has substantial practical value, providing opportunities to improve practice and training in numerous domains of human skill (e.g., Crowley, Naus, Stewart, & Friedman, 2003; Hoffman, 1984; Hoffman & Markman, 2001; Kundel, 2000; Kundel, Nodine, & Carmody, 1979; Lesgold, Rubinson, Feltovich, Glaser, Kopfer, & Wang, 1988; Pani, Chariker, & Fell, 2005; Wolf, Horowitz, & Kenner, 2005).

Our own recent work has focused on the interpretation of microscope slides in histology, the microanatomy of biological tissue. We have argued that microscopy is not solely a process of perceptual learning and recognition of the microscope images. Certainly interpreting a microscope slide begins with perception and at least some partial recognition. Success in this mode, however, is surprisingly brittle, due to variations in the tissue and its presentation in the slides. Very often, perception and recognition are incorporated into a goal-directed system that includes exploration of the slide, retrieval of anatomical knowledge from memory, and reasoning about the microscope slide in relation to anatomy. The final result is a high-level form of recognition in which the microscope slide is understood to be a particular sample from, or a point of view on, an instance of whole tissue. Overall, reasoning-driven recognition is the fundamental visual skill of the histologist.

These conclusions were based primarily on verbal protocols and structured interviews from the interpretation of microscope slides by graduates of a college course in histology. The present paper focuses on visual search behavior during the interpretation of the microscope slides. Data about visual search come from video recordings of the movement of the microscope stage -- movement that centers the object of interest in the microscope view -- during naturalistic identification and description of tissues. We will argue that the visual search data support the conclusions drawn from the analysis of verbal data.

Medical imaging is an area in which there has been intensive study over recent years, with the great majority of work directed toward cognition in x-ray radiology (see Krupinski, Kundel, Judy, & Nodine, 1998). The cognitive task under study typically is detection of abnormalities in tissue that is otherwise well-defined ahead of time. Microscopy is different in many ways from x-ray radiology, and there is a potential for cognition in this domain to differ from what is common in radiology. This is especially true where microscopy is used to learn and practice a basic science, such as histology. In this case, a large domain of conceptual knowledge is learned in part through exploration in the microscope (Ross, Kaye, & Pawlina, 2003). It is an empirical question just how domain-general learning of microanatomy and domain-specific detection of abnormalities are related (see also Crowley et al., 2003).

The Challenge of Microscopy for Cognition

Systematic use of a domain of visual structure for purposes of gaining information about a second, target, domain establishes a visual information system. In such a system, there is a mapping from the objects in the target domain into the structures in the information domain. The user must work the mapping in reverse, going from the information domain to the target domain. Whereas carefully designed visual information systems, such as musical notation, contain a straightforward mapping, exploratory visual information systems, such as those that employ x-ray machines, microscopes, and telescopes, can present a substantially greater challenge for the user.

Certainly objects in the information domain often look different from their counterparts in the target domain. Microscope slides are formed by taking a thin slice through the inte-
rior of a structure and staining it in order to make various structures clearly visible. With this transformation, a single curved tube can appear in the microscope as a set of separate circular and oval cross-sections. However, the challenges for learning microscopy in histology go well beyond this. A single tissue can appear in many different ways in a microscope slide, depending on the angle of cut, its position, and the type of stain that was used in slide preparation. Moreover, there is substantial variation among instances of a single type of tissue. Compounding the problem, tissues that are actually quite different can look the same if a single thin slice through their interiors is all the visual information that is given about them. Overall, the mapping from target to information domain in histology can be both one-to-many and many-to-one, creating a substantial challenge for cognition.

An analogous set of problems occurs in x-ray radiology, where the interpretation of an image can be quite challenging even for practitioners with years of experience (e.g., Nodine & Kundel, 1987).

For the present study, we treated movement of the microscope stage as analogous to eyemovements in normal perception. The microscope is used actively in histology as a tool for exploration, and histologists become highly skilled with it. The slide stage is moved smoothly in a horizontal plane with one hand using two knobs that provide continuous control. Magnification is changed often with the other hand. Standard practice is to gaze straight ahead into the microscope and to move the slide under the gaze. Subjectively, it appears that the microscope is being moved over the tissue.

We explored a series of questions about visual search in microscopy. Apart from the time at which an identification of tissue took place, can a record of visual search differentiate easier from more difficult slides? Can a record of visual search differentiate correct from incorrect answers? Can sets of visual search records differentiate more from less skilled individuals? Where cognition is driven immediately by perceptual recognition, records of visual search may provide such differentiations. Where visual recognition is feeding into a larger system that engages in hypothesis testing, records of visual search should have a more variable and indirect relation to success in the task.

Visual Search in the Microscope

Participants

Five undergraduate students, three females and two males, participated in two sessions, each lasting approximately one and one-half hours. All students were in the pre-medical or pre-dental curricula. Four of the students had received a grade of A in the course and one had received a grade of B. All students had completed the undergraduate course in histology within the previous year.

Method

Materials. Four histological slides were viewed through one head of a two-headed laboratory microscope. A slide from the scalp was expected to be easy for the students to identify and describe. The scalp is complex, with numerous intermingled structures. However, it contains several salient diagnostic structures (e.g., hair follicles), it was a tissue that the students had all studied in class, and the stain in this particular slide was familiar to the students.

A section from a tendon was a simple tissue that all the students had studied, and the stain was a familiar one. However, it was expected to be somewhat challenging to identify, because the collagen fibers that often can be seen in a tendon were not easy to discriminate in this slide.

A slide of the pancreas showed a tissue that the students were familiar with, but the stain on this slide was one with which the students were unfamiliar. The slide was moderately complex, with several structures to identify.

The epiglottis was a complex tissue that the students studied in the textbook but had not seen in a slide. The slide contained many structures common in other parts of the body, and the stain was one that was familiar to the students. Correct identification required knowing a configuration of structures rather than a single diagnostic characteristic.

Procedure. This study included two sessions for each participant. The first session consisted of a verbal protocol followed by a structured interview. For the verbal protocol, participants verbalized their thoughts as they viewed the four slides under the microscope. Participants were asked to “think aloud” as they viewed each slide. They were assured that they were not being tested; instead, the objective was to understand the natural process of slide reading. They were encouraged to change focus and magnification as needed, and to follow their own pace. After the verbal protocol was completed for all four slides, the structured interview took place. Each slide was viewed under the microscope a second time, and a series of questions was asked. The majority of questions referred to structures that had been omitted or misidentified earlier.

Results

Performance. We have previously reported basic performance data from this study (Pani et al., 2005). We present a brief review of performance here to provide background for an examination of visual search. Table 1 indicates which slides were identified by which participants during the verbal protocol.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Scalp</th>
<th>Tendon</th>
<th>Pancreas</th>
<th>Epiglott.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>p2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>p3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>p4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>p5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 1. Participant by tissue matrix. A ‘1’ in a cell indicates successful identification.
The marginal sums indicate that there was large variation in the difficulty of interpreting the microscope slides and large individual differences among the students. Analyses of verbal protocol data suggested that cognition took a variety of forms. One common form was the recognition of diagnostic features and immediate inference of the identity of the whole tissue. Hypothesis testing was equally common. Quite often it employed a use of negative evidence to disconfirm a hypothesis.

**Visual Search: Transcription.** Large printed images of the microscope slides were prepared for each slide viewed by each participant. The video recordings were viewed interactively at slow speeds. For each recording, the center of the microscope view was transcribed onto the image of the slide as a continuous directional path. Numbered location markers were placed at most changes of direction, initiation and termination of stopped motion, and moderately sized intervals along otherwise long segments. Elapsed time was associated with each numbered location on the path, and the difference in elapsed time was associated with each pair of adjacent locations. With this information, rate of motion over the slide and the location of pauses could be easily related to the morphology of the tissue. In addition, visual search behavior could be related to structures on the slides and to the progress of identification and description of the tissue. Twenty visual search records were generated for the five participants and four microscope slides. A portion of a visual search record is shown in Figure 1.

**Global properties of motion over the slides.** The global pattern of exploration of the slides was strikingly uniform, suggesting that the students had learned a standard form of exploration in the lab sections of the histology course. In particular, nearly every visual search record showed a traverse of the entire slide, movement along boundaries of the tissue, and motions over the center and inner portions of the tissue. Although individual global patterns of exploration could be explained post hoc in relation to individual instances of identification, the set of patterns could not be used systematically to differentiate the level of performance of the participants or the level of challenge of the slide.

The rate of motion of the microscope over the slide (to use the subjective description of the motion) was strikingly smooth and uniform across all visual search records. The rate of motion of the microscope for the five participants viewing the pancreas is illustrated in Figure 2. In general, participants started out relatively rapidly and then slowed down, with a variety of pauses along the way. The deceleration in the path was determined in part by changes in magnification as the participants moved from low to medium to high magnification. Even when the rate of motion was measured in terms of changes in visual angle, however, there tended to be an overall deceleration. The pattern of rapid movement followed by an extended deceleration was sometimes repeated for a single slide, suggesting that new areas of the slide were being explored. Patterns in the speed of movement over the slide or the smoothness of the velocity gradient did not differentiate the level of performance of the participants or the level of challenge of the slide.

![Figure 2. Cumulative distance in relation to cumulative time for the slide of the pancreas for each of the five participants. Rate of motion over the slide is indicated by the slopes of the curves. The two arrows indicate the points at which Participants 3 and 5 identified the tissue.](image)

**Dwell Time.** One of the primary measures in visual search over natural images is dwell time, those instances when a participant stops to study a particular image feature (e.g., Nodine & Kundel, 1987). Dwell time in the tracking of microscope motion could be detected by graphing the time interval be-
between indexed locations against the distance between them and looking for spikes in the graph, as illustrated in Figure 3. There were many instances of substantial dwell time in the visual search data, and the pattern of dwell time is very interesting, although not simple.

The verbal protocols, as well as the expert opinion of the histologist on the research team, suggested that certain structures in the images were diagnostic of the tissue. This would include, for example, the islets of Langerhans in the pancreas. One complexity in interpreting dwell time is illustrated by the search data for pancreas from the two participants who identified it during the verbal protocol. Participant 3 showed a classic pattern of clear dwell time on the islets of Langerhans prior to identification (see Figure 3). Participant 5, who was the highest performer of the group, did not show such a pattern. After more than two minutes of viewing, a single, relatively modest (11 sec.), dwell on an islets of Langerhans occurred during identification of the tissue. The temporal characteristics of this dwell did not differentiate this participant from others who failed to identify the tissue. On the other hand, this participant spent a very long time moving slowly through a region of the slide that had a high concentration of islets of Langerhans. This exploration time was quite large, but only when it was defined over a region of the slide rather than a single locus.

A second complexity arose from the fact that participants who did not identify the tissue sometimes dwelled on islets of Langerhans. These participants noticed the structures, considered them, and then misidentified them. Thus, dwelling on a diagnostic feature does not necessarily equate to correct identification (e.g., Manning, Ethell, & Donovan, 2004). However, if slow movement over a critical region counts for dwell time (as discussed above), then it is clear that participants who did identify pancreas spent more total dwell time on the islets of Langerhans than those who failed to identify it. The students who misidentified the islets of Langerhans did not return to them, while the students who ultimately identified them correctly returned to them on several occasions prior to final identification of the tissue.

The situation is different, however, for the epiglottis. The pancreas was challenging because the slide was created with a stain that is not common. The epiglottis was challenging because, although the individual structures of the epiglottis were familiar, the tissue as a whole had been learned only from the textbook. It had never been studied in the microscope. In the case of the epiglottis, there were participants who clearly dwelled on all of the diagnostic structures several times, discussed them clearly in the verbal protocol, but failed to remember that it was the epiglottis that would fit the description of the tissue. On the other hand, the one participant who successfully identified epiglottis during the verbal protocol did it after a single long dwell time (26 sec.) on its identifying features (elastin cartilage in conjunction with salivary glands). Thus, the pattern of dwell time associated with success in the pancreas was associated with failure in the epiglottis, and vice versa.

**Temporal Relation between Dwell Time and Identification.** A further important issue related to the viewing of diagnostic information is the temporal relation between viewing and identification. For easier slides, those that were identified correctly by everyone, there was a quick recognition of features, as indicated by the verbal protocol, and a quick response. Dwell time in these cases would only be impressive as a proportion of total elapsed time (which was often less than 10 sec.). These identifications are driven by perceptual recognition, typically recognition for one or more diagnostic structures within the slide (Pani, et al., 2005). For the more challenging slides, however, the pattern of data was quite different. For the slides that a minority of students identified successfully, all successful identifications took place after an extended period of time, with the minimum time being 34 sec. (Participant 5 viewing the epiglottis). Even a careful look at a diagnostic feature, such as the islets of Langerhans in the pancreas, typically was not connected closely in time with identification. Diagnostic features of the tissue were viewed numerous times over an extended period of time prior to identification.

**Case-study examples.** Because this was a small-sample naturalistic study of microscopy, it can be useful to view the data as a collection of case studies. Five brief descriptions of visual search in relation to performance are presented in the following. They are all from the pancreas, ranked in terms of the success in identification.

Participant 5 was the most skilled microscopist of the group. He identified the pancreas after about two minutes of viewing (125 sec.), mentioning the islets of Langerhans as one basis for his judgment. Participant 5 paused on islets of Langerhans briefly during identification (11 sec.), and for a much longer period afterward. Before identification, he spent substantial time moving around in an area that included numerous islets of Langerhans (and which included the 11 sec. dwell). Identification took place while viewing an islets of
Langerhans, but this was not the first time that the microscope was on or near an example of this structure. Indeed, the participant had explored the same critical region of the slide nearly 40 seconds earlier.

Participant 3 was a skilled microscopist who also identified the pancreas. She clearly dwelled numerous times on islets of Langerhans, and discussed them at length, but thought that they might be glomerular capsules of the kidney. She identified pancreas after a lengthy analysis of the slide in relation to known features of the pancreas and kidney. During the structured interview, she reversed her decision and decided that the slide was from a kidney.

Participant 4 did not identify pancreas during the verbal protocol. She identified it immediately in the structured interview after an islets of Langerhans was pointed out. Her visual search data, however, indicated that islets of Langerhans were at or near the center of view on numerous occasions. At one point, the microscope stopped quite near an islets for 10 seconds. Given her behavior during the structured interview, it is straightforward to infer that Participant 4 did not notice the islets of Langerhans during the verbal protocol. This can be attributed in part to their unusual look. In addition, however, she had earlier misidentified a different structure (adipose cells were interpreted as central veins of the liver) and was pursuing a hypothesis (liver) inconsistent with the presence of islets of Langerhans. Pursuit of an incorrect hypothesis easily could have contributed to her not noticing the islets. On the whole, the absence of significant dwell time on islets of Langerhans for this participant would support the idea that dwell time on diagnostic features is associated with successful performance for challenging slides.

Participant 1 noticed the islets of Langerhans, in association with a substantial dwell time, but explicitly rejected them as being islets of Langerhans. She did not offer glomerular capsules, as Participant 3 did. Rather, she searched her memory for “what has a nodule”, ultimately remembering that the thymus gland does. In the end, she offered no identification at all, saying that she was stumped. In comparison to the visual search of Participants 5 and 3, Participant 1 is remarkable for only dwelling on the islets of Langerhans once.

Participant 2 also did not identify pancreas. He scanned across islets of Langerhans on numerous occasions without remarking on them. Eventually he did pause for a long period of time on an islets of Langerhans, but misidentified it as a nerve bundle cut in cross-section. Again, this performance contrasts with the more successful students in the fact that he dwelled on islets of Langerhans only once.

Summary

The global characteristics of visual search of microscope slides were uniform across individuals and slides. Participants traversed the entire tissue, taking care to study borders of the tissue and interior regions. The speed of travel over the slide was remarkably smooth, with a tendency to begin moving relatively rapidly and then to slow down. On occasion, participants repeated this cycle for the same slide, apparently finding a new area to explore.

For the slides that were less challenging for the group of participants, identification of tissue was rapid, often requiring less than 10 seconds. Identification in these cases clearly was driven by perceptual recognition. The verbal protocols indicated that this recognition was nearly always based on noticing one or more diagnostic structures in the tissue.

For the challenging slides, even when recognition of diagnostic structures led to successful identification, identification often took an extended period of time. Viewing structures generally considered to be diagnostic of tissue was necessary but not sufficient for identification; viewing of those structures was involved in a variety of patterns of search and levels of success over varying periods of time. This variety included multiple excursions back to the same structures, long periods of exploration of regions of the tissue, misidentifications of the structures, failure to notice them, and forgetting (sometimes with eventual recall) of tissue descriptions.

In general, patterns of visual search could not be used reliably to differentiate more from less skilled participants at the level of the whole study or at the level of individual microscope slides. This appeared to be due to the importance of such higher level processes as retrieval of anatomical knowledge, anatomical interpretation of structures, and reasoning about the image that was extended in time.

Discussion

Identification of tissue in a microscope slide begins with visual perception and at least a partial recognition of structures in the tissue. In some cases, immediate recognition is all that is required (for discussion of diagnostic features in recognition, see Oliva & Schyns, 1997; Schyns, Goldstone, & Thibault, 1998). When processes of recognition are not definitive in reaching the goals of the viewer, processes of recall and hypothesis testing, including evaluation of any tentative recognitions, are joined with perception and recognition (e.g., Brooks, Norman, & Allen, 1991). In such cases, there will not be a tight spatiotemporal coupling between perception of structures and identification of the tissue. This was evident in the present study.

These findings are consistent with a number of studies of medical imaging in more diagnostic settings (e.g., Crowley et al., 2003; Nodine & Kundel, 1987). Some authors, for example, have cautioned that improving the visual display of medical images will not always improve success rates in detection, because the limiting factors for detection are sometimes cognitive rather than sensory (Manning et al., 2004; see also Kundel & Wright, 1969; Kundel & Nodine, 1983).

A histologist looking through a microscope is subjectively immersed in a microworld, and interpretation of the microscope slide nearly always feels like an instance of recognition. Subjectively, the slide seems like a sample, a point of view, as it were, that is derived from the whole tissue. Skilled histologists smoothly integrate anatomical knowledge and hypothesis testing with perception and recognition, so that successful identification of tissue often is a form of reasoning-driven recognition.
Immediate recognition and reasoning-driven recognition can be related systematically in terms of the mapping between information and target domains in a visual information system. For successful identification, the microscopist must believe that a one-to-one mapping can be maintained between structures in the slide and structures in the known target tissue. Any structure in the slide must have a counterpart in the understanding of the whole tissue, and any structure known to be in the whole tissue must either be seen in the slide or be justifiably absent. Deviations from a one-to-one mapping must be justifiable in terms of the uniqueness of the particular organism or the process of sampling slides from the whole tissue. For example, a microscopist might conclude that the slice through the tissue happened to miss a particular structure, that a particular stain does not make certain types of cell visible, or that a certain separation in the tissue is due to the mechanical process of preparing the slide.

When a one-to-one mapping from the slide to the anatomy of the tissue can be made confidently from perception and recognition of the slide, there is little need for reasoning. When the look of a structure is ambiguous, or not every structure on the slide can be mapped to knowledge, or known structures appear to be missing, reasoning becomes involved, either to explain how the mapping actually is adequate or to shift cognition toward establishing a new mapping.

Describing the use of a visual information system in terms of mappings between target and information domains suggests that it is similar in important ways to the generation and understanding of analogy. Mapping between structures is central to analogical processes also (e.g., Gentner, 1983). We believe that approaching visual information systems with questions borrowed from the study of analogy will prove to be fruitful (also see Gattis, 2004).

Acknowledgment
The authors were supported by grant 1 R01 LM008323-01A1, Histological Reasoning: Visual Cognition in Microanatomy, from the National Library of Medicine, National Institutes of Health, during completion of this work and preparation of this manuscript. The second author received support from the Grawemeyer Awards during conduct of this research.

References