A Cognitive Exploration of the “Non-Visual” Nature of Geometric Proofs

Peter W. Coppin*† Stephen A. Hockema‡
Faculty of Information
University of Toronto

Abstract

Why are Geometric Proofs (Usually) “Non-Visual”? We asked this question as a way to explore the similarities and differences between diagrams and text (visual thinking versus language thinking). Traditional text-based proofs are considered (by many to be) more rigorous than diagrams alone. In this paper we focus on human perceptual-cognitive characteristics that may encourage textual modes for proofs because of the ergonomic affordances of text relative to diagrams. We suggest that visual-spatial perception of physical objects, where an object is perceived with greater acuity through foveal vision rather than peripheral vision, is similar to attention navigating a conceptual visual-spatial structure. We suggest that attention has foveal-like and peripheral-like characteristics and that textual modes appeal to what we refer to here as foveal-focal attention, an extension of prior work in focused attention.

Keywords attention, visual thinking, proof, logic, geometry

1 Introduction

Why are geometric proofs usually “non-visual”? We asked this question as a way to explore the similarities and differences between diagrams and text (visual thinking versus language thinking [19]). We felt that the examples provided by text-based geometric proofs might be a microcosm for notation use in broader contexts, such as education, a field similarly traditionally dominated by text relative to visual-spatial information [13]. We believe that ongoing research to increase an understanding of the cognitive dimensions of visual-spatial notations relative to text could increase abilities to conceptualize, comprehend, and communicate ideas in education, public policy, and beyond by introducing principled approaches for using ergonomically appropriate notations relative to an intended communication or comprehension purpose (cf. [5, 8]).

*E-mail: petercoppin@gmail.com
†This material is based upon work supported under a National Science Foundation Graduate Research Fellowship and a University of Toronto Fellowship (during different time periods).
‡E-mail: steve.hockema@utoronto.ca

Despite a *prima facie* case that the subject matter of geometry and its underlying theories seems to be about spatial forms and relationships, geometric proofs are most often formally represented to people as text-based descriptions of geometric properties that demonstrate how a geometric relationship is necessarily true as a series of logical relationships. As Tennant [17] described: *"[The diagram] is only an heuristic to prompt certain trains of inference; . . . it is dispensable as a proof-theoretic device; indeed, . . . it has no proper place in a proof as such. For the proof is a syntactic object consisting only of sentences arranged in a finite and inspectable array."* (as quoted in [4])

For example, if block A is under block B and block C is above block B, then logic tells us that block A is below block C. Alternatively, we can easily induce that block A is below block C by observing a diagram, yet the logical text-based proof is considered more rigorous than a diagram [17]. Indeed, for generations, Euclid's *Elements* was considered to be flawed because of its reliance on diagrams. As Mumma [11] described: *"for some of Euclid's steps, the logical form of the preceding sentences is not enough to ground the steps. One must consult the diagram to understand what justifies it."* For this reason it is commonly felt that Euclid *"failed in his efforts to produce (an) exact, full explicit mathematical proof"* [11]. (We will show an example of this below in section 3.1.)

Barwise and Etchemedy [4] began to question the assumption that diagrams were less rigorous than (non-diagrammatic) proofs, bolstering their perspective by using evidence from cognitive psychology [9] that showed how maps enable problem solving more effectively in certain situations. Nonetheless, text-based "language thinking" and algebraic notations remain the dominant mode relative to diagrams throughout mathematics [6]. We asked: *"could there be ergonomic properties afforded by text relative to diagrams that encourages textual modes for proofs?"*

An inversion of our question might be phrased as follows: *what do the representational modes of sequential symbolic proofs relative to diagrams reflect about human cognition?* Though others [1, 6, 12] have explored related questions and have described the dominance of text over visual-spatial representations through historical explanations, we focus our attention on human perceptual-cognitive characteristics that may encourage proofs (and similar materials) to evolve towards text-based modes relative to visual-spatial modes because of the ergonomic affordances of text relative to diagrams.[1]

1.1 Related Work

Relative to well developed studies of language in linguistics and related fields, studies of formal visual representations are sporadic and fall across less connected fields [1, 4, 13]. Very little prior work was found that addressed our specific question (especially from a perceptual-cognitive perspective), however, some work with results that can be adapted to explore our question follow:

[1] We also narrow our focus to the presentation of proofs in their final form, as opposed to also considering the discovery and construction process of proving geometric theorems. While the two are obvious related, we believe there is enough to say about the former here that can stand on its own without considering the latter.

Coming from an information processing (cognitive psychology) perspective, Larkin and Simon [10] sought to explore the differences between information as diagrams versus sentences, concluding that sentences embody the characteristic of being indexed on a list, with each element "adjacent" only to the next element in the list. In contrast, diagrams are indexed by location on a plane, many elements may share the same location, and each element may be adjacent to any number of other elements; in this way, Larkin and Simon propose that diagrams may be more useful than sentences for solving certain kinds of problems because they can support more efficient computational processes. (Larkin and Simon include human neurological processes when they use the word "computational".) They also noted that this efficiency depends on the design of the diagram and the ability of the user to interpret the diagram.

Approaching the issue as mathematicians and logicians, Barwise and Etchemendy [4] begin their work by noting that in the field of mathematics and logic, diagrams are not considered valid parts of a proof, and are present only as a heuristic aid (Barwise and Etchemendy [4] citing Tennant [17]). A major purpose of [4] is to overturn this thinking, bolstering their case by citing cognitive psychologist Kosslyn [9], who used maps to justify visual presentations as valid problem solving tools, also making the point that sentences or visual representations offer advantages or disadvantages based on the purpose of the task at hand.

Barwise and Etchemendy conclude (like Larkin and Simon in [10]) by describing advantages offered by diagrams that are not offered by sentences, and by doing so offer a suite of differences between sentences and diagrams that extend [10]. For example, with diagrams relationships are often implicit, whereas with sentences, even the most trivial consequences must be inferred explicitly (as demonstrated in the introductory example). Additionally, they point out that a picture or diagram can support "countless facts" (by this they mean that a plurality of sentences can be constructed from a diagram).

More recently, Mumma approached the question of why geoemtric proofs are language-based by examining the so-called "flaws" in Euclid's proofs where he made use of diagrams and then seeking to provide a rigorous diagramatic foundation for these proofs [12]. In so doing, Mumma proposed three interrelated factors why Euclid's reliance on diagrams in his proofs is regarded as non-rigorous These were what he referred to as:

- the generality problem – proofs are meant to be more general than the *particular* instance in a diagram but how should we generalize from a particular diagram to a more general case?
- the modern mathematical understanding of continuity – diagrams may lead to simplistic and invalid assumptions about the continuity of lines, e.g. with respect to the existence of intersection points
- the modern axiomatic method, which requires that *all* axioms and deductive steps be explicitly specified, and raises suspicion about the assumptions embedded in diagrams.

Thus, like Barwise and Etchemendy, Mumma locates the answer in the subject matter and norms of the field and then attempts to argue that the above three problems can be overcome in a more carefully specified hybrid system. We will argue below that there is also another component to the answer based on basic facts about human cognition.

Along these lines, Shimojima and Katigiri [16] sought to support Barwise and Etchemendy [4] by gaining empirical evidence to show how diagrams reduce inferential load by drawing on newer discoveries in cognitive science such as Ballard et al.'s [2] theory of "deictic indices." Deictic indices are mental pointers to particular objects in external space. The theory suggests that through an attentional mechanism, people can maintain a small pool of such indices at once and can easily direct focal (mental) attention or gaze to any of these indices. In Ballard's et al.'s [2] words *"pointing movements are used to bind objects in the world to cognitive programs."*

Shimojima and Katigiri [16] describe how these indices can be used to keep track of (mental) "non-physical drawings" they may construct while doing inferences about diagrams. They suggest that reasoners can then navigate their attention through these "non-physical" (mental) drawings during reasoning tasks and that these non-physical drawings are assisted by real drawings (diagrams) and that this assistance reduces "inferential load" during reasoning tasks.

Thus, although none of these works specifically address our question about why proofs are "non-visual," we gain important insight regarding the differences between text and diagrams that can guide our inquiry. The thread running through each prior work listed above is that text/language/prose guides attention in ways that are different from visual-spatial representations. The next section will attempt to demonstrate this more explicitly.

1.2 Integrating and Extrapolating from Prior Research: Language Appears to Guide Attention through Visual-Spatial Structures

From this loose collection of interrelated work, we can extrapolate some general principles regarding the cognitive dimensions of illustrations relative to text.

Larkin and Simon suggested in [10] that a cognitive dimension of sentences is their list-like structure, in that each item on the list is only adjacent to the item before or after it on the list. In contrast, items in a diagram are adjacent to many items on a list. This view is synergistic with Barwise and Etchemendy, who suggested that a picture or diagram can support "countless facts" (by this they mean that a plurality of sentences can be constructed from a diagram). In other words, many sentences could be created by linking together elements in a diagram into a sentence (list-like structure).

In this way, we extrapolate that a list-like structure (as suggested by Larkin and Simon) can be inferred or induced from a diagram. Each sentence inferred from a diagram is like a path that guides attention through visual-spatial relationships in a diagram. This extrapolation is demonstrated by Shimojima and Katigiri's eye tracking study in [16] that showed how reasoners mentally guide their attention through a "non-physical drawing." They suggest that actual drawings thus support non-physical (mental) drawings, thus reducing inferential load. To summarize, it appears that sentences guide attentional paths through both physical and non-physical (mental) visual-spatial structures. Further, it appears that Shimojima and Katigiri demonstrated that rational language/propositional logic guides attention and motor movements (through eye fixations) through non-physical visual-spatial representations.

1.3 Why does "Language Thinking feel More Precise than Visual Thinking"?

At this point in the paper, we are almost ready to suggest a contributing reason to the answer of why "geometric proofs are text-based." We have suggested that sentences guide attention through visual-spatial structures. However, a question remains: *Why does navigating attention through a visual spatial structure guided by rational language (propositional logic) feel more "rigorous" than experiencing it as a diagram?* As described above, in [12] Mumma proposed a 3-part answer as to why the diagrams are considered less rigorous by the field in general. Here we narrow the question to focus on cognitive dimensions of individual mathematicians.

We propose that detailed scrutiny of visual-spatial structures (and perhaps concepts in general) requires what we will refer to as "higher resolution" foveal-like attention, even if those visual-spatial structures are conceptual (not directly sensed). We suggest that visual-spatial perception in the physical world, where an object is perceived with greater acuity through foveal vision rather than peripheral vision, is similar to attention navigating a conceptual visual-spatial structure. *We suggest that attention has foveal-like and peripheral-like characteristics.* Linkages to traditional (and synergistic) ideas of attention (e.g., [18]) will be described later in this paper.

To explain how navigating non-physical (mental) drawings may have dimensions that mimic visual-spatial perception of the external world through foveal and peripheral vision, we can turn to Barsalou's [3] theory of perceptual symbol systems where concepts are based on inherently modal neural patterns rooted in direct sensory experience. As Barsalou [3] describes:

> During perceptual experience, association areas in the brain capture bottom-up patterns of activation in sensory-motor areas.
>
> In a top-down manner, association areas partially reactivate sensory-motor areas to implement perceptual symbols.
>
> The storage and reactivation of perceptual symbols operates at the level of perceptual components not at the level of holistic perceptual experiences [3].

In other words, an experience of a geometric visual-spatial structure is inherently modal in that, if experienced through the eyes and visual cortex (for example), the memory of that experience would reflect the experiential mode (i.e. visual versus auditory experience). This means that concepts that emerge from the neurological patterns created from sensory (modal) experience reflect the characteristics of the experiential mode. This means that an experience of a visual-spatial structure that emerges as a concept uses much of the same neurological machinery used to perceive (experience) the visual-spatial structure.

This relationship between foveal attention and perceptual symbols is the basis for our theory detailed in the next section.

2 Theory

Why are geometric proofs (usually) "non-visual?" We propose that perceptual architectures associated with foveal (sharper, center view, but narrower field of view [FOV]), and peripheral (outside of the center view, less sharp, but a wider FOV) vision found in the human eye, in V1, and the rest of the visual cortex extend into the "deepest levels" of human cognition and are reflected both in conceptual structures and the architecture of attention that "probes" those conceptual structures. In this paper, foveal attention is analogous to (and parallels) foveal vision. Likewise, "peripheral perceptual-cognitive attention" (shortened to peripheral attention [PA] for the rest of this paper) is analogous to peripheral vision.

We suggest that foveal attention may by synonymous with "focused attention" as proposed by Treisman [18], who suggested *"attention must be directed serially to each stimulus in a display whenever conjunctions of more than one separable feature are needed to characterize or distinguish the possible objects presented."* By separable feature, she means primitives such as basic shapes, objects, and colors prior to integration into a conceptual "whole." Furthermore, Treisman uses a metaphor that easily maps to our description of foveal attention:

> Visual attention, like a spotlight or zoom lens, can be used over a small area with high resolution or spread over a wider area with some loss of detail. (Treisman [18] citing Eriksen and Hoffman [7])

We can extend the analogy in the present context to suggest that attention can either be narrowed to focus on a single feature, when we need to see what other features are present and form an object, or distributed over a whole group of items which share a relevant feature [18].

This *"narrowing of the spotlight"* is synonymous with what we mean by foveal attention. Relative to peripheral attention, we suggest that foveal attention is more precise and can detect more detail, paralleling Treisman's spotlight/zoom lens metaphor. Similar to how the eye must explore areas broader than the narrow FOV of foveal vision via a sequence of saccades, we suggest that focal attention must also sequentially walk through mental visual-spatial structures. However, perhaps differently from low-level saccades, we suggest that *language and language-thinking guides* attention through such structures in order to build more precise holistic ideas.

By building on Barsalou's notion that concepts arise from neural patterns that are rooted in modal experiences, and our own speculation (extrapolating from Triesman) that foveal-focal attention may have a limited "FOV," several explanations for why proofs are usually text-based and propositional are proposed:

Why proofs are often sequential: Like foveal vision that must saccade to different parts of a visual-spatial structure, we suggest that foveal-focal attention must also "saccade" to different parts of a conceptual visual-spatial structure due to foveal-focal attention's narrower FOV (rather than experiencing / attending to the whole structure at once).[2]

[2]It should be emphasized that we are using the notion of saccade here metaphorically; sequential movement of attention to various parts of a structure will probably bear no resemblence to the way an eye actually

Why a diagram usually cannot constitute a convincing "holistic" proof: The need for symbols to fall within the narrow FOV of foveal-focal attention means that diagrams, and the spatial relationships they embody, usually cannot be taken in (i.e. attended to) all at once. So they should instead be processed in a way that allows linkages between earlier perceptual memories and later percepts.

Why text is effective for proofs: External symbolic representations such as text are designed such that each symbol can sequentially fall within the narrow FOV of foveal vision [15] and therefore, foveal attention and foveal-focal attention.

Why propositional logic is used for proofs: Propositional structures in proofs may provide symbolic "short-cuts," serving as stand-ins for visual-spatial relationships that cannot all simultaneously be in the limited FOV of foveal attention. For example: The statement *"if C is below B"* references a perceptual symbol constructed from a previously considered image, and the statement *"if B is below A"* also references a perceptual symbol constructed from a previously considered image. The statement *"therefore C is below A"* references the two previous symbols in order to support construction of a new (mental) image that can serve as the basis for a new perceptual symbol (and that can be used in future propositional statements).

To summarize our theory; we suggest that a visual-spatial structure, such as a geometric structure (irregardless of whether it is presented as a diagrammatic representation) is often beyond the "FOV" of foveal-focal attention. The purpose of sequential symbolic representations such as text, organized as propositional statements, is to guide foveal-focal attention through a sequence of patterns in order to create perceptual symbols that are amenable to analytical neurological machinery.

Hence, in addition to being due to the norms of the field of mathematics, as well as many other social and mathematically technical reasons that have been proposed, we argue the answer to our initial question is *also* related to basic facts about how human cognitition works.

3 Thought Experiment

A thought experiment may help clarify the role of the less-diagrammatic notation style of propositional logic used in a geometric proof by imagining what a notation style would look like that was designed to guide narrow FOV foveal-focal attentional processes.

First, a notation style where symbols would fit the narrow FOV of foveal vision would seem to be appropriate, although it is perhaps not the only style that could work. For example, with a better understanding of processes linking perception and attention, we might be able to use more "bandwidth" in parallel by providing just the right cues to guide attention through a diagram (and hence, reduce or even avoid the need for "symbols" altogether). Yet text appears to be an example of a notation system naturally suited to the ergonomics of foveal vision [14].

---

saccades, such as to its ballistic nature for example. Further, the "sequence" implied by the word "sequential" here is not necessarily imply a *particular* ordering, especially not one that might correspond with the actual sequence of eye saccades. We assume that many cognitive and pragmatic factors play into determining in what order structures must be attended.

Second, geometric structures expressed through the notation system would need to be "serialized" as chunks/strings since more complex visual-spatial structures (i.e. diagrams associated with non-trivial geometric proofs) will presumably require a FOV that fall outside of the foveal attentional units of a notation system ergonomically designed for foveal vision [14].

Third, the notation style would need to deliver those serialized chunks/strings in ways that would direct foveal attention in specifically ordered trajectories and patterns to build a network of foveal-focal perceptual symbols in order to construct and mentally navigate a visual-spatial conceptual structure that falls outside of the "FOV" of foveal-focal attention.

This is because sequential patterns of foveal-focal "saccades" support the creation of perceptual symbols that can be referenced in later attentional "saccades." This would be a hierarchical structure of previously attended foveal-focal mental images (perceptual symbols) where latter parts of the conceptual structure reference previously attended foveal-focal mental images.

In other words, a notation system that was custom designed to guide foveal-focal attention through a visual-spatial conceptual structure would resemble the ergonomics of sequential symbolic (i.e. text based) proofs consisting of serialized sequential symbols (i.e. descriptions), and support the embedding of perceptual symbols, constructed from previously experienced foveal mental images, into other mental images and perceptual symbols (i.e. propositional logic).

3.1 Example

A more specific example may reveal the ergonomic characteristics and constraints described above. Proposition 35 from Book I of Euclid's Elements is a classic example that we suggest demonstrates the way that text appears to focus foveal-focal attention. It is also a useful example in that it is a hybrid proof, as will be described below, relying on both text and diagram.

Proposition 35 is that parallelograms that are on the same base and in the same parallels equal one another. Euclid's proof proceeds as follows:

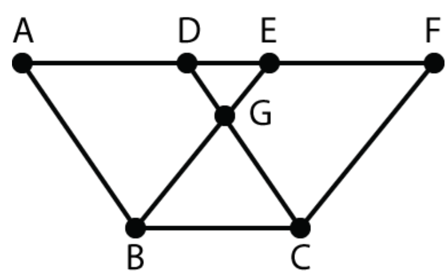

Figure 1: Diagram used in proof of Proposition 35

Proof.

(i) Let ABCD, EBCF be parallelograms on the same base BC and in the same parallels AF, BC.

(ii) Since ABCD is parallelogram, AD equals BC (Proposition 34). Similarly, EF equals BC.

(iii) Thus, AD equals EF. (Common Notion 1)

(iv) Equals added to equals are equal, so AE equals DF. (Common Notion 2)

(v) Again, since ABCD is a parallelogram, AB equals DC (Proposition 34) and angle EAB equals angle FDC (Proposition 29).

(vi) By side angle side congruence, triangle EAB equals triangle FDC (Proposition 4). Subtracting triangle EDG from both, we have that the trapezium ABGD equals the trapezium EGCF (Common Notion 3).

(vii) Adding triangle GBC to both, we have that ABCD equals EBCF (Common Notion 2);

□

Note how the text references aspects of the visual-spatial concept in "chunks," revealing the visual-spatial concepts serially, possibly ergonomically optimized for foveal-focal attentional processes. For example, the first line introduces the symbol "$ABCD$", which, in the presence of the figure, directs attention sequentially through the verticies A-B-C-D and then to the parallelogram as a whole, separable from the rest of the figure. However, the figure is not necessary for this step, for ABCD could also serve the role of a simple "word" (term) in the logical proof without actually *referring* to the geometric figure at all. Further, for everything specified early on in the proof—symbols and relationships—each line can be derived from the previous without making reference to the figure at all. Although the figure can still play a helpful illustrative role, it does not play a role in sanctioning particular inference steps. Indeed, as Mumma describes in [12], this is the case all the way up to step iv.[3]

For our purposes, the key point here is that in its linguistic form, devoid of the figure, the above proof is amenable to cognitive processes, and this is revealed in aspects of its design. The structure of the language guides the reader through the proof in "bite-sized" chunks – both in the the breakdown as to what constitutes an individual *step*, and the number of symbols involved in each step – that can be linked together and composed to gradually build up to the conclusion.

However, things get more complicated in steps iv through vii, in part, because the author of the proof (Euclid via translation) seems to be making the assumption that we will be using the figure to provide interpretation for the text statements, and thus both allow the engagement of our natural perceptual-cognitive chunking abilities and relieve

[3] Step iv contains a so-called "flaw" in that in order to determine what the "equals" are that have been added to AD and EF, one needs to realize that DE is a common segment shared by both AE and DF. The step relies on an understanding that DE has been "added" to both AD and EF and it is reflexively self-equal. However, while it is true, it was nowhere stated in steps i-iii that DE is a shared segment, i.e., that D lies on segment AE and that E lies on segment DF. In the above proof, this is knowledge that must be gleaned from the figure.

some of our memory burdens by using the figure as external memory. The proof can be rewritten more "rigorously" to eliminate the need to refer to the figure and to rely on text alone, but this presumably will also involve a simplification of steps iv through vii by breaking them down into more, explicit steps that each demand less cognitive work for the reader.

But could we go the other way? Could we prove the same thing by relying even *more* on diagrams and using much less text? If so, what would it take? In the next subsections we will attempt to do so in order to illustrate our thoery as to how deductive proofs require the explicit sequential guidance of attention through symbols, and demonstarte the difficulties that arise when we attempt to employ diagrams for this purpose.

3.2 Towards a more Visual Version

In the last subsection, we saw an example of a geometry proof that, while not entirely avoiding reliance on a diagram, was primarily text-based. In light of our argument above (that this is partly because text is better suited to guide foveal-focal attention sequentially through the appropriate perceptual symbols), here we seek to illustrate this point by describing the results of our attempt at a visual proof of Proposition 35. How could this be achieved without the aid of text? Or in other words, how can one draw attention without using text?

The most straightforward approach here might be to translate each line of text from the proof into a diagram seeking to achieve the same thing. In so doing, one might seek to direct attention to different parts or aspects of the diagram using graphical techniques such as highlighting with color, luminance (value), shading, or adding arrows (to name a few techniques). So, for example, to translate the step that establishes that AD equals EF (step iii), we might highlight both of those seqments in red and add some sort of connection beteween the two to denote equivalence. Further, we would need to somehow add the *justification* for this, somehow refering to Common Notion 1.

After we attempted to do this in a systematic way, several things become apparent. First, the guidance of attention within a diagram seemed harder to control relative to text, and even somewhat arbitrary (somewhat like "black magic") to a practitioner primarily trained to express ideas through text. The guidance of attention with a diagram seems to be something that requires a solid scientific understanding of how diagrams are visually processed and/or considerable artistic sophistication and skill. Additionally, relative to our experience with text, it was harder to be precise – to the "right" degree of precision – in a diagram. For example, how can we call attention to just the vertices, say, of a parallelogram, without also calling attention to the parallelogram itself via gestalt principles of perception? Similarly, as Mumma described in [12], it was harder to make general points when dealing with specific diagrams, and there seemed to be the potential to be misled by superfluous details in the diagrams.

So, a straightforward translation of a text-based proof into a visual one seemed to presents several difficulties. Yet there may be other approaches to visual proof that are not biased by the texual starting point, perhaps natively taking advantage of unique affordances of diagrams. We speculate about this in the next subsection.

An open question is how text and diagrams can best be used *together* to complement each other's strengths in a a hybrid approach. As noted, Euclid's proof was not entirely text based; the diagram was required for it to go through. It is instructive to consider the role the diagram played in this case. In our theory, a primary function of the diagram is to provide the basis for perceptual symbols to which the text can then refer and "navigate", i.e. guide attention through. We unpack this more in the next paragraph.

As an example of how the diagram works together with the text, consider step vi of the proof:

- By side angle side congruence, triangle EAB equals triangle FDC (Proposition 4). Subtracting triangle EDG from both, we have that the trapezium ABGD equals the trapezium EGCF (Common Notion 3).

First, note that the diagram serves to confirm the applicability of Common Notion 3 in a way similar to how the diagram was used to justify step iv. While the text here could be augmented to ostensibly stand on its own, it is still compelling to look at the trapezia ABGD and EGCF in the figure to at least *confirm that the text was properly understood*. (A reader can "double-check" that they did the subtraction in the right way by checking that the trapezia they obtained were indeed describable as ABGD and EGCF, that is, by unpacking the symbols—e.g., ABGD to A-B-G-D—to confirm the correct vertices were involved.) Next, note that the first sentence not only establishes an equivalence relationship, but also serves to focus attention on the two relevant triangles, EAB and FDC, as quasi-independent parts of the diagram, which then makes them available to mental manipulation (and thus, subsequent use in the text). The operation of subtracting triangle EDG from both EAB and FDC requires manipulating it as a "non-rivalrous" (reusable) symbolic unit. Further, as was previously discussed, [16] has provided evidence of how the diagram is used when mentally reasoning about such operations, with deictic indices being allocated to diagrammatic symbols from a limited pool that constrains the number of symbols to which we can simultaneouly attend. Thus, the mental subtraction requires manipulating holistic triangles as symbols – each as *one* thing – as opposed to a loosely coupled collections of their component vertices and edge segments which would quickly exhaust attentional resources. The diagram helps to reify these component parts into unitary symbolic wholes thanks to gestalt principles of perceptual grouping.

So one of the primary roles of the diagram seems to be to serve as the basis for perceptual symbols that can play a role in *compositional* manipulations, analogous to how words are composed via grammar in language. That is, the diagram, when it is being used effectively in a hybrid text-diagram proof system, still seems to be guiding attention in ways similar to text (sybolically and sequentially). Yet, diagrammatic symbols may afford more than textual symbols in that they can be unpacked and "inspected" when necesary, as the trapezia were when verifying the subtraction. This warrants more research into the affordances of diagrams relative to text in the context of proofs, as opposed to in more general contexts.

3.3 Other Hybrid Systems

In the last subection, we started by considering how we might constuct a visual proof from a primarily text-based one. But starting with text might bias us away from a proof strategy that might better take advantage of the affordances of diagrams. What if we were to start from scratch?

Ware points out in [19] that one of the more effective techniques used in HCI and Information Visualization to draw attention is motion. This implies that animations might be a useful tool in geometric proofs. However, since traditional (paper oriented) document formats do not ordinarily support dynamic animations, for the purposes of this paper a comic strip-like story-board will depict such a possible interaction by way of example. Figure 2 shows a possible story-board (or comic strip), corresponding to a potential animation sequence, for the proof of Proposition 35. This is just a tentative sketch to demonstrate the idea; it is has not been refined nor tested with mathematicians, nor has research been done on the proper balance between textual and non-textual elements in such a medium.

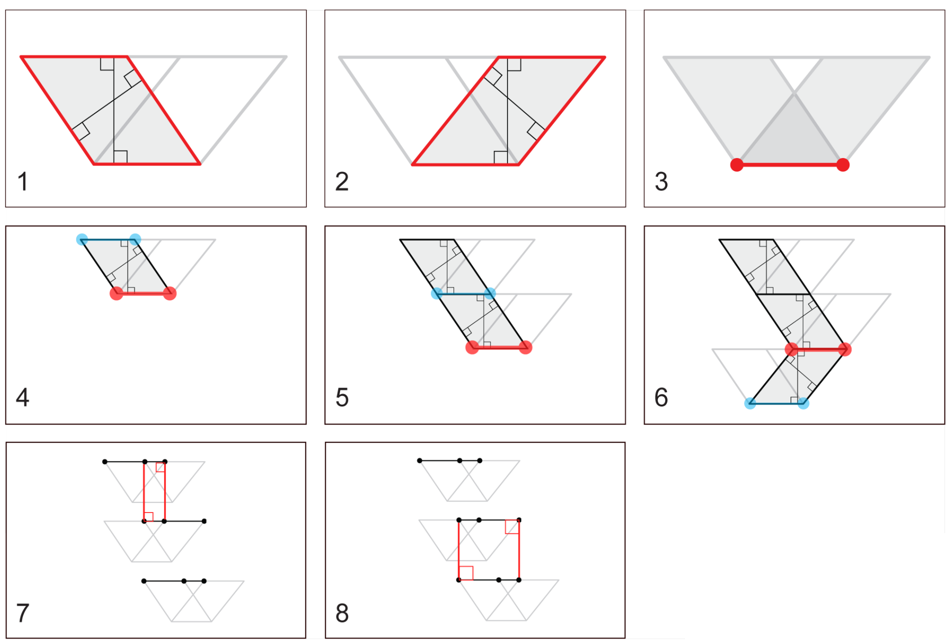

Figure 2: A possible "storyboard" (or "comics") hybrid proof

Storyboards (or comics), in their own right, also present a potentially useful hybrid approach for directing attention during a visual proof. In our theory, the storyboard has a useful property for proofs: the boxes help guide attention, much like the proof steps and symbols in a text-based proof. Further, the domain of comics provides many techniques to further refine the guidance of attention (such as cutaway shots) and help with the problems described above that arise when trying to use diagrams in proofs. For example, there are standard comics techniques to "zoom" in and out, both temporally and spatially, across panels, which may be useful in addressing the problems with precision.

One final point of interest here that emerged while creating this example is that this form of reasoning "felt" more inductive than deductive. Our research seeks to increase the understanding of how different media types might be better suited for diferent sorts of inference. Understanding ergonomic factors that enable a principled approach to attention guidance is an area of active future research for us.

4 Conclusion and Implications

We feel that demonstrating how propositional logic guides foveal-focal attention through non-physical drawings has implications far beyond notation techniques for geometric proofs; the issues surrounding the text-based nature of geometric proofs are a microcosm for issues facing other materials, such as textbooks, because many educational concepts that are also visual-spatial in nature use illustrations in a supporting role [13]. We feel that understanding how notation styles direct attention can enable the creation of materials suited for different purposes and for different kinds of learners.

For example, the effectiveness of the Barwise and Etchemedy vision of using visual materials for logical problem solving would be a function of how well the materials support the guiding of foveal-focal attention in specific patterns. Indeed, many of their ideas were rooted in classroom multimedia experiences geared for the teaching of logic. Using today's technology of the Web and multimedia (i.e. rollovers, mouse events, etc.), the sequential symbolic characteristics of comics, and beyond, many options exist to direct foveal-focal attention using visual materials in ways that may serve the attentional purposes provided by text-based proofs and text heavy materials. We feel that such techniques could be extended for other information presentations as well, finding uses in education, public policy, business, engineering, and beyond.

However, we suggest here that prose may direct attention in ways that may not be possible through visual-spatial notations such as illustrations and diagrams. Because text and other sequential symbolic "language" systems inherently require learning, they may by their very nature "bypass" or "overcome" low level pattern detection neurological machinery in order to trigger "top down" processes that amplify attentional neurological machinery in order to focus attention in specific ways on visual-spatial diagrams or conceptual structures. A more intuitive notation system that relied more on "natural," "hard-wired." or "gestalt-like" abilities might strengthen bottom-up patterns instead, at the expense of the intended abstract reasoning encouraged by such materials. Instilling neural patterns that overcome lower level neurological machinery may be the very nature of some aspects of education. At the same time, visual thinking has been proposed as a key part of "transformational thinking," and many great discoveries occurred through insights with a strong visual-spatial component [13] and by individuals who do not learn well through textual modes [15].

Thus, many open questions remain about the ergonomic properties of textual and diagrammatic modes.

Importantly, a better, more rigorous understanding of the explicit and implicit guidance of attention is a necessary step, and a direction for our current and future research.

References

[1] Arnheim, R., "Visual thinking," University of California Press, 2004.

[2] Ballard, D. H., M. M. Hayhoe, P. K. Pook and R. P. N. Rao, *Deictic codes for the embodiment of cognition*, BEHAVIORAL AND BRAIN SCIENCES **20** (1995), pp. 723—767.
URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.49.3813

[3] Barsalou, L. W., *Perceptions of perceptual symbols*, Behavioral and Brain Sciences **22** (1999), pp. 637–660.
URL http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=31859

[4] Barwise, J. and J. Etchemendy, "Visual information and valid reasoning," Mathematical Association of America, 1991 pp. 9–24.
URL http://portal.acm.org/citation.cfm?id=115667

[5] Blackwell, A. F., *Ten years of cognitive dimensions in visual languages and computing: Guest editor's introduction to special issue*, Journal of Visual Languages & Computing **17** (2006), pp. 285–287.
URL http://www.sciencedirect.com/science/article/B6WMM-4K5JBS2-1/2/d76e6d2adf8fbe31d1dcecf8ddfb7fca

[6] Brown, J. R., "Philosophy of mathematics," Routledge, 1999, 215 pp.

[7] Eriksen, C. W., *Temporal and spatial characteristics of selective encoding*, Perception and Psychophysics (1980).

[8] Green, T. and M. Petre, *Usability analysis of visual programming environments: A 'cognitive dimensions' framework*, Journal of Visual Languages and Computing **7** (1996), pp. 131–174.
URL http://www.scopus.com/inward/record.url?eid=2-s2.0-0030167097&partnerID=40

[9] Kosslyn, S. M., "Image and mind," Harvard Univ Pr, 1980.

[10] Larkin, J. H. and H. A. Simon, *Why a diagram is (sometimes) worth ten thousand words*, Cognitive science **11** (1987), pp. 65–100.

[11] Mumma, J., "Ensuring Generality in Euclids Diagrammatic Arguments," 2008 pp. 222–235.
URL http://dx.doi.org/10.1007/978-3-540-87730-1_21

[12] Mumma, J., *Proofs, pictures, and Euclid*, Synthese (2009, in press).
URL http://www.contrib.andrew.cmu.edu/~jmumma/list.html

[13] Ramadas, J., *Visual and spatial modes in science learning*, International Journal of Science Education **31** (2009), pp. 301–318.

[14] Rayner, K. and J. H. Bertera, *Reading without a fovea*, Science **206** (1979), pp. 468–469, ArticleType: primary_article / Full publication date: Oct. 26, 1979 / Copyright 1979 American Association for the Advancement of Science.
URL http://www.jstor.org/stable/1749329

[15] Schneps, M. H., L. T. Rose and K. W. Fischer, *Visual learning and the brain: Implications for dyslexia*, Mind, Brain, and Education **1** (2007), pp. 128–139.
URL http://dx.doi.org/10.1111/j.1751-228X.2007.00013.x

[16] Shimojima, A. and Y. Katagiri, "An Eye-Tracking Study of Exploitations of Spatial Constraints in Diagrammatic Reasoning," 2008 pp. 74–88.
URL http://dx.doi.org/10.1007/978-3-540-87730-1_10

[17] Tennant, N., *The withering away of formal semantics?*, Mind & Language **1** (1986), pp. 302–318.
URL http://dx.doi.org/10.1111/j.1468-0017.1986.tb00328.x

[18] Treisman, A. M. and G. Gelade, *A feature-integration theory of attention*, Cognitive psychology **12** (1980), pp. 97–136.

[19] Ware, C., "Visual Thinking: For Design," Morgan Kaufmann, 2008.