Chapter 2
Theories and Models in Graph Comprehension

Amy Rae Fox

Abstract Graph comprehension is the act of deriving meaning from graphs, an activity grounded in visuospatial reasoning that develops through a combination of instruction and practice. What we know about the mechanisms of graph comprehension stems from interleaving lines of inquiry in statistics, computer science, education, and psychology dating back to the 1980s. In this integrative review, I describe how models of graph comprehension evolved in response to developments in cognitive theory, offering a critical commentary on how foundational theories build upon each other, extending rather than replacing theoretical claims at different levels of analysis. I illuminate the landscape of contemporary research, before concluding with an argument for the role of visualization psychology in supporting theoretical integration across disciplinary boundaries.

2.1 Introduction

There is a conceptual paradox at the center of research on graph comprehension. The reason we employ graphical displays is that—in relation to text or tables of numbers—they seem effortless. Deriving meaning from a graph is described as "seeing" the information, equated with the facile fluency of perception. But this effortless access obscures a murky, error-ridden reality. Correctly reading a graph is much harder than we think. After 40 years of empirical research and theory building, we have learned that our ability to interpret a graph is influenced by a multitude of interacting factors affecting the display, the individual, and the situation.

In this chapter I offer a historical commentary on the development of graph comprehension research. I describe how theory in graph comprehension arose out of empirical research across disciplines and propose a role for visualization psychology in facilitating theoretical integration. This chapter will be useful for visualization

A. R. Fox (✉)
University of California San Diego, La Jolla, CA, USA

D. Albers Szafir et al. (eds.), *Visualization Psychology*,
https://doi.org/10.1007/978-3-031-34738-2_2

researchers looking to navigate the interdisciplinary milieu of graph comprehension, and students of behavioral and social sciences seeking a primer on this essential area of research.

2.1.1 What Kind of Graph Is a Graph?

The term external representation is used to indicate things in the world—subject to experience by human perception—that purposefully refer to other things. External representations can be constructed for any sensory modality and medium, though the visualization researcher is particularly interested in those employing *graphics* that can be *seen* on some *surface*. The text on this page is a visual external representation, with the letters of the alphabet functioning as symbols referring to sounds that you have learned to assemble into words from which you construct a certain understanding of what I intend to communicate. Similarly, a photograph is a visual external representation, referring via resemblance and analogy to the scene it depicts. A rich spectrum lies between these symbolic texts (describing the world) and analogous pictures (depicting the world). The design and interpretation of external representations belongs to the interdisciplinary realm of *semiotics*: the study of meaning-making (see Chap. 9). The focus of this chapter is a subset of external representations colloquially referred to as graphs (from the Greek graphē "writing, drawing"), charts, or plots: diagrams that convey relationships between sets of information via visual-spatial variables in a coordinate system (see [6, 62]). These are not to be confused with another set of representations referred to as "graphs": collections of edges that join pairs of vertices (à la "graph theory; node-link diagrams). Graphs are typically distinguished from *maps* which use scaled space to represent geographic relations. Both kinds of graphs belong to the larger class of diagrams: external representations that use space and simplified visual forms to convey relationships between their referents. Importantly, the use of these terms in empirical research is as fluid as the taxonomies that seek to structure them (see [25, 34, 53]). While the models and theories of comprehension reviewed in this chapter reference graphs specifically, it is reasonable to infer that the general purpose mechanisms of graph comprehension may also apply to the larger class of external representations.

2.2 An Abridged History of Theory in Graph Comprehension

As is often the case with interdisciplinary research, the study of graph comprehension arose from the needs of practice, rather than an invariable march of basic theory. The pioneering graphical inventions of Playfair, Minard, and Galton in the "golden age" of visualization were only made mainstream through inclusion in textbooks (e.g., [11]) and standards reports (e.g., [2]), through championing in professional texts (e.g., [78]) and essays in scholarly journals (e.g., [21, 45]). As the use of such "statistical graphics" spread, guidelines were needed for when and how they could be used to communicate effectively: a call for science to explain the art.

The earliest empirical investigations were published in statistics [22, 24, 82] and consisted of discrete comparisons between bar and pie charts, testing a viewer's performance in judging proportions. Concurrent work in educational psychology [85] tested secondary school students on their memory of facts learned from bar and line charts, pictographs, and tables. Studies of this kind were framed as empirical tests of guidelines offered in textbooks like that of Brinton [11] but were subject to methodological critiques of construct validity. In contextualizing their results, the authors tended to frame outcomes as properties of the representations themselves: *a bar chart is more effective at [X] than a pie chart*, while contemporary scholars would identify performance as arising from the *interaction between* the individual and representation. This subtle but important difference betrays that the focus of early efforts was on understanding the nature of the representations and their properties.

These types of point-to-point and application-grounded studies would continue for decades, in the absence of frameworks, theories, or models to guide causal or mechanistic investigation. The work was published in statistics, educational psychology, computer graphics, and the burgeoning field of HCI. This would be the case until three developments in the 1980s paved the way for a more coherent, additive body of research to unfold. First, Jaques Bertin's seminal work *A Semiology of Graphics* was translated from French to English by WJ Berg (under the supervision of Howard Wainer) in 1983. Bertin was the first to offer a concise language and structure for decomposing the questions we might ask about what a graphic is and how it might work. Second, post-cognitive revolution, substantial theories connecting visual perception to higher order cognition had been published in cognitive science—notably Marr [52] and Ullman [80]. Finally, the "mental imagery debate" was well underway, which saw leading cognitive scientists debating the nature of mental representation. This focuses on representation spurred interest in *external* representation and in particular how graphics are leveraged for problem solving and communication (e.g., [46]).

In the sections that follow, I describe a progression of theoretical development that has shaped the trajectory of graph comprehension research—work that directly addresses the fundamental question: *how are humans able to read graphs?* Our focus will be on the elaboration of general *theory*—accounts of the mechanisms through which our interaction with statistical graphics unfold—rather than individual empirical contributions. We will see examples of theory reasoned from personal experience, appeal to logic, and theory reasoned from experimental evidence. A substantial body of theory has been developed in information visualization and education that addresses the application of visualization and diagrammatic representations more broadly, though (cognitive) theory in graph comprehension can be construed as its foundation, the backbone of investigations exploring specific phenomena observed within those interactions. Questions like *what kind of graph is most effective for decision-making?* or *how can we help learners correctly interpret a graph?* rely on general purpose mechanisms of graph comprehension, just as questions of effective linguistic communication rely on the underlying mechanisms of reading and speech comprehension. Figure 2.1 summarizes early theoretical contributions, including a number of general taxonomic grammars and computational efforts that are not discussed in further detail.

The reader will notice that our understanding of graph comprehension did *not* progress via development of *competing* models and theories. Rather, research has unfolded as a progressive elaboration of a vast problem space, with works that shed light on disparate aspects or tasks, and others that expand on prior theory at different levels of detail, iterating rather than refuting. Half of the challenge is deciding what questions need to be answered, and here lies the power and difficulty of such interdisciplinary inquiry.

2.2.1 A Semiology of Graphics: Bertin

> To utilize graphic representation is to relate the visual variables to the components of the information. With its eight independent variables, graphics offers an unlimited choice of constructions for any given information. (...) The basic problem in graphics is thus to choose the most appropriate graphic for representing a given set of information. — Bertin [6, p. 100]

Jacques Bertin (1918–2010) was a French cartographer, born in the suburbs of Paris and educated in the School of Cartography at the Sorbonne. An esteemed map-maker, he contributed to new methods of cartographic projection as the head of research at France's National Center for Scientific Research (CNRS) [58]. Yet his most widespread legacy would be the first and most far-reaching effort to provide a theoretical foundation to the design of information graphics, first offered in the text *Sémiologie Graphique* [5].

Bertin's volume resists concise summary,[1] though its most oft-cited concepts in contemporary writing, are the *visual variables* and *levels of organization*, which taken together form a table of perceptual properties: a heuristic for information-visual mapping (Fig. 2.2a). Bertin organized the tools at our (external) representational disposal in terms of space (two *planar dimensions*: location on a surface) and the visual (*retinal*) properties along with marks positioned within the space can vary: size, value, texture, color, orientation, and shape. In short, the visual variables

[1] Any attempt to summarize the 400 page volume would be too brief, and this author is convinced that although widely cited, the depth of Bertin's intellectual contributions is underestimated on account of opaque linguistic constructions. Bertin also contributed theory on *levels of reading* [p. 141], *stages of processing*[140], *functions of graphics*[p. 160], and *information processing*[p. 166]. The motivated reader is strongly encouraged to give "Part 1. Semiology of the Graphic Sign-System" a close reading [6].

Early Theoretical Contributions to Graph Comprehension

Year	Author	Key Contributions
1967	Bertin	**visual variables; levels of organization**
1981/2	Pinker	early version of Pinker 1990, as MIT report
1983	Bertin	english translation by WJ Berg
1984	Cleveland & McGill	**ordering of elementary perceptual tasks** (codes); (*re-articulates Bertin's visual variables with partial accuracy rankings*)
1985	Kosslyn	*Book review in the J. Amer. Statistics Assoc contained thorough but accessible primer of contemporary information processing psych as applied to graphics*
1986	Mackinlay	codification of graphic design criteria in a form that can be used by the presentation tool, including expanded (theoretical) ranking of elementary codes
1987	Cleveland & McGill	expanded set of **elementary codes** with refined accuracy rankings
1987	Simkin & Hastie	**judgement tasks; elementary mental processes** *(demonstrates interaction of encoding & task; positions Cleveland & McGill in context of Pinker & information procesing)*
1989	Kosslyn	**analytic scheme for deconstructing graphs; acceptability principles** *(thorough treatment, framing common graphical intuitions in terms of information processing)*
1990	Pinker	***first general process account***; *(schema-theoretic account from information processing perspective)*
1993	Lohse	*computational (symbolic, GOMS) production-system model predicting scanpath & response time from question & graph*
1994	Gillan & Lewis	*computational Mixed Arithmetic-Perceptual (MA-P) model derived from task analyses*
2002	Shah & Freedman	**construction-integration model of graph comprehension** (*builds upon Pinker 1990 to integrate iteration & prior-knowledge driven processing)*
2002/3	Peebles & Cheng	*ACT-R/PM based computational model capable of predicting scanpaths on cartesian graphs under questions*
2008	Trafton, et. al	*argues for explicit inclusion of 'spatial processing' and 'cognitive integration' in existing models*

Fig. 2.1 Early influential theories, frameworks, and models in Graph Comprehension [32, 49, 59, 60, 76]

(A) Bertin (1967, 1983)

LEVEL OF THE VARIABLE

VISUAL VARIABLES	ASSOCIATIVE *(similar)*	SELECTIVE *(different, groups)*	ORDERED *(ordered)*	QUANTITATIVE *(proportional)*
	Position	Position	Position	Position
	Size	Size	Size	Size
	Color (value)	Color (value)	Color (value)	
	Texture	Texture	Texture	
	Color (hue)	Color (hue)		
	Orientation	Orientation		
	Shape			

(B) Cleveland & McGill (1984, 1987)

DATA TYPE

ELEMENTARY PERCEPTUAL TASKS: QUANTITATIVE *(1984)*	ELEMENTARY CODE: QUANTITATIVE *(1987)*
Position (*along a common scale)*	Position (*along a common scale)*
Position (*along a non-aligned scale)*	Position (*along a non-aligned scale)*
Length, Direction, Angle	Length
Area	Angles
Volume, Curvature	Slopes*
Shading, Color (saturation)	Areas
	Volumes
	Densities
	Color (saturation)
	Color (hue)

(C) Mackinlay (1986)

DATA TYPE

PERCEPTUAL TASKS	NOMINAL	ORDINAL	QUANTITATIVE
	Position	Position	Position
	Color (hue)	Density	Length
	Texture	Color (saturation)	Angle
	Connection	Color (hue)	Slope
	Containment	Texture	Area
	Density	Connection	Volume
	Color (saturation)	Containment	Density
	Shape	Length	Color (saturation)
	Length	Angle	Color (hue)
	Angle	Slope	
	Slope	Area	
	Area	Volume	
	Volume		

Fig. 2.2 Four contributions ranking perceptual accuracy of visual-spatial encodings. Bertin (**a**) was reasoned phenomenologically, Cleveland and McGill (**b**) derived from experimental studies with quantitative proportion judgments, which (**c**) Macklinlay [51] extended for nominal and ordinal data reasoning from existing psychophysics studies, not empirically validated in the context of graph comprehension

offer eight channels into which information can be mapped. Bertin argued these channels have varying capacities for adequately representing different aspects of information: a correspondence between the nature of the information and perceptual requirements for discerning it in graphical form. In an orthogonal scheme, he posited four *levels of organization* that govern what *about* some information we might seek to perceive. Selective perception involves discerning categorical belonging; associated perception grouping like instances; and ordered perception discerning step-wise order and quantitative perception discerning the absolute value of an instance or numeric ratio between instances. Bertin asserted that to map data to a visual variable, the level of organization of the data must correspond to the capacity of the visual variable (Fig. 2.2a). Any mismatch is a source of "graphic error" [6, p. 64].

Bertin envisioned a unifying framework that could govern the design of all kinds of graphics. A CNRS colleague reflected that it was the exposure to hundreds of representations from different scientific domains—brought to Bertin for advice—that endowed him with the sort of global perspective required to write a text as comprehensive as Sémiologie Graphique [7]. In modern parlance, we would say Bertin offered a structured design space for mapping information-to-graphical marks. Though it is important to note that these ordered mappings were inferred from a combination of logical reasoning and perceptual experience rather than experimental evidence. Bertin's treatise is partially descriptive: structuring his observation of the components of graphical communication, and prescriptive: offering guidelines for how and when certain mappings should be made. In justification of the levels of organization assigned to each variable, Bertin offers a test, a sort of phenomenological self-check (or to the researcher, suggested experimental task) that should convince the reader. In this way, the classification of visual variables can be read as a set of hypotheses for controlled psychophysics experiments. The continued influence of Bertin's work should remind us of the value of the kind a priori theorizing required to construct such a theoretical framework. He did not conduct experiments or build models to explain data, but rather imposed a coherent logical structure on a disorganized set of phenomena growing rapidly in importance. Though perceptual experiments would follow, Bertin's visual variables still stand as the most common starting point for information-graphic mapping in visualization design. His work is widely cited in the pioneering research in computer graphics and information visualization, as well as the psychological studies of graphical perception that would begin in earnest in the 1980s.

2.2.2 Elementary Structures in Graphical Perception: From Cleveland and McGill to Simkin and Hastie

> We do not pretend that the items on our list are completely distinct tasks; for example, judging angle and direction are clearly related. We do not pretend that our list is exhaustive;

> for example, color hue and texture (Bertin 1973) are two elementary tasks excluded from the list because they do not have an unambiguous single method of ordering from small to large and thus might be regarded as better for encoding categories rather than real variables. Nevertheless the list ... is a reasonable first try and will lead to some useful results on graph construction. — Cleveland and McGill [16, p. 532]

The Semiology of Graphics would not be published in English until 1983, and as graphic displays of information became prevalent in American statistical journals in the early 1970s, calls were made for more systematic inquiry. A "theory of graphical methods" was needed [21, p. 5] in order to overcome the state of "dogmatic and arbitrary" design guidance of the time [45, p. 29]. William Cleveland and Robert McGill were statisticians at Bell Labs when they answered this call, publishing a series of empirical studies in the Journal of the American Statistical Association (JASA) which they described as theory for the relative accuracy for a set of *elementary perceptual tasks* readers perform to extract the values of real variables from statistical graphs [16]. In subsequent years, Cleveland and McGill refined their terminology, replacing *perceptual tasks* [16] with *graphical-perceptual tasks* [17], *basic graphical judgments* [18], and finally, *elementary codes* [19], with influential publications spanning venues of statistics, HCI, and popular science. Claims made in their initial 1984 work were tested by additional experiments and deeper engagement with contemporaneous theories of vision, resulting in the much refined 1987 publication ranking accuracy of an expanded set of *elementary codes* (Fig. 2.2b).[2] These codes describe channels available for mapping quantitative information to graphic form. In this sense, the authors re-articulated the visual variables described by Bertin [5, 6] and further ordered them according to human accuracy in making quantitative relational judgments. Cleveland and McGill's variables do not match those of Bertin and, however, are admittedly neither exhaustive nor mutually exclusive [16, p. 532]. One explanation for this discrepancy is their having conceived of the codes on the basis of their personal experience with statistical graphs, while Bertin set out to theorize a structure that could account for the visual-spatial properties of all graphic marks on 2D surfaces.

Cleveland and McGill's approach was partially deductive—structured a posteriori from personal experience and perceptual theory (e.g., [74]) and inductive, generalizing from reviews of psychophysical experiments (e.g., [4]), and their own original studies. It is perhaps most accurate to characterize their studies as tests of Bertin's hypotheses for the appropriate visual variables for quantitative perception. The experimental task asked participants—presented with two marked graphic components—to indicate "what percentage the smaller is of the larger" (p. 539), an operationalization of Bertin's test for quantitative perception: "ask the reader the value of the larger sign if a value of one is attributed to the smaller sign" [6, p. 69].

[2] Nonetheless, the more preliminary 1984 publication remains the most widely cited of their works, with nearly eight times as many citations as the 1987 elaboration [as reported by Google Scholar and Web of Science, January 2021]. This observation reinforces the importance of tracing the intellectual history of theoretical works to find their most mature form and should serve as a warning against cherry-picking references.

While Bertin reasoned that only the planar dimensions (spatial location) and size can adequately communicate quantitative information, Cleveland and McGill give us the relative accuracy of ten encodings for the same task. Their experimental data support Bertin's hypothesis that spatial location (e.g., position along common scale, position along non-aligned scales) can carry this information most accurately. If *length* is imputed as the size variation of a line [6, p. 71] and area the size variation of a point, then the data support Bertin's conclusions about the size variable, but not in relation to direction (Bertin's orientation for line) or angle (potentially construed as shape). There is enough discrepancy suggested in the empirical results to warrant further scrutiny of Bertin's criteria for judging a variable as applicable to a particular level and of the experimental tasks themselves.

Four years later, Northwestern University psychologists David Simkin and Reid Hastie offered JASA a contextualization of Cleveland and McGill's elementary codes, under a framework of information processing psychology [72]. Simkin and Hastie emphasized that performance of graphical perception depends not only on the way information is encoded but also on the judgment tasks performed by the human beings for whom the graphs are intended. Building upon Follettie [26], they differentiated between measurement, discrimination, proportion, and comparison judgments (Fig. 2.3a). It is important to note that all of Cleveland and McGill's studies used proportion judgments. Follettie, and later Simkin and Hastie, brought awareness to a whole new range of judgment tasks for which statistical graphs are used. Most importantly, they demonstrated that choosing a graphic mapping for a variable of data should not only depend on the data type (Bertin's level of organization) but also on the judgment task the designer wants the reader to perform. They offered empirical demonstrations of the interaction between elementary codes and judgment tasks (e.g., comparison judgments were most accurate with simple bar charts (position along common scale) while proportional judgments were most

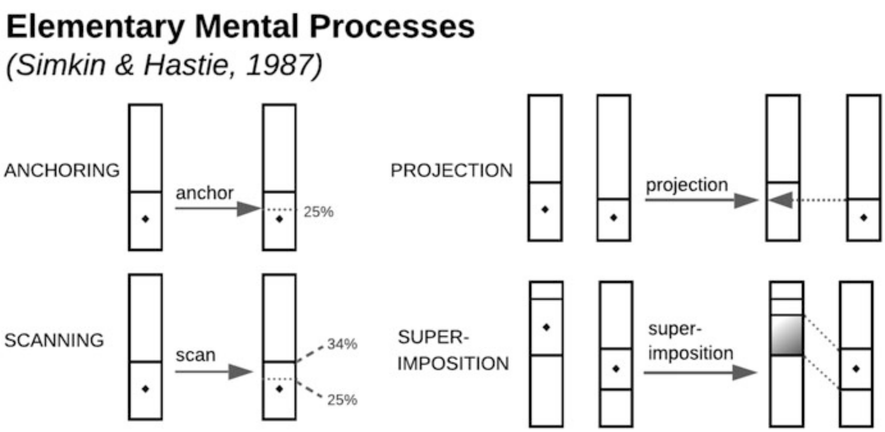

Fig. 2.3 Schematic diagram of Simkin and Hastie's theorized Elementary Mental Processes, adapted from (1987)

accurate with simple pie charts (angles)). Moving beyond encoding, they theorized four *elementary mental processes* that could—in an algorithmic sense—explain relative error and response rates across tasks (Fig. 2.3b). The elementary mental processes can be construed as visual data extraction steps: ordered in procedures that are executed by the perceptual system in order to accomplish a judgment task.

Over the course of the 1980s, the use of statistical graphics in publishing and data analysis surged with the development of software packages that made simple visualizations accessible for personal computer users. The cross-fertilization of empirical research between perceptual psychology and statistics demonstrated how demand for design recommendations can drive applied research questions that in turn inspire basic science research. Though the decade began with a focus on mapping information to visual forms, it would end with sophisticated hypotheses about how such mappings would interact with tasks, governed by perceptual rules, to elicit comprehension.

2.2.3 The Rise of Process Theories

Prior to 1980, there had been very little systematic research on the psychology of graph comprehension [84]. Over the course of the 1980s, methods and theories from cognitive psychology began to permeate the community in statistics concerned with graphical perception. Simkin and Hastie, notably, were psychologists, though they published their seminal work in the Journal of the American Statistical Association (JASA) rather than a journal of applied cognition or perception. Their contribution stood in direct conversation with the earlier work of Cleveland and McGill in the same venue. In [43], psychologist Stephen Kosslyn published in JASA a review of five books on charts and graphs, including Bertin [6], Tufte [77], and Chambers [13]. Rather than a straightforward critique however, Kosslyn offered a thorough primer on relevant concepts from cognitive psychology contextualized with respect to graph reading. He provided a sketch of contemporary visual information processing [52] and the distinction between short- and long-term memory [3, 47] before addressing the extent to which the practical guidance offered by each book comported with aspects of cognitive theory. Although its citation count pales in comparison to the aforementioned works, the importance of Kosslyn's contribution cannot be overstated. In this cross-disciplinary fertilization, he offered—like Bertin—a structure for thinking about the scope of what questions might be asked of graphical performance. He shared a simple (conceptual, process) model of visual information processing (Fig. 2.4) in which graph perception would be situated. To an application-focused community of statisticians *using* graphics, he brought a concise summary of relevant psychological constructs. While previous efforts focused on structural questions of encodings and tasks, Kosslyn drew attention to the way that graph reading unfolds as a *process*.

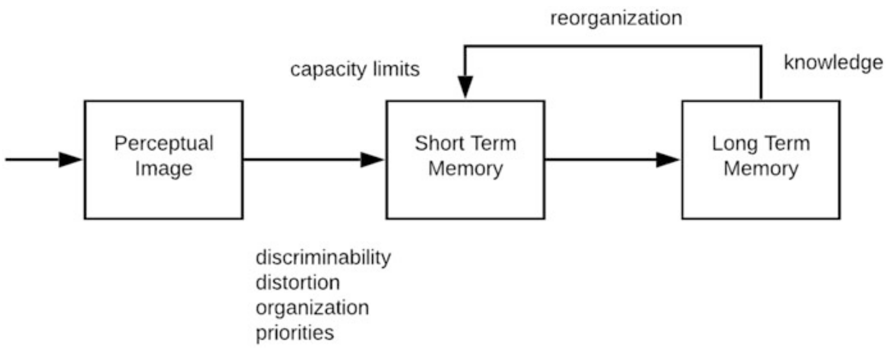

Fig. 2.4 A process description of visual information processing, adapted from [44]. The same figure appeared (without linguistic annotation of the important characteristics) in [43]

But Kosslyn's influence would not end there. In [44] he published an analytic scheme for deconstructing graphs[3] into constituent parts, which could then be analyzed at the levels of: syntactics (configuration of marks), semantics (the meaning that arises from configurations), and pragmatics (conveyance beyond direct interpretation of symbols). This contribution was more structural than procedural, offering a schema for evaluating graphs with respect to acceptability principles reasoned from cognitive theory. But in doing so, he would make reference to a forthcoming publication from his former graduate student Steven Pinker, one that would go on to stand as the most widely cited theory of graph comprehension.

2.2.3.1 A Theory of Graph Comprehension: Steven Pinker

While experimental psychologist Steven Pinker is most widely recognized for his popular science books on language and human nature, he got his start in the late 1970s as a doctoral student studying visual cognition with Stephen Kosslyn at Harvard. His chapter "A Theory of Graph Comprehension" in the book *Artificial Intelligence and the Future of Testing* would influence research on the design and function of visual-spatial displays across psychology, education, and computer science for decades [62]. In fact, the ideas were influential *before* publication, with earlier versions of the theory cited via MIT technical reports from the early 1980s.

Pinker's theory consists of a series of computational processes that propagate representations of information across components of a theorized human cognitive architecture (Fig. 2.5). He proposes that graph interpretation begins with construction of a *visual array*: a relatively raw, minimally processed representation of the

[3] Kosslyn makes a distinction between charts (specifying discrete relations between discrete entities) and graphs (a more constrained form, requiring at least two scales associated via a "paired with" relation).

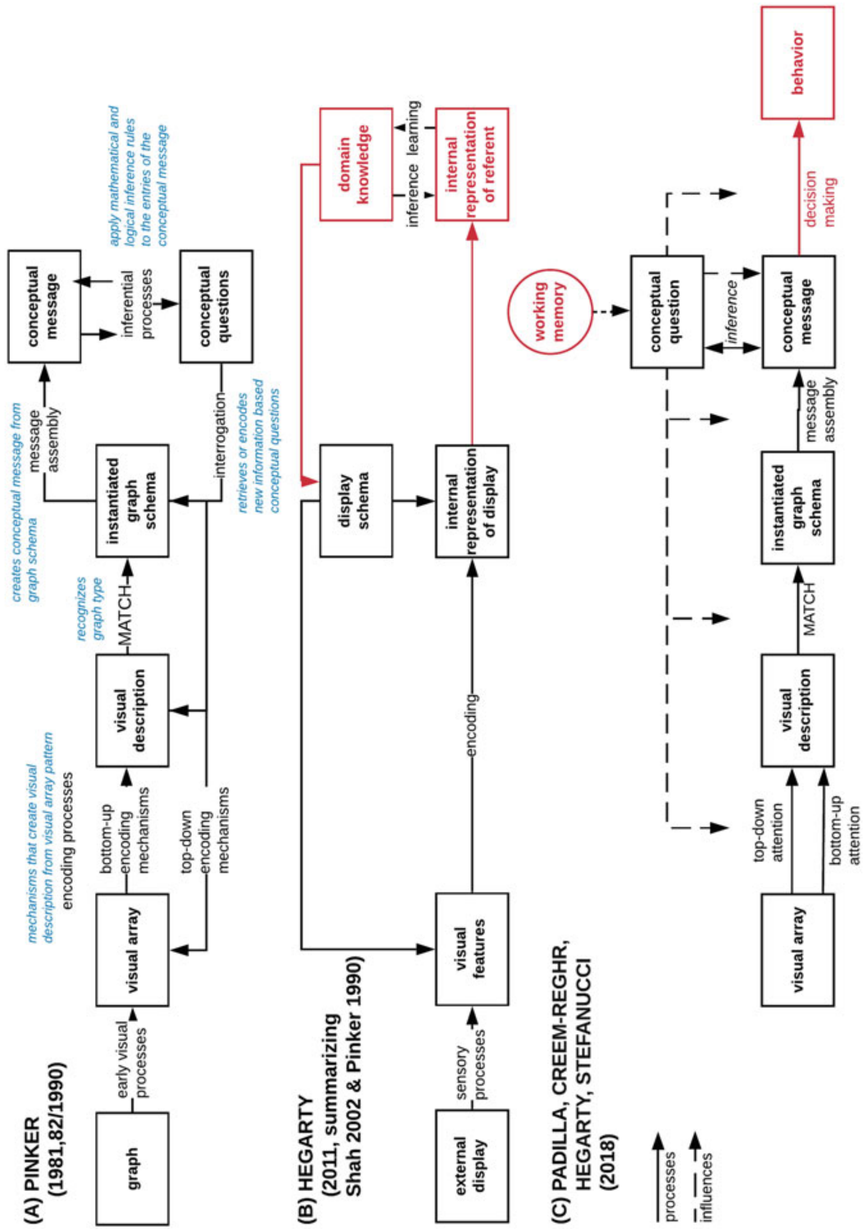

Fig. 2.5 Three versions of Information Processing accounts of Graph Comprehension. Italic annotations in blue indicate clarifications, and red indicates changes from prior models. In reading these diagrams, it is important to recognize they represent processes, not components. The boxes in Pinker, for example, indicate representations of information, not theorized cognitive structures, like working memory or executive control. The diagrams are not schematics for the structure of a cognitive system, but schematics of how information is processed, and care must be taken to avoid inadvertently reifying them into component structures, which might serve an *implementation* level of analysis

information made available to the nervous system via patterns of intensity on the retinas. The visual array is then *encoded* into a *visual description*: a symbolic, structural representation of the scene in a form more efficient for computation with knowledge in memory. A *MATCH process* then compares the visual description with the contents of memory in order to select the correct *graph schema*—a sort of placeholder indicating the structural relation of information for that particular class of graph. Once *instantiated*, information from the visual description is structured according to the relations of the selected schema. By this point, the external representation of the graph has been transformed into an internal representation in some structured, symbolic form that can be interrogated (searched) in order to extract information. Pinker uses the term *conceptual question* to refer to the information the reader wishes to derive from the graph and *conceptual message* the information that is actually extracted. A *message assembly process* searches the instantiated graph schema for information to translate to the form of the *conceptual message*. But processing capacity limitations prevent all the information from being automatically translated to messages. Rather, the *interrogation* process searches the graph schema for information matching the conceptual question. If it is found, message assembly takes over. But if not, *interrogation* can traverse the prior stages of representation (the visual description, then visual array) until the desired information is found, a top-down search that may require re-encoding the visual array. Finally, Pinker appeals to a general class of (logical, mathematical, and qualitative) *inferential processes* that operate on the conceptual message in service of answering the conceptual question.

Pinker's approach was deeply situated in the tradition of information processing, expressing an orientation toward a computational theory of mind. His explanation functions at Marr's *algorithmic* level of analysis—specifying representations and procedures for transforming them [52]. He offers an exceptionally detailed account of the properties of the representations he proposes (especially the visual description) and how they comport with cognitive theory in vision, memory, and attention. The 1990 publication is not an easy read, and it is my personal opinion that its scope is often misunderstood and contribution inadvertently reified *as* its diagrammatic representation of information processing.[4] Figure 2.5a is adapted from Pinker's Figures 4.14 and 4.19 which he characterizes as "representing the flow of information specified by the current theory" [62, p. 104]. The diagram depicts the order of representations and names of processes that transform them but fails to adequately describe re-encoding of the visual array (by re-attending to the graph) or the timecourse of decay of any representation based on the capacity limits of short (i.e., working) memory (e.g., [62, p. 89]). This leads to the misconception that Pinker does not address the role of working memory or proposes that an entire

[4] Just as we are drawn to graphs of empirical results, we are drawn to diagrams of theoretical offerings. The readers are warned against assuming that a diagram *entirely represents* a theoretical account, and writers encouraged to explicitly describe the representational role of diagrams in the scope of their theory.

graph is encoded in a single linear process. Rather, it is more appropriate to construe the diagrammatic representation as a snapshot of the flow of information through a single iteration of a bottom-up (perceptually driven) loop. We are similarly left wondering "where" in the mind his representations exist. This is not explicitly defined in the process diagram nor the text, but it can be reasonably inferred that all posited internal representations exist in short term (i.e., working) memory, as this is where processing would occur in the context of the cognitive theories he references (with the exception of the uninstantiated graph schema, likely in long-term memory).

Most importantly, justification for the theory rests on a single proposition: that graph comprehension exploits general purpose cognitive and perceptual mechanisms. Pinker's chapter was not the culmination of decades of empirical experimentation with graphs, but rather, the application of contemporaneous theories of vision, memory, and attention to the phenomenon of graph comprehension. This statement is not offered in critique, but in observation of the variety of ways that theory is developed. In this case, refutation rests on change to theories of vision, attention, and memory or evidence that graph comprehension is sufficiently different from the phenomena used to construct those theories to warrant special purpose cognitive mechanisms.

2.2.3.2 A Construction-Integration Model: Shah and Colleagues

An alternative to refuting a theory is refining it, by elaboration (specifying detail) or contextualization (situating in larger scope). In the late 1990s and early 2000s, Priti Shah and colleagues arguably did both: zooming out to describe the iterations of information processing when comprehending a graph and zooming in to elaborate the influence of "top-down" factors.

While prior experimental work focused on the perceptual aspects of graph comprehension, Cognitive Psychologist Priti Shah's mid-1990s dissertation work emphasized the role of *cognitive processes* in graph comprehension. Though contemporary Cognitive Science resists a precise delineation between perception and cognition, in graph comprehension a distinction is typically drawn between sources of information. Perception—information arriving via the senses—is referred to as "bottom-up" processing, while prior knowledge and computation over internal representations is referred to as "top-down" processing. Like Pinker, Shah, and her colleagues reasoned that graph comprehension would make use of general purpose cognitive processes rather than some special graphics engine in the mind. Drawing inspiration from Walter Kintsch's well-regarded Construction-Integration Theory [41], Shah elaborated how the processes of constructing meaning with a graph might proceed in the same fashion as constructing meaning from text or linguistic discourse.

Along with Patricia Carpenter, Shah first drew attention to the timecourse of information processing when reading a graph [12, 67]. Prior perceptual accounts tended to emphasize holistic pattern recognition processes that allow the readers to make the sort of quick proportional judgments used in studies of graphical perception. Carpenter and Shah employed more complex tasks, asking the readers to describe graphs and answer comprehension questions. Performance on these tasks, accompanied by measurements of eye fixations, revealed a more iterative procedure was taking place: one that involved a serial identification of visual chunks, followed by inferences and reasoning, repeated until the task goal had been accomplished. Along with evidence of differential task performance based on prior knowledge of semantic content, their studies provided support for the claims that (1) successful graph interpretation depends not only on appropriate information-to-graphical encoding but also on prior knowledge and skill of the graph interpreter and (2) graph comprehension is an iterative, multi-stage process. Publications in 2002 drew more strongly from CI Theory, characterizing the timecourse of processing in terms of two phases: an initial *construction phase*, where visual chunks activate relevant prior knowledge and are integrated into a coherent representation, and an *integration phase*, where inferences are made over the (coherent) representation (Fig. 2.6a) [30, 68]. The phases follow in order, though can be repeated, and integration can be followed by further construction, as necessary (Fig. 2.6b).

The astute reader will ask how Shah's Construction-Integration Model relates to Pinker's [62] Theory of Graph Comprehension. The answer depends on one's interpretation of each text. In a 2005 review, Shah and colleagues describe their model as differing from Pinker's in that it specifies that prior knowledge (and in turn, expectations) is activated by the encoding of visual chunks, which serve as a top-down constraint on inferential processing [69]. Pinker also describes the activation of prior knowledge, though in slightly different terms. Specifically, the MATCH process "searches" prior knowledge in order to instantiate an appropriate schema (prior knowledge structure) for the type of graph being perceived [62, p. 101]. In this way, the prior knowledge of graph type is activated by the (symbolic) visual description of the graph (the encoded visual chunk). Since inferential processes act on the instantiated graph schema, this prior knowledge serves to constrain interpretation. What Pinker does not explicitly describe is the activation of prior *domain knowledge*, or any understanding the reader has about the information being represented by the graph, though a generous interpretation would be that he includes this constraining influence under the scope of *inferential processes* (p. 103), a catch-all term to describe all of the higher order processing (logical, mathematical, judgments, and decisions) that one performs *on* the instantiated graph schema. If Shah's *coherent representation* is equated with Pinker's *instantiated graph schema*, then the two accounts are congruous. They are consistent in appealing to general purpose mechanisms, to describing a serial process of encoding, some form of integration with prior knowledge, and inferential processing. They both posit the existence of internal representations: Pinker gives a specific account of a plausible form of these representations, Shah requires only that they exist, leaving the CI model with less explanatory power for mechanisms, but greater robustness to change in the perennial debate on the nature of internal representation. It is *this* author's reading that these two accounts of graph comprehension are highly compatible, serving to elaborate different aspects of graphical processing at different levels of

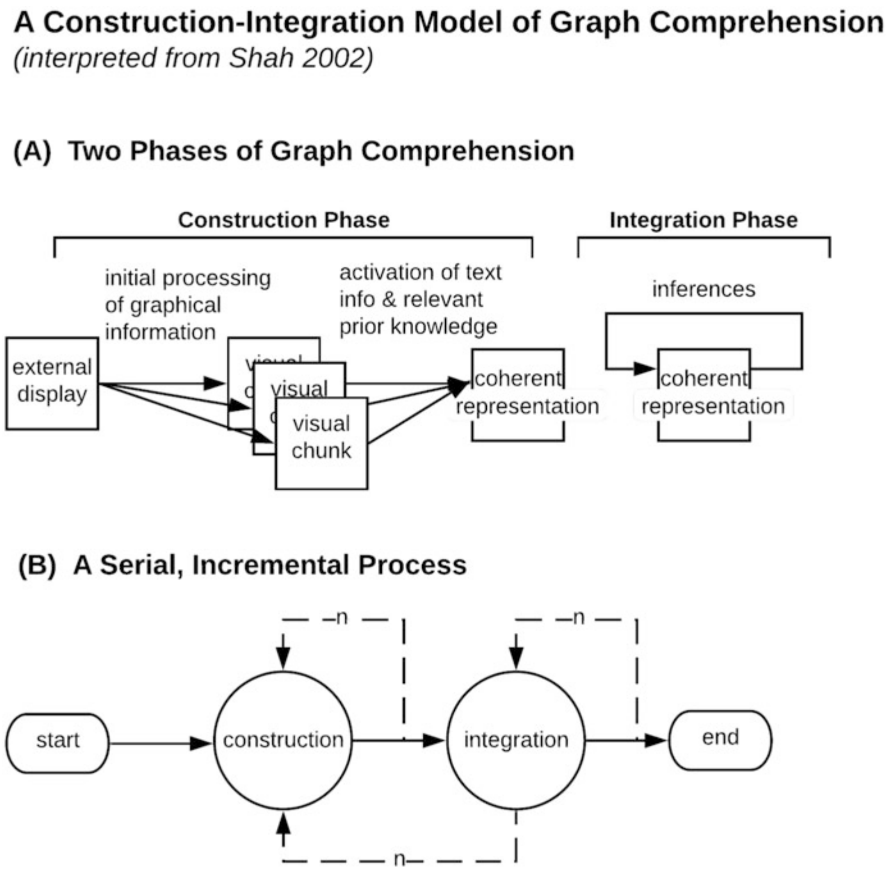

Fig. 2.6 A Construction-Integration Model of Graph Comprehension, derived from the text description in [30, 68]. (**a**) describes two distinct phases of comprehension: the first involves encoding visual chunks, while the second involves higher order cognitive processing over the working internal representation. (**b**) describes how integration follows some number of iterations of construction, before processing is either complete and ready for integration

specificity. While Pinker attends to a computationally plausible encoding structure for graphical information, Shah attends to the more global timecourse of processing and iterations of "perceptual" and "cognitive" efforts. They both offer testable predictions about how factors of the graphical display *and* the graph reader should differentially influence task performance.

2.3 The Landscape of Contemporary Research

Statistical graphics have never been more prevalent than they are today in scientific inquiry, business operations, or popular media. With such a wealth of applications, it is a good time to be a Visualization Psychologist but is not easy to *study* the psychology of visualization because as an applied area of inquiry, both students and scholars alike must navigate an opaque disciplinary milieu. The readers can find relevant empirical research in venues as distinct as journals and conferences of science or math education, learning science, information and library science, cognitive, educational, perceptual or (general) experimental psychology, vision science, cognitive science, and of course computer science—where the conference triad *InfoVIS, SciVIS, and VAST* claim some epistemic authority of the subject matter by virtue of naming rights.

In the two decades since Shah's Construction-Integration model, we have not seen similar overarching, general process accounts of comprehension. Rather, the researchers across these fields have progressively elaborated a complex ecosystem of factors that influence performance on graph comprehension tasks. We can organize these factors into three groups: those pertaining to the display, the individual, and the situation.

Display Factors The research on display characteristics tends to center on determining the most ideal encoding of information, a question of design. Bertin offered the first experientially deduced guidelines for mapping data to graphic marks [5, 6],[5] some of which were experimentally tested using relational judgment tasks and ranked by Cleveland and McGill [16, 19] and further extended by Mackinlay [51] who ranked encodings according to theorized perceptual accuracy for communicating quantitative, versus ordered, verses categorical data (see Fig. 2.2c). If humans were perceptual computers, this might be the crux of visualization psychology. But we are, of course, more delightfully nuanced creatures. Contemporary research has demonstrated that effectiveness of encodings depends not only on the capacity of a particular type of mark to carry a certain type of information but also on what *about* that information the designer wants the reader to perceive most effortlessly. Ensemble encoding, for example, relies on characteristic performance of the visual system to inform encoding choice when the goal is to facilitate, for example, identification of an outlier, versus recognition of a statistical mean, or apprehension of clusters within the data [75]. Design choices within a particular encoding strategy are nuanced as well, as evidenced by research on the use of color. Color hue has been shown to be particularly effective for encoding data for nominal or absolute value judgments, while color brightness is superior to hue when encoding the same data for *relative* judgments [10, 55]. The plot thickens—design choices become more complex—when visualizing more than one variable and the interactions between

[5] The oft-overlooked footnote to these heuristics is that the rankings are meant to apply when the reader's *task* is an "elementary reading" (extracting a specific value).

encoding strategies need be considered. Smart and Szafir recently demonstrated that the shape of a graphic mark significantly influences perception of color and size [73]; whatever the designer's most informed intentions, their efforts can be thwarted by interactions between decisions they make. Similarly, visual saliency (how "attractive" an area is to the eye) has been shown to influence how humans attend to visual stimuli [38]; though recent efforts to computationally reconcile bottom-up saliency models top-down "cognitive" models have proven ineffective at predicting gaze behavior [48]. While display characteristics were the focus of the earliest research in graph comprehension, they receive no less attention in modern research efforts. Designers need practical guidance on when and how to use animation [8, 79] and 3D [68], how to use signals or instructions to augment a display and scaffold comprehension [1, 28, 42, 54], and how to use interaction most effectively [61, 66]. Since the time of Cleveland and McGill, research on display characteristics has become increasingly nuanced, revealing more factors that influence how a display should be designed and the interactions between them.

Individual Factors Research on individual differences, or factors that give rise to differential performance with the same graphic display, is most common in cognitive and educational psychology and learning science. As Carpenter and Shah argued, "individual differences in graphic knowledge should play as large a role in the comprehension process as does variation in the properties of the graph itself" [12, p. 97]. But what is meant by *graphic knowledge*? In empirical work, graph knowledge is tightly entwined with graph reading abilities and expertise. The terms graphicacy, graphical literacy, graph sense, graphical competence, and representational competence are used throughout the literature in psychology and education to refer to a reader's ability to understand (and potentially create) information displayed graphically. If graph comprehension is the act of deriving meaning from a graph, then *graphicacy* is its educational flip side: the ability to perform a graph comprehension task. Some have treated this ability as a foundational step in cognitive development, akin to numeracy and literacy [31]. Others treat the ability as a practice, implicating the importance of experience and socio-cultural influences [64, 65]. In education in particular, the researchers have pursued general learner characteristics that might serve as pre-requisites or predictors of these graphing abilities, including mathematical ability [23], working memory [12], and spatial reasoning [81]. Ulrich Ludewig's recent doctoral dissertation offers a thorough reconciliation between perspectives of graph comprehension and graphicacy [50]. It is slightly easier to differentiate between ability and knowledge with respect to specific graphs, for example, domain knowledge of the information represented in a particular graph, and knowledge of that particular representation's graphical formalisms. The act of graph reading requires that we use our knowledge of a graph's formalisms to perform some task (e.g., extract a value, detect a trend), thereby "learning" something about the domain. In my own research, I have demonstrated that this procedure is not reciprocal. It is much more difficult to use prior knowledge of a domain to "reverse engineer" understanding of a graphical formalism, such as may be required to understand an unfamiliar or unconventional type of graph [28, 29]. A reader's understanding of the concepts represented in a graph has been shown to guide not only the reader's interpretation of the display [63] but early perceptual processing as well [68]. In some cases, a reader's expectations seem to "inoculate" them from true relations presented in the data or lead them to over or underestimate the magnitude of relations. Conversely, domain knowledge has been shown to support comprehension by making the readers more likely to ignore "noise" in data [86]. More recently, Jessica Hullman and colleagues have explored the role of prior beliefs [37, 40] and even judgments of expectations of others [36] on graph interpretation. Taken together, the research on characteristics of individuals has provided strong evidence for "top-down" influences on graph comprehension.

Situational Factors Factors that change comprehension performance of an individual with a particular display depending on the *situation* are the least structured, thus least understood pieces of this factorial puzzle. Affect (emotion) and motivation clearly influence human performance of any task, and although these are characteristics of an individual, we classify them as situational because they are more situationally variable—in the context of a repeated measures study, for example—than the relatively stable[6] factors like prior knowledge or ability. *Task* is the most studied situational factor, though it is at present a hierarchical concept poorly operationalized across the literature. The term "task demand" is used to indicate a variety of contextual factors, from a relatively low-level step of information extraction (i.e., a micro-step in a larger process, such as identifying a location of interest in a graph), to a specific task or goal provided to a reader in an experiment (e.g., extract a value, compare two points, characterize a trend), to the context of some cognitive activity (e.g., analyzing data, making a decision, forecasting, solving a problem), and to the communicative intent of the designer (e.g., to inform, educate, entertain, persuade, etc.). In the beginning, there was but a single task: Cleveland and McGill's proportional judgments [16, 19]. Folettie, followed by Simkin and Hastie, elaborated further judgments (measurement, discrimination, and (non-proportional) comparison) [26, 72]. Bertin also addressed tasks, proposing three "levels of reading" [6, p. 141]. Other tripartite classifications have been proposed in the same vein, all structuring how much of the depicted information the reader need attend to, and how explicit or precise their response should be [5, 6, 23, 31, 83]. In their application of ensemble encoding theories to visualization, Szafir and colleagues offer a parallel taxonomy of four tasks-types that require visual aggregation [75]. These can be partially but not entirely mapped onto the extant tripartite structures. The most complete deconstruction of the concept of task can be found in Brehmer and Munzner's, "Multi-Level Typology of Abstract Visualization Tasks," which surveyed an impressive volume of prior task frameworks in computer graphics and visualization, visual analytics, human–computer interaction, cartography, and information retrieval [9]. A fruitful undertaking for

[6] Variability, of course, depends on the scope of time under consideration.

visualization psychology would be to extend this typology to include the tripartite classifications that grew out of education, the lower level tasks elaborated in vision science, and higher level "communicative context" that is evident in the structure of the field of visualization itself [27]. A strong underlying assumption of much research in graph comprehension (and visualization writ-large) is that the graph designer's goal is to clearly communicate, "the truth" of some data to the reader. Thus, the graph should be maximally informative and minimally difficult—the graphical equivalent of Grice's maxims for communication. But research in learning science has taught us that sometimes difficulty is *desirable*. Perhaps if my graph is for *learning*, I might encode data differently so as to scaffold a reader's process of discovery and more deeply engage with the data. Alternatively, if the context of my communication is *persuasion*, I might use more signals to direct reader's attention than I would if the context were exploratory analysis. The role of communicative context is seen structurally through the emergence of specialized workshops at the IEEE VIS conference but has not yet been systematically investigated across a full range of communicative tasks. My own theoretical intuition—reasoned from design experience and engagement with the literature—is that situational factors are those that present mediating or moderating influences on other individual and display characteristics, at either the time of design or comprehension.

A primary challenge facing designers and researchers alike is the sheer number of factors found to influence comprehension and the fact that they are typically studied in limited clusters, inconsistently operationalized between studies and across disciplines. This makes it difficult to conceive of the complex interactions that may exist between factors and how to go about constructing nuanced guidelines for designers. The most comprehensive summaries of factors can be found in [31, 33, 70] and [35], which features a concise set of empirically grounded principles for display design that would make a useful addition to the wall of any graph designer.

2.4 What Remains to Be Discovered

The good news is that "the state of our (sub) discipline is strong." The bad news is that it is difficult to navigate and even more difficult to *integrate*. In the two decades since the last publication of a general process theory of graph comprehension [68], the march of empirical research has only quickened, offering insight into factors that affect graph comprehension, but in forms too piecemeal to be fruitfully and consistently applied. There are myriad open questions to be answered, from how exactly factors interact to influence performance to how performance is expressed in different forms of cognitive activity: decision-making vs. problem solving, forecasting, learning, or creative construction. We need to explore our boundaries: how does interaction with the narrowly defined class of "graphs" compared to the broader class of diagrams or external representations, in general? (see [14, 15] for thorough treatments). And our field too must address the challenge of traversing "lower levels" of explanatory analysis: there is a tremendous gulf of explanation between conceptual models of graph comprehension and understanding of how these processes are enacted in the body.

Hegarty [71] and more recently Padilla [56] have convincingly argued for the importance of *cognitive models* in guiding visualization research. Hegarty suggests they are useful for predicting the effectiveness of designs and informing design decisions. Padilla argues that cognitive models can be used to promote innovation and evaluate validity of empirical research designs. In sum, they can bridge an important gap and presuming they are communicated in an appropriate venue, well-articulated models can help ensure that the "state of the art" in basic research is available to guide applied efforts in design and instruction. But what kinds of models do we need, and what makes a model *cognitive*?

Those seeking easy answers to these questions will fall quickly down a philosophical rabbit hole. Models in science come in all shapes and sizes, with differing levels of analysis and varieties of explanation. In the social and behavioral sciences alone, one finds component and structural models, conceptual models, computational models, and task-analytic and mathematical models. Models differ in what *aspect* of a phenomenon they explain (e.g., structures, relationships, processes), how they are justified (e.g., by phenomenological, experimental or task-analytic empirical evidence, by logic or appeal to reason), and the way they are represented (conceptually: typically via words and diagrams or computationally: via math and/or computer programs). The importance of clearly conceptualizing and subsequently articulating the scope and purpose and form of a model cannot be overestimated, as the failure to do so can have tragic consequences for the intellectual trajectory of a field.

Take, for example, [62] Theory of Graph Comprehension. Setting aside for the moment that it is characterized as a theory and not a model,[7] a quick inspection of its diagrammatic representation (Fig. 2.5a) will reveal no mention of memory. Does this mean that Pinker believed memory was not involved in graph comprehension? *No*, it means that the reader needs clarification on what aspect of the phenomenon Pinker's model explains: a propagation of representations and the processes that transform them. Close reading of the accompanying text reveals what was likely obvious to readers at the time: all of the representations and processing take place *in* some form of memory. Pinker might have chosen to represent this in the diagrams by locating the representations (boxes) inside other graphics representing memory structures. This would have been advantageous for subsequent theorists looking to position their own ideas in relation to his but would also have changed the type of model, from the flow information processing to the flow of information processing *and* component structures—taking on an additional Marrian level of analysis [52]. In applying Pinker's model to a specific cognitive activity (decision-making), Padilla and colleagues have done well to clearly articulate the role of memory,

[7] Theories are typically treated as superordinate to models, though their exact relation is a topic of debate in philosophy of science.

as well their interpretation of the construct of memory itself [57], implicating a multiple component conception where "a multicomponent system (...) holds information temporarily and mediates its use in ongoing mental activities" [20, p. 1160]. While these details may be superfluous for those keen to *apply* the model, they are absolutely essential for the ongoing intellectual dialogue expressed via works of scholarship that move our science forward. Imagine next year a ground-breaking study is published in a journal of experimental psychology that questions the multicomponent conception of working memory, supporting a rival account with implications for how visual attention is directed. Changes to the underlying constructs on which a model or theory rests should necessitate its re-evaluation, no different from the need for testing and upgrading software when the libraries on which they are built mature.

The obvious difficulty is that constructs are transient, under-specified, and certainly not versioned like packages of code. Too often the precise conceptualization of constructs is held as tacit knowledge instantiated in encapsulated research labs, propagated through limited networks via the exchange of students and postdoctoral scholars.[8] Too little space is allocated in our written scholarship to descriptions of what we *specifically mean* by the terms we use, a symptom of a drive toward innovation and novelty over depth of explanation. I propose that in theoretical scholarship we should strive to be a little more like academic philosophy, where precision and justification in language is not only valued but demanded. We should be novel in our applications, but religiously rigorous in our theory. Models and theories should exist in direct dialogue with those that come before, explaining *exactly* how and why they differ and offer sufficiently impactful differences to be worthy of inclusion in the scientific canon.

In this onerous challenge stands a role for visualization psychology: as a mediator between disciplines (computer science, psychology, and education) and between professions (basic and applied research, design, and instruction). As a community, visualization psychology can position itself at the intersection of these goal-driven efforts and moderate the construction of *reference models*, intended to integrate theory across disciplines and levels of analysis that is specifically related to our phenomena of interest. We need not be concerned with explaining precisely how memory or attention are instantiated by the body but should take responsibility for maintaining enough awareness of the progression of those basic theories, so we can apply and as needed update our own models of how such cognitive phenomena drive the performance of graph (and visualization) comprehension.

[8] see Kaiser [39] for a fascinating intellectual history of this phenomenon with respect to dialects of Feynman diagrams.

References

1. C. Acarturk, C. Habel, and K. Cagiltay. Multimodal Comprehension of Graphics with Textual Annotations: The Role of Graphical Means Relating Annotations and Graph Lines. In G. Stapleton, J. Howse, and J. Lee, editors, *Diagrammatic Representation and Inference*, Lecture Notes in Computer Science, pages 335–343. Springer Berlin Heidelberg, 2008.
2. American Statistical Association. Joint Committee on Standards for Graphic Presentation. *Publications of the American Statistical Association*, 14(112):790–797, 1915.
3. J. R. Anderson and G. H. Bower. *Human Associative Memory*. Human associative memory. V. H. Winston & Sons, Oxford, England, 1973.
4. J. C. Baird. *Psychophysical Analysis of Visual Space*. Pergamon Press, 1970.
5. J. Bertin. *Sémiologie Graphique. Les diagrammes—les reseaux—les cartes.* Mouton, The Hague, 1967.
6. J. Bertin. *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin Press, Madison, WI, 1983.
7. S. Bonin. Le développement de la graphique de 1967 à 1997. *Cybergeo : European Journal of Geography*, Nov. 2000.
8. J.-M. Boucheix and E. Schneider. Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction*, 19(2):112–127, 2009.
9. M. Brehmer and T. Munzner. A Multi-Level Typology of Abstract Visualization Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013.
10. L. A. Breslow, J. G. Trafton, and R. M. Ratwani. A perceptual process approach to selecting color scales for complex visualizations. *Journal of Experimental Psychology: Applied*, 15(1):25–34, Mar. 2009.
11. W. C. Brinton. *Graphic methods for presenting facts*. The Engineering magazine company, 1914.
12. P. A. Carpenter and P. Shah. A Model of the Cognitive and Perceptual Processes in Graph Comprehension. *Journal of Experimental Psychology: Applied*, 4(2):75–100, June 1998.
13. J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey. *Graphical Methods for Data Analysis*. Wadsworth Publ. Co, Belmont, CA, USA, 1983.
14. P. C.-H. Cheng. What Constitutes an Effective Representation? In M. Jamnik, Y. Uesaka, and S. Elzer Schwartz, editors, *Diagrammatic Representation and Inference*, Lecture Notes in Computer Science, pages 17–31, Cham, 2016. Springer International Publishing.
15. P. C.-H. Cheng. A Sketch of a Theory and Modelling Notation for Elucidating the Structure of Representations. In A.-V. Pietarinen, P. Chapman, L. Bosveld-de Smet, V. Giardino, J. Corter, and S. Linker, editors, *Diagrammatic Representation and Inference*, volume 12169, pages 93–109. Springer International Publishing, Cham, 2020.
16. W. S. Cleveland and R. McGill. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.
17. W. S. Cleveland and R. McGill. Graphical Perception and Graphical Methods for Analyzing Scientific Data. *Science*, 229(4716):828–33, Aug. 1985.
18. W. S. Cleveland and R. McGill. An Experiment In Graphical Perception. *International Journal of Man-Machine Studies*, 25(5):491–500, Nov. 1986.
19. W. S. Cleveland and R. McGill. Graphical Perception: The Visual Decoding of Quantitative Information on Graphical Displays of Data. *Journal of the Royal Statistical Society Series a-Statistics in Society*, 150:192–229, 1987.
20. N. Cowan. The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, 24(4):1158–1170, Aug. 2017.
21. D. R. Cox. Some Remarks on the Role in Statistics of Graphical Methods. *Applied Statistics*, 27(1):4, 1978.
22. F. E. Croxton and R. E. Stryker. Bar Charts versus Circle Diagrams. *Journal of the American Statistical Association*, 22(160):473–482, Dec. 1927.

23. F. R. Curcio. Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18(5):382–393, Nov. 1987. Publisher: National Council of Teachers of Mathematics.
24. W. C. Eells. The Relative Merits of Circles and Bars for Representing Component Parts. *Journal of the American Statistical Association*, 21(154):119–132, June 1926.
25. Y. Engelhardt. *The Language of Graphics*. PhD thesis, University of Amsterdam, 2002.
26. J. Follettie. Real-World Tasks of Statistical Graph-Using and Analytic Tasks of Graphics Research. *unpublished paper presented at the annual meeting of the National Computer Graphics Association, Anaheim, CA.*, 1986.
27. A. R. Fox. A Psychology of Visualization or (External) Representation? In *Proceedings of the Workshop on Visualization Psychology @ IEEE VIS*, 2020.
28. A. R. Fox and J. D. Hollan. Read It This Way: Scaffolding Comprehension for Unconventional Statistical Graphs. In P. Chapman, G. Stapleton, A. Moktefi, S. Perez-Kriz, and F. Bellucci, editors, *Diagrammatic Representation and Inference*, Lecture Notes in Computer Science, pages 441–457. Springer International Publishing, 2018.
29. A. R. Fox, J. D. Hollan, and C. M. Walker. When Graph Comprehension Is An Insight Problem. In *Proceedings of the Annual Conference of the Cognitive Science Society*, Montreal, Canada, 2019.
30. E. G. Freedman and P. Shah. Toward a model of knowledge-based graph comprehension. In *Diagrammatic Representation and Inference*, pages 18–30, 2002.
31. S. N. Friel, F. R. Curcio, and G. W. Bright. Making Sense of Graphs: Critical Factors Influencing Comprehension and Instructional Implications. *Journal for Research in Mathematics Education*, 32(2):124–158, 2001.
32. D. J. Gillan and R. Lewis. A Componential Model of Human Interaction with Graphs: 1. Linear Regression Modeling. *Human Factors*, 36(3):419–440, Sept. 1994.
33. N. Glazer. Challenges with graph interpretation: a review of the literature. *Studies in Science Education*, 47(2):183–210, Sept. 2011.
34. R. L. Harris. *Information Graphics: A Comprehensive Illustrated Reference*. Oxford University Press, 1999.
35. M. Hegarty. The Cognitive Science of Visual-Spatial Displays: Implications for Design. *Topics in Cognitive Science*, 3(3):446–474, 2011.
36. J. Hullman, E. Adar, and P. Shah. The Impact of Social Information on Visual Judgments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1461–1470, New York, NY, USA, 2011. Association for Computing Machinery.
37. Hullman, J. , Kay, M., Kim, Y. , and Shrestha, S. Imagining Replications: Graphical Prediction & Discrete Visualizations Improve Recall & Estimation of Effect Uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):446–456, Jan. 2018.
38. L. Itti and C. Koch. Computational modelling of visual attention. *Nature Reviews. Neuroscience*, 2(3):194–203, Mar. 2001.
39. D. Kaiser. *Drawing theories apart: the dispersion of Feynman diagrams in postwar physics*. University of Chicago Press, Chicago, 2005.
40. Y.-S. Kim, L. A. Walls, P. Krafft, and J. Hullman. A Bayesian Cognition Approach to Improve Data Visualization. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14, Glasgow Scotland Uk, May 2019. ACM.
41. W. Kintsch. Modeling comprehension processes: The construction-integration model'. In *Comprehension. A paradigm for cognition.*, pages 93–120. Cambridge University Press, Cambridge, 1998.
42. N. Kong and M. Agrawala. Graphical overlays: Using layered elements to aid chart reading. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2631–2638, 2012.
43. S. M. Kosslyn. Graphics and Human Information Processing: A Review of Five Books. *Journal of the American Statistical Association*, 25(4):R134–R136, 1985.
44. S. M. Kosslyn. Understanding charts and graphs. *Applied Cognitive Psychology*, 3(3):185–225, 1989.

45. W. Kruskal. Visions of Maps and Graph. *Auto- Carto II, Proceedings of the International Symposium on Computer Assisted Cartography, ed. J. Kavaliunas, Washington, D.C.: U.S. Bureau of the Census and American Congress on Survey and Mapping.*, pages 27–36, 1975.
46. J. Larkin and H. Simon. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 99:65–99, 1987.
47. P. H. Lindsay and D. A. Norman. *Human Information Processing*. Academic Press, New York, second edition, 1977.
48. M. A. Livingston, L. E. Matzen, A. Harrison, A. Lulushi, M. Daniel, M. Dass, D. Brock, and J. W. Decker. A Study of Perceptual and Cognitive Models Applied to Prediction of Eye Gaze within Statistical Graphs. In *ACM Symposium on Applied Perception 2020*, pages 1–9, Virtual Event USA, Sept. 2020. ACM.
49. G. L. Lohse. A cognitive model for understanding graphical perception. *Human-Computer Interaction*, 8(4):353–388, 1993.
50. U. Ludewig. *Understanding Graphs: Modeling Processes, Prerequisites and Influencing Factors of Graphicacy*. PhD thesis, Universität Tübingen, Tübingen, Germany, 2018.
51. J. Mackinlay. Automating the Design of Graphical Presentations of Relational Information. *ACM Transactions on Graphics*, 5:110–141, 1986.
52. D. Marr. *Vision*. The MIT Press, July 1982.
53. M. Massironi. *The Psychology of Graphic Images*. Psychology Press, 2001.
54. P. D. Mautone and R. E. Mayer. Cognitive aids for guiding graph comprehension. *Journal of Educational Psychology*, 99(3):640–652, 2007.
55. D. H. Merwin and C. D. Wickens. Comparison of Eight Color and Gray Scales for Displaying Continuous 2D Data. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 37(19):1330–1334, Oct. 1993.
56. L. M. Padilla. A Case for Cognitive Models in Visualization Research : Position paper. In *2018 IEEE Evaluation and Beyond—Methodological Approaches for Visualization (BELIV)*, pages 69–77, Oct. 2018.
57. L. M. Padilla, S. H. Creem-Regehr, M. Hegarty, and J. K. Stefanucci. Decision making with visualizations: a cognitive framework across disciplines. *Cognitive Research: Principles and Implications*, 3(1):29, July 2018.
58. G. Palsky. Jacques Bertin, from classical training to systematic thinking of graphic signs. *Cartography and Geographic Information Science*, 46(2):189–193, Mar. 2019.
59. D. Peebles and P. C. H. Cheng. Extending task analytic models of graph-based reasoning: A cognitive model of problem solving with Cartesian graphs in ACT-R/PM. *Cognitive Systems Research*, 3(1):77–86, Mar. 2002.
60. D. Peebles and P. C. H. Cheng. Modeling the effect of task and graphical representation on response latency in a graph reading task. *Human Factors*, 45(1):28–46, 2003.
61. W. A. Pike, J. Stasko, R. Chang, and T. A. O'Connell. The Science of Interaction. *Information Visualization*, 8(4):263–274, 2009.
62. S. Pinker. Theory of Graph Comprehension. In R. Freedle, editor, *Artificial Intelligence and the Future of Testing*, pages 73–126. Erlbaum, Hillsdale, NJ, 1990.
63. Y. Postigo and J. I. Pozo. On the Road to Graphicacy: The Learning of Graphical Representation Systems. *Educational Psychology*, 24(5):623–644, Oct. 2004.
64. W.-M. Roth. Toward an Anthropology of Graphing. In *Toward an Anthropology of Graphing: Semiotic and Activity-Theoretic Perspectives*, W.-M. Roth (Ed), pp. 1–21. Springer, Netherlands, 2003. https://doi.org/10.1007/978-94-010-0223-3_1
65. W.-M. Roth, L. Pozzer-Ardenghi and J. Y. Han. *Critical Graphicacy: Understanding Visual Representation Practices in School Science*. Springer, Netherlands, 2005. https://www.springer.com/gp/book/9781402033759
66. K. Sedig and P. Parsons. Interaction Design for Complex Cognitive Activities with Visual Representations: A Pattern-Based Approach. *AIS Transactions on Human-Computer Interaction*, 5(2):84–133, 2013.

67. P. Shah. A Model of the Cognitive and Perceptual Processes in Graphical Display Comprehension. In *Reasoning with diagrammatic representations II*, M. Anderson (Eds), pages 94–101. AAI Press, Menlo Park, CA, 1997.
68. P. Shah. Graph Comprehension: The Role of Format, Content and Individual Differences. In M. Anderson, B. Meyer, and P. Olivier, editors, *Diagrammatic Representation and Reasoning*, pages 173–185. Springer London, London, 2002.
69. P. Shah, E. G. Freedman, and I. Vekiri. The Comprehension of Quantitative Information in Graphical Displays. In P. Shah, editor, *The Cambridge Handbook of Visuospatial Thinking*, pages 426–476. Cambridge University Press, New York, NY, 2005.
70. P. Shah and J. Hoeffner. Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, 14(1):47–69, Mar. 2002.
71. P. Shah, R. Mayer, and M. Hegarty. Graphs as Aids to Knowledge Construction: Signaling Techniques for Guiding the Process of Graph Comprehension. *Journal of Educational Psychology*, 91(4):690–702, 1999.
72. D. Simkin and R. Hastie. An Information-Processing Analysis of Graph Perception. *Source Journal of the American Statistical Association*, 82(398):454–465, 1987.
73. S. Smart and D. A. Szafir. Measuring the Separability of Shape, Size, and Color in Scatterplots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI ’19, pages 1–14, New York, NY, USA, 2019. Association for Computing Machinery.
74. S. S. Stevens. *Psychophysics*. John Wiley, New York, 1975.
75. D. A. Szafir, S. Haroz, M. Gleicher, and S. Franconeri. Four types of ensemble coding in data visualizations. *Journal of Vision*, 16(5):11–11, Mar. 2016.
76. S. B. Trickett and J. G. Trafton. Toward a comprehensive model of graph comprehension: Making the case for spatial cognition. *Lecture Notes in Computer Science*, 4045 LNAI:286–300, 2006.
77. E. Tufte. *Visual Display of Quantitative Information*. Graphics Paper Press LLC, Cheshire, CT, first edition, 1983.
78. J. W. Tukey. *Exploratory Data Analysis*. Addison-Wesley, 1977.
79. B. Tversky, J. B. Morrison, and M. Betrancourt. Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57(4):247–262, Oct. 2002.
80. S. Ullman. Visual routines. *Cognition*, 18(1):97–159, Dec. 1984.
81. Velez, M.C. , Silver, D. , and Tremaine, M. Understanding visualization through spatial ability differences. In *VIS 05. IEEE Visualization, 2005.*, pages 511–518, Oct. 2005.
82. R. von Huhn. Further Studies in the Graphic Use of Circles and Bars: I: A Discussion of the Eells’ Experiment. *Journal of the American Statistical Association*, 22(157):31–36, Mar. 1927.
83. H. Wainer. Understanding Graphs and Tables. *Educational Researcher*, 21(1):14–23, Jan. 1992.
84. H. Wainer and D. Thissen. Graphical Data Analysis. *Annual Review of Psychology*, 32(1):191–241, 1981.
85. J. N. Washburne. An experimental study of various graphic, tabular, and textual methods of presenting quantitative material. *Journal of Educational Psychology*, 18(6):361–376, Sept. 1927.
86. J. C. Wright and G. L. Murphy. The Utility of Theories in Intuitive Statistics: The Robustness of Theory-Based Judgments. *Journal of Experimental Psychology: General*, 113(2):301–322, 1984.