Distributed Cognition as a Theoretical Framework for Information Visualization

Article *in* IEEE Transactions on Visualization and Computer Graphics · November 2008

DOI: 10.1109/TVCG.2008.121 · Source: PubMed

Distributed Cognition as a Theoretical Framework for Information Visualization

Zhicheng Liu, Nancy J. Nersessian, and John T. Stasko, *Member, IEEE*

Abstract—Even though information visualization (InfoVis) research has matured in recent years, it is generally acknowledged that the field still lacks supporting, encompassing theories. In this paper, we argue that the distributed cognition framework can be used to substantiate the theoretical foundation of InfoVis. We highlight fundamental assumptions and theoretical constructs of the distributed cognition approach, based on the cognitive science literature and a real life scenario. We then discuss how the distributed cognition framework can have an impact on the research directions and methodologies we take as InfoVis researchers. Our contributions are as follows. First, we highlight the view that cognition is more an emergent property of interaction than a property of the human mind. Second, we argue that a reductionist approach to study the abstract properties of isolated human minds may not be useful in informing InfoVis design. Finally we propose to make cognition an explicit research agenda, and discuss the implications on how we perform evaluation and theory building.

Index Terms—Information visualization, distributed cognition, interaction, representation, theory and methods

1 Introduction

Card, Mackinlay and Shneiderman define information visualization (InfoVis) as "the use of computer-supported, *interactive*, *visual representations* of abstract data to *amplify cognition*" ([6], emphasis added). Representation and interaction, as noted in the definition, are two important dimensions of InfoVis and have received considerable attention in InfoVis research. Understanding cognition, although fundamentally central to these dimensions and the field in general, has received far less scrutiny and is often considered as the responsibility of cognitive scientists. Although there has been some progress in cognitive science, there are no fully developed theories sufficient to explain and predict the effectiveness (or lack thereof) of interactive visual design, which InfoVis research can simply adopt. Existing design guidelines such as Shneiderman's mantra "overview first, zoom and filter, then details-on-demand" [27] have been useful in offering guidance to practitioners, but they also lack the rigor and power that formal theories or methodologies should provide [9]. As InfoVis designers, we often have only our own intuition and experience to depend on. Even though InfoVis research has matured technically in recent years, an important problem for the field remains the lack of an underlying theory or even a systematic framework for guiding design and investigation [26].

In this paper we argue that the "distributed cognition" framework which has emerged in cognitive science over the past 20 years has the potential to serve as a theoretical framwork for InfoVis. Using the distributed cognition framework (hereafter called DCog) to examine representation and interaction, the two key issues in InfoVis, we seek to establish the appropriateness and relevance of DCog for InfoVis research. We propose to adopt DCog's perspective of cognition as an emergent property of interaction among humans and artifacts, and discuss how the DCog framework could lead to new ways of thinking about and doing InfoVis research. Here we are using the notion of framework as proposed by the cognitive scientist Ryan Tweney in distinguishing between *theories*, for which the emphasis is on "hypotheses formation" and "testability," and *frameworks*, which "attempt to reconstruct a model of the world that meets criteria other than testability as such, is interestingly related to our theories of the world, and reduces the complexities of the real-world process in a way that permits anchoring the framework to the data" [33]. To this definition, we add that framework analyses have the potential to lead to new hypotheses that can be explored and tested in experimental research. We argue further that it might be necessary to make cognition a research agenda for InfoVis researchers, and this can have implications on the research questions we ask and the methodologies we adopt.

2 Distributed Cognition: An Environmental Perspective

DCog is a theoretical framework that originated largely with the work of Edwin Hutchins and colleagues at UCSD in the mid 1980s. In this section we provide a short description of the origin and the philosophy of this framework. We begin with the conventional view that tools amplify cognition. Visual tools, ranging from multiplication on pencil and paper to diagrams, do seem to amplify cognition as we can accomplish more with than without them. From this traditional perspective, artifacts are *scaffolds* for cognition, i.e. they make cognitive tasks easier or more efficient, and in some instances they provide a means of accomplishing cognitive tasks that could not be performed without the tools. However, cognition involves only the manipulation of information abstracted and represented "in the head", and artifacts serve to assist these processes.

Traditional research in cognitive science has adopted this perspective and thus takes an individual human as the unit of analysis. The fundamental assumption is that cognition is information processing inside the brain, and research is primarily based on a framework that models human information processing as a three-stage perception-cognition-action process. It is acknowledged that the environment, which encompasses the material, social and cultural dimensions of the world that humans are embedded in, provides the indispensable contents for thinking. However all important information about the environment is processed through perception and abstracted and stored in memory as representations or symbols. This assumption places the focus of cognitive science research on investigating how these symbols are stored and processed inside the brain. This approach - often called "reductionist" - largely ignores the environment, assuming that after we have a better understanding of the individual cognizer, we will be able to understand how cognition interacts with the complexity of the environment [29]. Based on these assumptions, the traditional experimental approach studies how humans perceive and store abstract information. One widely familiar result of this approach is the famous limitation on working memory of "the magical number 7 ± 2" [21] .

DCog questions such approach of studying human cognition. Tak-

- *Zhicheng Liu and John Stasko are with the School of Interactive Computing and the GVU Center at the Georgia Institute of Technology, E-mail: zcliu@cc.gatech.edu and stasko@cc.gatech.edu.*
- *Nancy Nersessian is with the School of Interactive Computing and the School of Public Policy at the Georgia Institute of Technology, E-mail: nancyn@cc.gatech.edu.*

Manuscript received 31 March 2007; accepted 1 August 2007; posted online 27 October 2007.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

ing sides with Cole and Griffin [8], Hutchins argued that the notion that tools *amplify* or *scaffold* cognition is misleading. When we remember something by writing it on a piece of paper and reading it later, our memory is not amplified. Instead we are using a different set of cognitive skills including writing, reading and interpreting. If we look only the *product* of the cognitive work, it seems that cognition is amplified since we can do things that we could not do without them. However, if we look at the *process* of accomplishing the task, none of our cognitive abilities has been amplified. Hutchins puts it this way: "tools permit us to transform difficult tasks into ones that can be done by pattern matching, by the manipulation of simple physical systems, or by mental simulations of manipulations of simple physical systems. Tools are useful precisely because the cognitive processes required to manipulate them are not the computational processes accomplished by their manipulation." ([18], p.170-171)

One might argue that amplifying cognition simply means that we can do things better with tools than without them, which is obviously true. However, as Hutchins suggests, it is important to make a fundamental distinction between *the cognitive properties of systems* composed of individuals manipulating tools and *the cognitive properties of individual minds*. The expression "amplifying cognition" has an undertone that blurs this distinction, thus it tends to give rise to the inadvertent conception that these are of the same nature, only that the former is more powerful than the latter. This can lead to the over-attribution problem where "we ascribe to individual minds in isolation the properties of systems that are actually composed of individuals manipulating systems of cultural artifacts" [18]. A subsequent mistake would be to isolate a human from the environment while looking for cognitive properties that arise from interacting with artifacts within the human.

There are of course cognitive properties associated with the human brain and these are part of what it means to have a "cognitive system", but much of the "intelligence" that we encounter in everyday life is fundamentally embodied in the cognitive system comprising individuals and their environment. DCog is part of a wider movement within contemporary cognitive science to account for the role of the environment (social, cultural, and material) in shaping and participation in cognition. These "environmental perspectives" [22] emphasize that cognition is embodied, enculturated, situated in local interactions, and distributed or stretched across humans and artifacts. As Nersessian has argued, "[s]uch construal leads to a shift in analytical approach from regarding cognitive and cultural factors as independent variables to regarding cognitive and cultural processes as integral to one another" ([22], p.126). Thus, DCog proposes that ascribing "cognition" to mental activities alone is misplaced. The environment provides *affordances* as well as *constraints* that essentially shape human cognition and behavior: "[culture, context and history] are fundamental aspects of human cognition and cannot be comfortably integrated into a perspective that privileges abstract properties of isolated individual minds" ([18], p.354). From this environmental perspective, cognition is more *an emergent property of interactions* between an individual and the environment through perception and action rather than a property bounded inside an individual. In the rest of this paper, the term "cognition" will be used with this connotation.

In studying cognition, we cannot thus separate the environment from the human, neither can we separate perception and action from cognitive processes, as with the traditional view. The unit of analysis should not be a human individual, but a *cognitive system* composed of individuals and the artifacts they use to accomplish a task. Consequently, purely laboratory experiment-oriented research methodology might no longer be appropriate in many cases of investigation. Rather, to understand the properties and processes of a cognitive system, a first step is often to provide descriptions of cognitive tasks through ethnographic field study in a real problem setting.

3 Steps Toward a Theoretical Framework

We recognize that the above account risks giving superficial treatment to both traditional cognitive science and DCog. Of necessity, the contrast between traditional cognitive science and DCog provided here is over-simplified. In particular, we wish to emphasize that the contrast is not primarily methodological: it is not a distinction between experimental laboratory work and observational field work - both of which are needed, as we will discuss in later sections. DCog does not reject experimental methods, the distinction is not one between quantitative and qualitative methods. The contrast is essentially perspectival: it is about a shift of outlook from an internal symbol processing view of cognition to a situated view that recognizes cognition as a property of interaction.

In addition, we are not claiming that DCog as a theory is more "correct" or "true" than traditional cognitive science theory. Both are still very much under development. In scientific fields like physics, for example, theories are best understood as providing interpretations, explanations or models of natural phenomena, and many of the existing theories have been adequately tested and show consistency. However, this does not mean that they are not subject to revision or even replacement in the future, as the history of science well documents. Stephen Hawking, in his book *A Brief History of Time*, points out that "any physical theory is always provisional, in the sense that it is only a hypothesis; you can never [conclusively] prove it. No matter how many times the results of experiments agree with some theory, you can never be sure that the next time the result will not contradict the theory" [15]. The same can be said about theories in cognitive science. Every theoretical perspective makes philosophical assumptions upon which empirical work is built. To understand the role of what are considered "theories" in a design discipline, Bederson and Shneiderman [4] summarize five ways these can help practitioners and researchers:

- *descriptive*: identify key concepts and provide a conceptual framework
- *explanatory*: rhetorically support explaining relationships and processes to support education and training
- *predictive*: make predictions about performance in existing and new situations
- *prescriptive*: provide guidelines and warnings for design
- *generative*: facilitate creativity and discovery in future research

For a design discipline, it is often argued that we take a more pragmatic attitude toward theories: the value of any theory is not "whether the theory provides an 'objective representation' of reality, but whether the particular practice in question can be informed ... by using the general propositions of the theory" [3]. In other words, "theories are judged not by their claims to truth, but by their ability to do work in the world" [2]. Our claim is primarily that DCog provides greater descriptive and explanatory power for understanding the role of artifacts and in this context, specifically InfoVis in cognitive processes. Thus, this paper neither tries to invalidate or falsify traditional cognitivism notion, nor does it advocate the validity of DCog as a theory. We do not want here to join that part of the debate between these two perspectives. Rather, our main argument is that current understanding of how human perception and cognition work under the traditional model has, thus far, been inadequate in informing InfoVis as a design discipline, and so we need to examine what the DCog perspective might be able to contribute to InfoVis.

4 The Distributed Cognition Framework

4.1 An Overview

From his field study on the behavior of crews involved in ship navigation, Hutchins conceptualizes some theoretical constructs that are applicable to studying human cognition in other domains. His account focuses on the distribution of cognitive processes across internal (within the bounds of the human body) and external structures and across members of a social group within a problem-solving system. In particular, he examines the *propagation of information* as *representation states* across a series of *representational media* that are brought into *coordination* with one another. The propagation of representation states is accomplished by both human and artifact components.

To account for the social and cultural dimensions of the environment, DCog examines members of a team and the artifacts they use as a cognitive system. The focus is on the interdependency and coordination between people who use artifacts together. The cognitive properties of such a system depend more on the social organization in the team than the cognitive properties of individual members.

4.2 A Scenario

To illustrate these concepts better and to explain how DCog can be of theoretical value to InfoVis, we present a real life scenario here. In the next two sections we then will use theoretical concepts from DCog to examine two fundamental aspects of InfoVis, representation and interaction, in the context of this scenario.

In October 2007, as the winner of the Visual Analytics Science and Technology (VAST) Symposium contest, we attended a live session to work with a professional intelligence analyst using a visual analytic system we developed called Jigsaw [30]. The analyst was given a fabricated dataset consisting of around 400 news stories and a few web pages and images. There were two embedded story threads in the dataset, one having to do with endangered species and the other having to do with threats to US citizens. The analyst's goal was to develop reasonably specific hypotheses and evidence collections for threats to endangered species with the help of Jigsaw.

Since we were only given two hours and the analyst was not very familiar with Jigsaw's interface and functionality, we conducted the visual analysis collaboratively. The analyst took control the analysis process, specifying the information to be visualized and retrieved, and a member of our team helped to realize his commands by interacting with the system.

Due to the limited length of this paper, it is not possible to provide a detailed account of the two-hour long visual analysis process, hence only a vignette is developed as follows:

> (Two major visualizations of Jigsaw, the Graph View and the Document View, were placed on the two monitors in front of the analyst and the developer.)
> Analyst: *(Jotted "butterfly" in his notepad after reading a report mentioning that children in a Swiss valley are getting paid for killing off butterflies)* "Show me all the reports related to butterflies"
> *(The developer typed the keyword "butterfly" as the keyword in the search field in the main window, which was placed together with the Document View on the left monitor, and pressed RETURN to issue a query)*
> Analyst: *(looking at the Graph View on the right monitor where reports mentioning butterfly are now shown as white nodes)* "Can I see the entities mentioned in these reports?"
> *(The developer selected all the report nodes and issued an EXPAND command to make all the entities in these reports visible as nodes connected to their respective reports in a semantic graph. The entity nodes are colored according to the their types (place, person, organization or date). The entity and report nodes are largely randomly drawn on the screen, with edges that link them intersecting each other.)*
> Analyst: "Oh, there are so many of them, which ones are important"?
> *(The developer clicked on the "circular layout" button to re-layout the graph. In this visualization, report nodes are distributed on the circumference of an invisible circle, entities connected to multiple reports are placed toward the center of the circle while entities appearing in only one report are drawn around the periphery (See [12]))*
> Analyst: (*seeing that Osaka and Texas connect to multiple reports and thus appear near the center of the circle*) "Can I see the reports mentioning Texas"?
> *(The developer selected Texas and right clicked to bring up a pop-up menu in which he selected "add to other views"; the Document View on the left monitor, initially blank, displays relevant reports in tabs)*
> Analyst: *(finding nothing interesting after reading through the reports)* "Try Osaka instead."
> *(The developer selected Osaka in the Graph View on the left monitor and repeated the action, the Document View on the left monitor refreshes and displays reports mentioning Osaka in tabs. It turns out that a Japanese man from Osaka has been smuggling endangered species of butterfly into the United States and attempted to sell them to collectors. These species of butterfly are protected by the Convention on International Trade in Endangered Species of Wild Fauna and Flora and the federal Endangered Species Act. A raid on this man's home reveals that he has been rearing these butterflies in a studio.)*
> Analyst: "Hmm, that Osaka-based smuggler seems interesting" *(jotting "Osaka smuggler" in his notepad).*

5 Representation: Internal and External

The above account is far from an ethnographic analysis. Our intention is to use it to illustrate the concepts of representation and interaction in the DCog framework. The role of representation in cognitive tasks has been extensively studied in cognitive science. Herbert Simon used mathematical theorem proving as an example to illustrate that proving a theorem is akin to making a sequence of rule applications to re-represent the axioms until they become the target proposition. He then argued that all mathematics can be viewed as changes in representation that make explicit that which is previously implicit and he extended this view to all cases in problem-solving: "solving a problem simply means representing it so as to make the solution transparent" ([29], p.133). Traditional cognitive science and DCog both subscribe to this view, hence given a cognitive task, both try to identify representational structures and explain how these representations are manipulated. Yet while traditional cognitive science solely focuses on finding internal representations and abstract rules inside the brain, DCog would argue that you get a better explanation of the phenomena by bounding the unit of analysis differently to encompass a cognitive system composed of individuals and the artifacts they use. As a result, DCog puts an emphasis on observable representations that are external to the human mind but within a cognitive system.

5.1 A System Level Account: External Representations

If we adopt the DCog perspective to look at the scenario, the analyst, the developer, Jigsaw and its related software/hardware computing infrastructure(e.g. CPU, RAM, OS), two LCD monitors, a mouse, a pen and a notepad are all parts of a cognitive system that accomplished an information foraging task. Representational media, such as the screens and the notepad, can be seen as internal to the system, and the cognitive activity of information foraging carried out on the media are processes internal to the system. Because the cognitive activity is distributed across multiple people and artifacts, many of these representations and processes are directly observable. Figure 1 summarizes the sequences of change in external representation states on representation media.

In between these stages there may be intermediate observable representations of information, for example, when the information "butterfly" is propagated from analyst's notepad to the search field on the screen, there is a verbal representation of the same piece of information in the analyst's instruction to the developer. Such intermediate representations can take various other forms such as gesture and pointing. In any case, the subsequent propagation of the digital text representation of "butterfly" to the screen is done so that next steps of propagation can continue.

5.2 An Individual Level Account: Internal Representations

As yet, the above account says nothing about any cognitive processes in the human components of the cognitive system. The change in external representation states of course does not happen automatically, but by identifying observable external representations in a cognitive system, we can have an indication of when and where internal representations are being processed so that we can be more directed in exploring them.

Information	Representation	Media
butterfly	written text	notepad
butterfly	text in search field	left screen
reports containing butterflies	white report nodes	right screen
reports containing butterflies and all the entities in the reports	semantic graph of white report nodes and colored entity nodes	right screen
Osaka and Texas and peripheral information	semantic graph of white report nodes and colored entity nodes, with important entities nodes shown more prominently	right screen
reports containing Texas	textual displays	left screen
reports containing Osaka	textual displays	left screen
Osaka smuggler	written text	notepad

Fig. 1. Propagation of Observable Representations across Observable Media

While the analyst was reading the documents containing "Texas", internal representations such as frame, schemata, mental models, propositions, lexical models and mental images are activated and created in memory and are used to make sense of the novel writings displayed on the screen. More data, such as in-depth observations and interviews would be necessary to determine how the analyst reached the conclusion that documents mentioning "Texas" might not be worth pursuing. Nevertheless, it is reasonable to assume that the processing of internal representations are tightly coupled with the processing of external textual representations on the screen.

5.3 Distributed Representations

Just how external and internal representations are interrelated and how they determine the cognitive properties of a cognitive system remain open problems for DCog. Zhang and Norman [39, 40] have approached this problem by studying some classical problem-solving tasks, such as the Tower of Hanoi problem, in which different representations of the same problem, called "problem isomorphs" ("orange", "donut" and "coffee", Figure 2), are implemented. These delimit the same problem space but have different distributions across internal and external media. A standard representation of the Towers of Hanoi problem is usually the donut version. The task is to move the three disks from one configuration to another, following three rules: 1) only one disk can be transferred at a time; 2) a disk can only be transferred to a pole on which it will be the largest; 3) only the largest disk on a pole can be transferred to another pole.

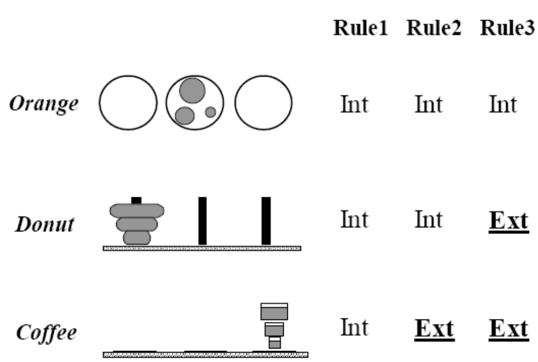

Fig. 2. Isomporphs of The Tower of Hanoi Problem

In the donut version, rule 3 is external in the sense that the physical constraints in the representation guarantee that the rule is followed: a smaller donut cannot be moved without moving a larger one on top of it first. In the coffee version, cups filled with coffee replace the donuts. In addition to rule 3, rule 2 also becomes external as a smaller cup cannot be placed on top of a larger cup without the coffee spilling out of the cup. In the orange version, no rules are incorporated into the physical constraints since one can freely move any oranges from one plate to another, and so all the rules have to be represented in human memory.

The results of Zhang and Norman's experiments show that the externalization of information is associated with significantly shorter task completion time and lower error rates. In summarizing their interpretation of the experimental results, they argue that external representations serve as more than inputs to the mind or memory aids. Differently distributed representations evoke different situated cognitive strategies, and it is the situated strategies rather than abstract mental plans that determine cognitive behavior. Citing Gibson's ecological approach to visual perception which argues that much information in the environment as invariants can be directly picked up without the mediation of any internal representations [11], they hypothesize that perceptual processing of external information is more efficient than processing of internally represented information. Distributed cognitive tasks are thus accomplished by means of an interplay between perceptual processes acting with external representations and processes acting with internal representations. The traditional cognitive approach, which separates perception and action from cognition, "postulates complex internal representations to account for the complexity of behavior" and "leads to unnecessarily complex accounts of cognition" for everyday tasks [40].

6 Interaction: Propagation by Coordination

Interaction is an essential part of InfoVis, but we do not have a clear understanding of the nature of interaction or of how to analyze it fully. Previous work has focused on interaction techniques that users can perform on visualization systems such as selecting, zooming and panning. However, interaction in a cognitive system is not a one-way process from humans to InfoVis systems. As we have seen, external and internal representations reciprocally interact and together they form a tightly coupled system. It has become recognized that we cannot overlook the effect of interaction on humans, and attention needs to be given to what users achieve by using interaction techniques rather than how the techniques provided by InfoVis systems work [37].

As discussed in the previous section, a cognitive system accomplishes a task by transforming and propagating representations across representational medias. The transformations of external representations are directly observable, but we cannot understand how they are transformed without resorting also to internal representations. DCog has not yet tried to describe the mechanisms that implement any internal representations as this is beyond the reach of all cognitive science approaches at present. Instead, DCog claims that in interacting with artifacts, the individuals in a cognitive system, as cognitive agents, provide internal representations that are required to coordinate the observable representations in solving the problem.

What does coordination mean? What are the general principles of coordination? What is involved in coordinating representations and constraints? These are central questions that DCog needs to address, as Kirsh has stated: "the study of distributed cognition is very substantially the study of the variety and subtlety of coordination" [19].

Questions about coordination are not part of traditional cognitive science because external representations are considered simply as stimuli that are taken in as input, and the reasoning process is conceived as, for instance, a process of applying condition-action rules to propositional representations of a situation. This conception places the focus of research on discovering a series of transformations of representations that can serve a path from a beginning state to a goal state. Kirsh and Maglio call this path a sequence of "pragmatic actions", which are actions whose primary function is to bring the agent closer to its physical goal [20].

In their study of how people play the Tetris game, however, they discovered that players performed a variety of actions that were not of immediate pragmatic value. Players often have little time to determine where to place a Tetrazoid, yet they rotate it multiple times more than necessary. The extra rotation appears to play a functional role in the player's computation of the goal placement, and it is distinct from pragmatic actions that advance them toward a goal state. Kirsh and Maglio call this kind of action an "epistemic action" [20] (Figure 3). Epistemic actions enable humans to make use of environmental structures or to create structures in the environment that link with internal structures so that a coordinated interplay between internal and external representations is achieved. Hence the purpose of taking some actions is not for the effect they have on the environment but for the effect they have on the humans. Coordination, as a result, is not the discovery of an external *a priori* rule or schema but an emergent property, achieved by moment-by-moment pragmatic actions and epistemic actions.

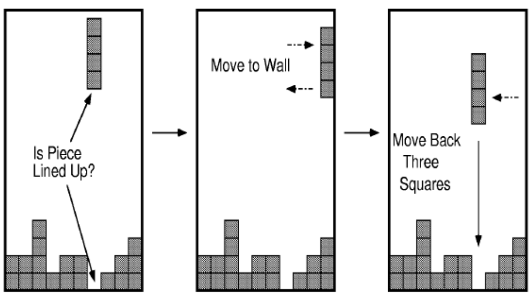

Fig. 3. Translation of a Tetrazoid as an Epistemic Action

7 Visualization at a Social Level

So far we have only been examining DCog at the individual level, focusing on the interplay between internal and external representational structures. However, to say DCog can be applied to tasks involving a single user does not mean a single user and an InfoVis system would make a complete cognitive system. In real life situations a cognitive system might actually include more people and artifacts that before were considered irrelevant. This view is supported by our contest experience described in section 4.2. First, although there was a clear division of labor in which the analyst drove the analysis process and the developer translated his intention into interaction with Jigsaw, the developer participated in the analysis from an analyst's perspective as well. When the analyst had explored all potential suspects and still did not have a viable lead, the developer suggested to look at the amount of money involved in some of the transactions. In addition, the analyst made several comments during the process such as, "This would be when I pick up the phone and ask a financial expert" or "I would now send an investigator to collect more information". From the DCog perspective, all analysis tasks are inherently social and embodied in context, even if they are the main responsibility of just one person. This can mean that the standard one-person-one-computer InfoVis model is over-simplified and by detaching the analysis task from its embodied context could lead to negative consequences.

Recent work on social visualization can be seen as an effort to expand the boundaries of individual-computer cognitive system. Visual exploration of data, for example, often does not end as an individual effort. Previous research has shown that users were eager to share visualizations with their friends through storytelling on applications that were intended for single users only [34]. With the understanding of the social nature of visual data analysis, projects such as Many Eyes [35] look at a larger cognitive system that includes the user and friends and potential others who will be interested in the data. In such a cognitive system, the propagation of representation states is greatly enhanced by the WWW infrastructure.

Here we can see again that interaction and coordination are important research issues for social visualization. What are the examples of coordinative mechanisms and processes between people mediated by artifacts? What are the roles of representations in the process? Answering these questions will help us design better systems to support social visual exploration of data.

8 DCog as a Theoretical Framework

The idea of using DCog as a framework for investigating human-computer systems will not be a new idea to many HCI and CSCW researchers, where attempts have been made to use DCog to inform research [17, 14]. However, we believe that it is still worthwhile to raise awareness about important assumptions and theoretical constructs of DCog in a way specific to the InfoVis community and in the context of our effort to search for a theoretical framework to guide research. Furthermore, we argue that DCog can shed light on future research directions and methodologies for InfoVis researchers.

8.1 In the Context of Recent Research Development

Identifying an underlying theoretical framework has been a research topic in the InfoVis community for several years [9, 26]. Among these efforts, linguistic theory, data-centric predictive theory, information theory and scientific modeling have been considered as potential candidates, but DCog has not yet been considered an important theoretical framework that deals with the central issue of cognition. Arguably the field of InfoVis has already accepted an "external cognition" approach, as represented by the foundational work by Card et al [6], and by discussions of visual representations by Don Norman [24]. These works do highlight the importance of external representations in accomplishing cognitive tasks, and they quite likely can be subsumed under the DCog framework. However, important theoretical constructs of DCog, such as the claims that cognition is an emergent property of interaction and interaction involves coordination and cognitive coupling, are missing from these works.

In light of these considerations, we think it is useful to make the theoretical constructs about cognition and interaction more explicit. The external cognition approach, for example, has placed much emphasis on representations, but interaction has played a lesser role, although it is acknowledged as being important. As Yi et al. pointed out, "without interaction, an InfoVis technique or system becomes a static image or autonomously animated images" [37]. Interaction makes InfoVis a powerful tool, but understanding the nature and mechanisms of interaction is difficult. DCog's perspective that interaction is the propagation of representation states in a cognitive system through coordination is helpful. Kirsh's work on epistemic action and pragmatic action is valuable in leading us in further exploration along this direction.

Furthermore, there is a growing interest in the topic of insight in the InfoVis community. It is widely acknowledged that "the purpose of visualization is insight, not pictures. The main goals of this insight are discovery, decision making, and explanation" [6]. In trying to understand how effectively our designs support this claim, InfoVis researchers are beginning to ask what insight is, how to measure insight and how insight is generated [25, 38]. Questions about insight are intrinsically about the nature of human cognition, and we believe a wider use of the DCog framework can be beneficial to research works on insight. A central problem in characterizing insight is the nature and processes of "conceptual change". Insights are often generated when people's conceptions undergo a transformation, and we usually construe such a transformation as a sense-making or reasoning process. Recent work on conceptual change advocates the role of internal/external representational coupling in creating insight and conceptual changes [23].

8.2 Framework, not Theory

As we have noted, DCog is a framework under development and cannot yet provide full explanations of the general principles of cognitive coupling, coordination, externalization or interaction, and so it cannot be simply applied to issues in InfoVis. Scientific theory building is a bootstrapping process, and so research in InfoVis from a DCog perspective can itself contribute to the development of the cognitive theory. At present DCog is more a theoretical framework in the sense that it provides a set of fundamental assumptions about the subject matter and points out important research questions to be investigated. The works on distributed representation and epistemic action, as we have mentioned in Sections 5 and 6, are research based on the fundamental assumptions made in the DCog framework. What DCog can provide is a way of developing detailed analysis of InfoVis as it participated in cognitive systems, which can provide insights on how to improve a design [14]. DCog also does not address all issues of importance to InfoVis, for instance the nature of visual perception. Here InfoVis has drawn valuable insights from various sources such as traditional cognitive science research on preattentive processing [16] and Tufte's works on information design [32].

In sum, we are advocating that DCog be seen as one theoretical framework guiding InfoVis research, most particularly on the issues of representation and interaction. DCog, as Halverson put it, provides one pair of colored glasses: "we put them on and the world is tinted. The change brings some objects into sharper contrast, while others fade into obscurity." [14] DCog can assist InfoVis in sharpening its objects and methods. In the next section, we look at current approaches in evaluation and theory building and discuss potential new ways of thinking and doing research based on the DCog framework.

9 Reflections on Current Practices in InfoVis

9.1 Evaluation

The main purpose of evaluation in traditional HCI design is to measure a system's usability and utility as ways of suggesting potential design problems to be corrected. As Greenberg and Buxton argue, good usability in successful products often happens after, not before, usefulness [13]. Thus it might not make much sense to do usability evaluation for novel techniques or systems, rather the focus of evaluation should be on usefulness instead. Yet determining usefulness of new designs is hard and we are still not sure how to best do that. In many traditional experimental methods, some control variables are identified and human subjects are brought in as operators of a system. The goal is to measure and compare system performances by varying control variables. Such an approach is arguably influenced by scientific methods in natural sciences. However we are often left perplexed after these experiments, not able to explain the results obtained. Furthermore, it is not impossible that doing these experiments in another lab with another group of participants could produce distinctively different results. Many have come to realize that using scientific methods does not necessarily mean we are doing good science [13].

One major criticism of laboratory-based experimental evaluation is that it ignores the situated nature of actual user tasks. As a result, there have been efforts to move away from short-term experimental evaluation to longer-term *in situ* observational evaluation [28]. However, it is possible that the crux of the problem of the experimental approach lies neither in the setting nor the methodology *per se*, but our fundamental assumptions about the nature of cognition. It is true that most human cognitive activities are inherently social and situated, yet in the end we still act locally. Considering that many tasks that InfoVis systems try to support are accomplished by single individuals, it is easy to form a cognitive system comprising a human and an InfoVis system. There is thus no reason why we cannot do controlled experimental evaluation in a laboratory setting. Yet before we design experiments, we need to realize that it might not be so easy to pin down the "control variables" and "dependent variables". Many traditional experimental methods can be described as treating an external system as some variable input to be fed into a human cognitive black-box and comparing the "output" produced without much analysis of the distribution of internal and external representations to which DCog calls our attention. Although some attempts have been made to complement the experimental approach by exploring what the user is thinking based on think-aloud protocols or interviews, as far as we have been able to determine, there has been no serious analysis of how a user's internal model interacts with external representations.

Long-term in-depth evaluation in a naturalistic setting presents a different set of problems. We can gather convincing evidence to show that InfoVis systems are indeed useful in actual work practices and document how we improve a design by working closely with actual users. However, the question arises how do we transfer what is learned from doing *in situ* evaluation to inform future design efforts?

Furthermore, with the goal of validating and improving a design in mind, evaluation is too often used as an existence proof for a system or technique, and in fact it would be surprising if the researcher could not come up with a single scenario where the new system or technique will be somehow better [13]. Evaluation results do not seem to play a significant role in our ongoing research efforts. How then, is evaluation of any sort contributing to the development of the field of InfoVis?

9.1.1 Cognition as a Research Agenda: Towards a Science of Interaction

Evaluation can and must play a more important and justifiable role in the development of the field of InfoVis. While validation and presentation of the utility of InfoVis systems are important, we propose that another primary goal of doing evaluation is to understand the nature of cognitive processes that can be carried out by a system comprising humans and InfoVis. A naive partitioning between InfoVis and cognitive science might be a contributing factor in preventing InfoVis research from reaching its full potential. Thus we advocate that research on cognition become part of the agenda of InfoVis.

Based on DCog's fundamental assumption that cognition is an emergent property of interaction, interaction, then, is a major focus in the study of cognition. Recently a call has been made for creating a new science of interaction for the field of visual analytics, which shares fundamental similarities with InfoVis [31]. But what exactly does a "science of interaction" mean and how to achieve it is in need of clarification. DCog as a theoretical framework can guide us in this endeavor toward a science of interaction. For instance, humans have limited attention span and mental resources. When a task is too complex and demanding, we may simply give up or procrastinate. However, we are also highly reflective, creative agents with the ability to adapt the environment instead of adapting ourselves for cognitive tasks. The flexibility and limitations of human cognitive resources are important facts that have not been emphasized in InfoVis work on interaction. Formalism based abstractions on task environments or task procedures are the major approaches for studying how people interact with artifacts and other people, yet these are known to be flawed in making unreasonable assumptions about the predictability of human behaviors and characterizing the environment as a fixed set of choice points [19]. How do we then proceed in a scientific study of interaction without resorting to these inadequate assumptions?

As we have discussed, in the DCog framework, it is through interaction that a cognitive system can possess cognitive properties to compute a task. Interaction is the propagation of representations across a series of representational media that are brought into coordination inside a cognitive system. Coordination and cognitive coupling are thus the central issues to be addressed. To reach its full potential, a science of interaction should not be just a taxonomy of interaction techniques or a framework of the abstracted task procedures; it should be a scientific approach to understand *how cognition emerges as a property of interaction between external and internal representations*. The research problem then, is not to discover, for example, the assumed condition-action schemata that govern human behavior. Instead we need to ask research questions such as the following:

- What are the nature and mechanisms of coordination and cognitive coupling?
- How do people develop interaction strategies during sensemaking and analytical reasoning?
- How does interaction with visual structures enable turning information into meaningful understanding?

9.1.2 Methodological Issues

What are the methods we can use to investigate such questions? A systematic and appropriate methodology is needed. A "divide-and-conquer" methodology that first breaks down a complex system into examinable parts and then puts the parts back together forms the basis of many well-developed areas of "hard-science". The great success it has enjoyed in these areas makes it a tempting choice. However, in the scenario described earlier, the cognitive abilities required to perform the information foraging task were found neither in the human individuals nor in Jigsaw alone. The cognitive abilities belong to the entire system, but they cannot be reduced to any part of the system - this is what "emergence" means. Through the interaction between the humans and Jigsaw, insights are generated and the cognition emerges from the entire system that is *larger than the sum of its parts*. Thus, approaches that try to break down the system into visualizations, humans and interaction techniques cannot give us a satisfactory account of how InfoVis helps us accomplish tasks.

The challenge, then, is to take the cognitive system as a unit of analysis and account for the mechanisms through which tight coupling or coordination between internal and external representations is achieved. And the objective is to analyze how meaning and understanding are constructed from interacting with external representations. "Cognitive ethnography" [17, 18] is the approach used by DCog to answer such questions, and it can be adapted for use in InfoVis research. Ethnography as a methodology has its roots in anthropology, where meaning, cultural model and perhaps intersubjectivity are among the primary subject matters of investigation. DCog has developed *cognitive ethnography* to focus on the role of interaction in cognitive activities [17]. In situ fieldwork, interviews, observations and video-taping are primary methods associated with ethnographic work. But many have pointed out that ethnography in its essence is an analytic strategy for assembling and interpreting the data obtained from fieldwork [1, 10]. By identifying various forms of coordinative mechanisms and the natural evolution of coordinative strategies, cognitive ethnography attends to how these interactions make sensemaking and meaning construction possible. This implies that we may want to, for example, explain how invented visual representations come to possess meaning and constitute shared cultural models. Visual cultural models then become an integral part of InfoVis research. Special attention needs to be paid to how these visual cultural models become externalized and internalized through interaction occurring both during the time of task completion as well as in past learning experiences.

Using an ethnography-based methodology does not require rejecting controlled experiments. As Hollan et al. have argued: "Sometimes the interpretations generated by the cognitive ethnography are underconstrained by the available observations. In these cases, the findings of a cognitive ethnography may suggest experiments that can resolve issues that cannot be resolved by observation alone. In this way, cognitive ethnography guides the design of experiments which contribute to the refinement of the distributed cognition theory." [17]

9.1.3 Theory Building

Most research that addresses theoretical issues in InfoVis has involved building taxonomies of visual representations, interaction techniques or user tasks. Taxonomy building is an empirically-driven strategy and can be viewed as a "bottom up" approach. Without an overarching theoretical framework, however, "bottom up" efforts risk being overwhelmed with details and not looking in the right places that must be addressed in theory developing. DCog as a theoretical framework can complement the "bottom up" approach with a "top down" one. That is, starting with an interpretive framework highlighting important issues that are brought "into sharper contrast", "bottom up" efforts become more directed and focused.

Theory building, of course, is rooted in empirical research. But how do we determine what aspects of empirical observation must be addressed when we are building taxonomies? Many existing taxonomizing efforts have focused on single aspects such as data [5], task-by-data-type [27] or visualization [7]. The DCog framework adds to this an integrative approach that addresses the interplay between internal representation and external representation. Such an approach expands the range of empirical observations relevant to developing taxonomic vocabularies. Evaluation might provide a good venue where empirical data can be collected to build such a taxonomy.

9.2 Beyond Evaluating InfoVis Systems

It is generally acknowledged that InfoVis systems are useful for reducing cognitive load and facilitating knowledge crystallization tasks. A quick scan of recent InfoVis and visual analytics papers will find case studies of the claimed benefits of deploying a system for a specific problem domain. These case studies often illustrate scenarios that give an account of how interactive visual representations generate insights. While these accounts focus on the positive effects, it is possible that we have overlooked potential negative effects brought about by InfoVis systems.

Hollan et al. point out that design of new digital artifacts risks destroying many of the most valuable aspects of current ways of doing things which we do not yet fully understand [17]. The process of interaction, as described in many InfoVis case studies, resembles a series of "pragmatic actions" which advance towards the end state where insights are revealed. InfoVis system design, in a way, can be seen as the effort to facilitate epistemic actions that turn hidden information into explicit visual representations. However in doing so, we might inadvertently construct an interaction space that eliminates the possibility of performing other important epistemic actions not supported by the system. Might it be the case that some InfoVis systems inhibit users from developing creative coordination strategies when performing a task? Furthermore, can digital interfaces provide affordances upon which users perform a range of epistemic actions? If not, how do we enable the interoperability between digital interfaces and other analog media so that computers do not create an artificial barrier that separates possible user actions into two disjoint kinds?

Answering such questions requires going beyond studying how InfoVis systems can help in a task. It is also important to study any possible adaptations or appropriations users make in their interactions with the systems. Many of the current InfoVis systems, however, are not flexible enough to support user customization and appropriation. So the DCog framework can push us to resolve outstanding technical questions. Furthermore, in thinking about the possible InfoVis design space, the study of cognition needs also focus on current work practices that might be amenable to being supported by InfoVis systems, with an emphasis on representations of information and the coordinating constraints that organize representations in a cognitive system.

10 Conclusion

A mature field is often supported by theories that possess descriptive, explanatory, predictive, prescriptive and generative powers. These theories, however, are seldom developed within a short time span. Physical theories in ancient Greek times were primarily descriptive with very limited predictive and prescriptive powers. Aristotle's "the five elements" and "the four causes" were thought to be the important theoretical constructs to be studied. It took nearly two thousand years for physicists to identify crucial concepts that should be the foci of study and the appropriate methods to study them. Before the birth of classical mechanics, many physical theories were domain specific and only addressing a limited set of phenomena such as tides (Galileo's theory of tides) and planetary motion (Kepler's Laws). Only when adequate understandings of specific phenomena were achieved could Newton propose generic theories on motion that possess great explanatory and predictive powers.

We believe that there are lessons to be learned from the history of science. We are still in the early stage of developing an understanding of human cognition. A first step thus is naturally a descriptive enterprise. Identifying important concepts to study and appropriate methods to study them is a prerequisite that guides the descriptive approach, and it might be that we need to start from examining cognition in specfic domains. But that does not mean that what we find is only about that specific domain. As Hutchins has remarked: "There are powerful regularities to be described at the level of analysis that transcends the details of the specific domain. It is not possible to discover these regularities without understanding the details of the domain, but the regularities are not about the domain specific details, they are about the nature of cognition in human activity." (as quoted in [36])

InfoVis is a research area that is closely related to the study of human perception and cognition. Successful InfoVis theories cannot ignore the importance of human cognition to understand InfoVis. To successfully incorporate the human component into InfoVis theories, we need a framework that can identify important theoretical concepts and suggest appropriate methodology for this endeavor. We have been arguing that DCog provides a more useful framework to address the central issues of representation and interaction than does the traditional cognitive science framework. It provides the kinds of interpretive concepts and methods needed to analyze the role of InfoVis in complex cognitive tasks. While it has limitations in providing prescription and prediction, it is useful in helping us understand phenomena that are important for InfoVis, and our designs can be informed with this understanding. And reflexively, understandings of cognition deriving from InfoVis research can feed back into the development of a general science of cognition.

Acknowledgements

We thank the reviewers for their insightful comments and criticisms on an earlier version of this paper. This research is supported by the National Science Foundation via Award IIS-0414667 and the National Visualization and Analytics Center (NVAC^{TM}), a U.S. Department of Homeland Security Program, under the auspices of the Southeast Regional Visualization and Analytics Center.

References

[1] B. Anderson. Work, ethnography and system design. In A. Kent and J. G. Williams, editors, *The Encyclopedia of Microcomputers*, volume 20, pages 159–183. Marcel Dekker, New York, 1997.

[2] S. Barab and K. Squire. Design-based research: Putting a stake in the ground. *Design-Based Research: Clarifying the Terms*, 13(1):1–14, 2004.

[3] J. Bardram. Designing for the dynamics of cooperative work activities. In *Conference on Computer-Supported Cooperative Work*, Seattle, WA, 1998. ACM.

[4] B. B. Bederson and B. Shneiderman. Theories for understanding information visualization. In *The Craft of Information Visualization: Readings and Reflections*, chapter 8, pages 349–351. Morgan Kaufmann, 2003.

[5] J. Bertin. *Graphics and Graphic Information Processing*. Walter de Gruyter, Berlin, 1981.

[6] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in information visualization: using vision to think*, chapter 1, pages 1–34. Morgan Kaufmann Publishers Inc, 25 January 1999.

[7] E. Chi. A taxonomy of visualization techniques using the data state reference model. In *Proceedings of InfoVis '00*, Salt Lake City, UT, 2000.

[8] M. Cole and P. Griffin. Cultural amplifiers reconsidered. *The social foundations of language and thought*, pages 343–364, 1980.

[9] B. Craft and P. Cairns. Beyond guidelines: What can we learn from the visual information seeking mantra? In *Proceedings of the Ninth International Conference on Information Visualization (IV'05)*, pages 110–118. IEEE Computer Society, 2005.

[10] P. Dourish. Implications for design. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 541–550, 2006.

[11] J. Gibson. *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates Inc, US, 1986.

[12] C. Görg, Z. Liu, N. Parekh, K. Singhal, and J. Stasko. Jigsaw meets Blue Iguanodon - The VAST 2007 Contest. In *Proceedings of IEEE VAST '07*, pages 235–236, Sacramento, CA, October 2007.

[13] S. Greenberg and B. Buxton. Usability evaluation considered harmful (sometimes). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Florence, Italy, 2008.

[14] C. A. Halverson. Activity Theory and Distributed Cognition: or What does CSCW need to do with Theories? *Computer Supported Cooperative Work*, 11:243–267, 2002.

[15] S. Hawking. *A Brief History of Time*. Bantam, 10 anv edition, 1998.

[16] C. Healey, K. Booth, and J. Enns. High-speed visual estimation using preattentive processing. *ACM Transactions on Computer Human Interaction*, 3(2):107–135, 1996.

[17] J. D. Hollan, E. Hutchins, and D. Kirsh. Distributed cognition: Toward a new foundation for human-computer interaction. *ACM Transactions on Computer-Human Interaction*, 7(2), 2000.

[18] E. Hutchins. *Cognition in the Wild*. MIT Press, Cambridge, MA, 1994.

[19] D. Kirsh. Distributed cognition: A methodological note. *Pragmatics and Cognition*, 14(2):249–262, 2006.

[20] D. Kirsh and P. Maglio. On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4):513–549, 1994.

[21] G. A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63:81–97, 1956.

[22] N. Nersessian. The cognitive-cultural systems of the research laboratory. *Organizational Studies*, 27(1):125–145, 2006.

[23] N. J. Nersessian. Mental modelling in conceptual change. In S. Vosniadou, editor, *Handbook of Conceptual Change*. Routledge, London, 2008. In Press.

[24] D. Norman. *Things that Make Us Smart: Defending Human Attributes in the Age of the Machine*, chapter 3. Addison-Wesley, 1994.

[25] C. North. Visualization viewpoints: Toward measuring visualization insight. *IEEE Computer Graphics and Applications*, 26(3):6–9, 2006.

[26] H. C. Purchase, N. Andrienko, T. Jankun-Kelly, and M. Ward. Theoretical foundations of information visualization. In A. Kerren, J. T. Stasko, J.-D. Fekete, and C. North, editors, *Information Visualization - Human-Centered Issues and Perspectives*, chapter 3. Springer, 2008.

[27] B. Shneiderman. The eyes have it: A task by data type taxonomy of information visualizations. In *Proc. IEEE Visual Languages '96*, pages 336–343, 1996.

[28] B. Shneiderman and C. Plaisant. Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. In *BELIV : BEyond time and errors: novel evaLuation methods for Information Visualization*, Venice, Italy, 2006.

[29] H. A. Simon. *The Sciences of the Artificial*. MIT Press, 1996.

[30] J. Stasko, C. Görg, Z. Liu, and K. Singhal. Jigsaw: Supporting investigative analysis through interactive visualization. In *Proceedings of IEEE VAST '07*, Sacramento, CA, October 2007.

[31] J. J. Thomas and K. A. Cook. *Illuminating the Path: the Reseach and Development Agenda for Visual Analytics*. IEEE Computer Society, Los Alamitos, CA, 2005.

[32] E. R. Tufte. *The Visual Display of Quantitative Information, Second Edition*. Graphics Press, Cheshire, CT, 2001.

[33] R. D. Tweney. A framework for the cognitive psychology of science. In B. Gholson, W. R. S. Jr., R. A. Neimeyer, and A. C. Houts, editors, *Psychology of Science: Contributions to Metascience*, pages 342–366. Cambridge University Press, New York, 1989.

[34] F. B. Viégas, danah boyd, D. H. Nguyen, J. Potter, and J. Donath. Digital artifacts for remembering and storytelling: Posthistory and social network fragments. In *Proc. of the 37th Hawaii International Conference on System Sciences*, 2004.

[35] F. B. Viégas, M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. Many eyes: A site for visualization at internet scale. In *IEEE Information Visualization '07*, pages 1121–1128, Sacramento, CA, October 2007.

[36] D. D. Woods. Toward a theoretical base for representation design in the computer medium: Ecological perception and aiding human cognition. In J. Flack, P. Hancock, J. Cairn, and K. Vincente, editors, *Global Perspectives on the Ecology of Human-Machine Systems*, pages 157–188. Lawrence Erlbaum Associates, 1995.

[37] J. S. Yi, Y. ah Kang, J. T. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, 2007.

[38] J. S. Yi, Y. ah Kang, J. T. Stasko, and J. A. Jacko. Understanding and characterizing insights: How do people gain insights using information visualization. In *Proceedings of BELIV '08*, Florence, Italy, April, 2008.

[39] J. Zhang. The Interaction between Perceptual and Cognitive Processes in a Distributed Problem Solving Task. *Working Notes of the 1993 AAAI Fall Symposium on Games: Planning and Learning*, 1993.

[40] J. Zhang. and D. Norman. Representations in Distributed Cognitive Tasks. *Cognitive Science*, 18(1):87–122, 1994.