A Multi-Level Typology of Abstract Visualization Tasks

Matthew Brehmer and Tamara Munzner, Member, IEEE

Abstract—The considerable previous work characterizing visualization usage has focused on low-level tasks or interactions and high-level tasks, leaving a gap between them that is not addressed. This gap leads to a lack of distinction between the ends and means of a task, limiting the potential for rigorous analysis. We contribute a multi-level typology of visualization tasks to address this gap, distinguishing why and how a visualization task is performed, as well as what the task inputs and outputs are. Our typology allows complex tasks to be expressed as sequences of interdependent simpler tasks, resulting in concise and flexible descriptions for tasks of varying complexity and scope. It provides abstract rather than domain-specific descriptions of tasks, so that useful comparisons can be made between visualization systems targeted at different application domains. This descriptive power supports a level of analysis required for the generation of new designs, by guiding the translation of domain-specific problems into abstract tasks, and for the qualitative evaluation of visualization usage. We demonstrate the benefits of our approach in a detailed case study, comparing task descriptions from our typology to those derived from related work. We also discuss the similarities and differences between our typology and over two dozen extant classification systems and theoretical frameworks from the literatures of visualization, human-computer interaction, information retrieval, communications, and cartography.

Index Terms—Typology, visualization models, task and requirements analysis, qualitative evaluation

1 INTRODUCTION

Consider a person who encounters a choropleth map while reading a blog post in the aftermath of last year’s American presidential election. This particular map is static and visually encodes two attributes, can-...
2 Background Context

The primary limiting factor in using extant classification systems as tools for analysis is that we cannot easily distinguish between the ends and means of tasks. Making this distinction is a central problem for practitioners during the abstraction phase of design studies [60] and during the analysis phase of qualitative user studies [34].

For instance, a number of extant classification systems mention the word derive [2, 14, 23, 37, 50, 66]. Is derive a task, or the means by which another task is executed? A user may derive data items as an end in itself, for example to reduce the number of dimensions in a dataset, or as a means towards another end, such as to verify a hypothesis regarding the existence of clusters in a derived low-dimensional space. The ends-means ambiguity exists for many terms found in extant classification systems: consider filter [2, 11, 19, 23, 29, 31, 37, 42, 50, 51, 57, 61, 82], navigate [23, 64, 75], or record [23, 42, 66].

The first step towards distinguishing ends from means involves asking why a task is performed separately from how a task is performed, a question that is central to the organization of our typology.

The separation of why and how does not in itself resolve all confusion. Consider sort, another term appearing in extant classification systems [2, 19, 23, 37, 50]. Sorting has an input and an output; in some cases, it is items of data within a single view [55]; in others, views themselves may be sorted [8]. In both cases, the sorted output can serve as input to subsequent tasks. The next step in distinguishing ends from means is thus characterizing what the task’s inputs and outputs are, allowing us to describe sequences of interdependent tasks.

To illustrate how the ends-means ambiguity arises during the course of analysis, we will now attempt to use representative extant classification systems to describe two example tasks:

Example #1: recall the example stated in the paper’s introduction, that of a casual encounter with an electoral map in which a user compares two regions; election results for each state are encoded as a choropleth map based on two attributes, candidate and margin of victory. Furthermore, we know that this user is familiar with American geography and map reading.

While the dimensions of this description are similar to the questions of why, how, and what, the description is incomplete, particularly in its classification of goals and objectives. Roth’s taxonomy provides us only with a partial sense of how the comparison is performed: retrieve does not tell us about whether the user knows the spatial location of the regions to be compared a priori. The goal, procure, does not provide us with any higher-level context or motivation for why the user is procuring; specifically, the user’s casual interest in these two regions is lost. Finally, Roth’s taxonomy imposes a spatial constraint on operands, leaving us unable to fully articulate what is being compared.

Example #2: in evaluation studies [34], it is sometimes necessary to perform a comparative analysis of a task being performed using different visualization systems. Consider a user of a tree visualization tool whose interest relates to two nodes in a large tree, and her intent is to present the path between these nodes to her colleagues. SpaceTree [20] and TreeJuxtaposer [46] are two tree visualization tools that allow users to locate paths between nodes by means of different focus + context techniques. Both systems allow for path selection, in which the encoding of selected paths differs from that of non-selected paths. The systems differ in how elements of the visualization are manipulated: TreeJuxtaposer allows the user to arrange areas of the tree to ensure visibility for areas of interest, while SpaceTree couples the act of selection by aggregating and filtering unselected items.

As in the previous example, task descriptions from extant classification systems seldom answer all three questions: why, how, and what. Using Heer and Shneiderman’s taxonomy of interactive dynamics for visual analysis [23], we might describe this task as being an instance of data and view specification (visualize and filter) as well as view manipulation (navigate and select). This description tells us how, but it doesn’t specify why the task is performed.

We might complement Heer and Shneiderman’s description with one based on Lee et al.’s taxonomy of graph visualization tasks [37], in which this task would be classified as a topology task, namely one of determining connectivity and subsequently finding the shortest path. As the scope of Lee et al.’s taxonomy is specialized, we are provided with a clear indication of what the user’s interest is, this being a path.

Unfortunately, this description provides only a partial account of why the task is performed; we are not provided with a high-level motivation beyond determining and finding a path.

Both descriptions do not relate the user’s actions to the high-level goal of presenting information to others. Second, and more importantly, these descriptions fail to distinguish how this task is performed using SpaceTree from how it is performed using TreeJuxtaposer.

Summary: these examples demonstrate our inability to comprehensively analyze tasks using state-of-the-art systems for classifying user behaviour. Note that we are not directly criticizing these classification systems; we acknowledge that their scope is often deliberately constrained, with some focusing on low-level tasks, interactions, or operations [2, 5, 10, 12, 13, 14, 15, 19, 29, 37, 59, 61, 72, 73, 75, 78, 82, 83], while others focus on high-level tasks or goals [3, 11, 31, 39, 51], or on user behaviour in specialized domains [37, 57, 65]. We lack guidance on how to integrate these disjoint bodies of work, to compose task descriptions that draw from all of them. This integration is the aim of our typology, which will allow practitioners to describe tasks that address critical questions posed during visualization design and evaluation, namely why, how, and what.

It could be argued that a classification of tasks should focus solely on the goal of the user, or why visualization is used; users are often not immediately concerned with how a task is performed, as long as their task can be accomplished. We argue that by classifying tasks according to how they are performed, in addition to why they are performed and what they pertain to, we can improve communication between visualization practitioners working in different application areas, facilitating tool-independent comparisons, the analysis of diverging usage strategies for executing tasks [74, 84], and improved reasoning about design alternatives.

3 Typology of Abstract Visualization Tasks

Our multi-level typology of abstract visualization tasks, represented in Figure 1, is encapsulated by three questions: why the task is performed, how the task is performed, and what does the task pertain to (Figures 1a-c). Complete task descriptions, such as those for Examples #1-2 (represented in Figure 2), must include nodes from all three parts of this typology. We denote this work as a typology, rather than a taxonomy, as the former is appropriate for classifying abstract concepts, while the latter is appropriate for classifying empirically observable events [6].

This structure, while unusual relative to many extant classification systems, mirrors the analytical thinking process undertaken in design studies [41, 43, 60]. Why, what, and how are also used in Cognitive Work Analysis [74], particularly for relating abstractions within a work domain, as well as in Aigner et al.’s analysis of techniques and systems for visualizing time-oriented data [1], which asks what is presented?, why is it presented?, and how is it presented?.

We will introduce why before how, as this order reflects the translation of empirically observable domain problems into abstract tasks and subsequently into design choices: practitioners first identify why
what?

a task is to be performed, and then must decide upon how the task is to be supported. We then discuss the what part of our typology, which considers the input and output of tasks. Our typology supports the description of complex visualization usage as a sequence of interdependent tasks, where the output of a prior task may serve as the input to a subsequent task, as we demonstrate in the case study of Section 4.

For clarity, we first present our typology in its entirety with minimal discussion of the previous work that informed its organization, and then focus on these connections in Section 5 and in Table 1.

3.1 Why?

The why part of our typology, shown in Figure 1a, allows us to describe why a task is performed, and includes multiple levels of specificity, a narrowing of scope from high-level (consume vs. produce) to mid-level (search) to low-level (query).

Consume: Visualizations are used to consume information in many domain contexts. In most cases this consumption is driven by either a need to present information or to discover and analyze new information [79]. However, there are many other contexts in which visualizations are simply enjoyed [17, 54, 65], where users indulge their casual interests in a topic.

Present refers to the use of visualization for the succinct communication of information, for telling a story with data, guiding an audience through a series of cognitive operations. Presentation using visualization may take place within the context of decision making, planning, forecasting, and instructional processes [18, 40, 57]. Presentation brings to mind collaborative and pedagogical contexts, and the means by which a presentation is given may vary according to the size of the audience, whether the presentation is live or pre-recorded, and whether the audience is co-located with the presenter [32].

Discover is about the generation and verification of hypotheses, associated with modes of scientific inquiry [50]. Scientific investigation may be motivated by existing theories, models, and hypotheses, or by the serendipitous observation of unexpected phenomena [4].

Enjoy refers to casual encounters with visualization [54, 65]. In these contexts, the user is not driven by a need to verify or generate a hypothesis; novelty stimulates curiosity and thereby exploration [16, 65, 67, 69]. This motivation is notably absent from previous classification systems, as shown in Table 1. Casual encounters with visualization can be fleeting, such as in the earlier example of encountering a static choropleth electoral map while reading a blog post. Conversely, these encounters might be immersive and time-consuming experiences, such as in museum settings [54].

Produce: we use produce in reference to tasks in which the intent is to generate new artefacts. These artefacts include but are not limited to: transformed or derived data, annotations, recorded visualization interactions, or screen shots of static visualizations. Examples of produce in previous work include the production of graphical annotations and explanatory notes to describe features of time-series graphs [81], or the production of graphical histories in Tableau [21], the latter being meta-visualizations intended to improve users’ analytical provenance. Additional examples of produce involving derived data and annotations are featured in the case study of Section 4.

It is important to note that the products of a produce task may be used in some subsequent task that may or may not involve visualization. For example, some visualization tools for analyzing high-dimensional data allow users to produce new categorical attributes for labelling clustered data points in a dimensionally-reduced coordinate space; these attributes might be used later for constructing a predictive model.

Search: Regardless of whether the intent is to present, discover, or merely enjoy, the user must find elements of interest in the visualization. While terms relating to search and exploration are often conflated [40, 69], we have imposed a characterization of search that depends on what is being sought. We classify them according to whether the identity or location of the search target is known a priori. Whether the identity of the search target is known recalls Andrienko and Andrienko’s concept of references and characteristics [5]: searching for known reference targets entails lookup or locate, while searching for targets matching particular characteristics entails browse or explore. Consider our earlier example of a user who is familiar with American geography and is searching for California on an choropleth map; we would describe this as an instance of Lookup. However, a user who is unfamiliar with American geography must locate California.

In contrast, the identity of a search target might be unknown a priori; the user may be searching for characteristics rather than references [5]; these characteristics might include particular values, extremum, anomalies, trends, or ranges [2]. For instance, if a user of a tree visualization is searching within a particular subtree for leaf nodes having few siblings, we would describe this as an instance of browse because the location is known a priori. Finally, explore entails searching for characteristics without regard to their location, often beginning at an overview level of the visualization [37]. Examples include searching for outliers in a scatterplot, for anomalous spikes or periodic patterns in a line graph of time-series data, or for unanticipated spatially-dependent patterns in a choropleth map.
Query: Once a target or set of targets has been found, a user will identify, compare, or summarize these targets. If a search returns known or reference targets [5], either by lookup or locate, identify returns their characteristics. For example, a user of a choropleth map representing election results can identify the winning candidate and margin of victory for the state of California. Conversely, if a search returns targets matching particular characteristics, either by browse or explore, identify returns references. For instance, our election map user can identify the state having the highest margin of victory.

The progression from identify to compare to summarize corresponds to an increase in the amount of search targets under consideration [5, 10, 72], in that identify refers to a single target, compare refers to multiple subsets of targets, and summarize refers to a whole set of targets. As with explore, summarize is also often associated with overviews of the data [37]. Continuing with the choropleth map example, a user identifies the election results for one state. compares the election results of one state to another, or summarizes the election results across all states, determining how many favored one candidate or the other, or the overall distribution of margin of victory values.

3.2 How?

We now turn our consideration to the how part of our typology, which contains methods, defined as families of related visual encoding and interaction techniques [45]. This part of our typology, shown in Figure 1b, is likely to be most familiar to readers, as it contains a number of methods associated with interaction techniques that are well-represented by several extant classification systems [19, 42, 57, 82]. We distinguish between three classes of methods: those for encoding data, those for manipulating existing elements in a visualization, and those for introducing new elements into a visualization.

Encode: the majority of visualization tasks rely on how data is initially encoded as a visual representation. A full enumeration of visual encoding techniques for various data types in terms of methods is beyond the scope of this paper and appears elsewhere [45].

Manipulate: the following methods affect existing elements of a visualization, modifying them to some extent. These methods represent families of interrelated techniques incorporating both interaction and visual encoding. We consider visual encoding and interaction techniques in a unified way because many methods incorporate aspects of both [41, 43], such as focus + context techniques [20, 46].

Select refers to the demarcation of one or more elements in a visualization, differentiating selected from unselected elements [56]. Examples range from directly clicking or lassoing elements in a single visualization to brushing methods used to highlight elements in visualization systems incorporating multiple linked views [77].

Navigate methods include those that alter a user’s viewpoint, such as zooming, panning, and rotating. Other methods trigger details-visualization, modifying them to some extent. These methods represent the following methods affect existing elements of a visualization, and those for introducing new elements into a visualization.

Import pertains to the introduction of new elements to the visualization, including new data elements. In some environments, these new data elements might be loaded from external sources, while others might be manually generated.

Derive methods compute new data elements given existing data elements. Aggregating data often implies deriving data, however this may not always be true: we further specify that derived data must be persistent, while aggregated data need not be. For instance, a user might derive new attributes for tabular data using a multidimensional scaling algorithm.

Finally, record methods save or capture visualization elements as persistent artefacts. As such, record methods are often associated with produce. These artefacts include screen shots, annotations, lists of bookmarked elements or locations, parameter settings, or interaction logs [63]. An interesting example of record is that of assembling a graphical history [21], in which the output of each task includes a static snapshot of the visualization, and these snapshots accumulate in a branching meta-visualization. Recording and retaining artefacts such as these are often desirable for maintaining a sense of analytical provenance, allowing users to revisit earlier states or parameters settings.

3.3 What?

Previous work has reached no agreement on the question of what constitutes a visualization. Many classification systems do not address it at all; others discuss what implicitly, as indicated by the parenthetical terms in Table 1. Of those that classify what, some focus on the level of the entire dataset, such as tables composed of values and attributes or networks composed of nodes and links [61]. Others allow more precise specification of data–attribute semantics, such as categorical, ordinal, and quantitative [11]. A few classification systems include not only data but also views as first-class citizens [13, 14, 23, 75]. Specific examples of what as classified in previous work include:

- Values, extremum, ranges, distributions, anomalies, clusters, correlations [2].
- Graph-specific objects [37]: nodes, links, paths, graphs, connected components, clusters, groups.
- Time-oriented primitives [1]: points, intervals, spans, temporal patterns, rates of change, sequences, synchronizations.
- Interaction operands [75]: pixels, data [values, structures], attributes, geometric [objects, surfaces], visualization structures.

are synonymous with outright deletion. As an example of temporary filtering, consider a user examining an age histogram based on population census data. First, she decides to exclude males, then further adjusts her filter criteria to focus solely on unemployed females. Finally, she revises the gender criteria to focus on unemployed males. A common example of permanent filtering, or deletion, is that of manually selecting and removing outliers resulting from errors in data entry. Alternatively, consider a scatterplot in which some points are labelled with manually generated categorical tags. Deleting a tag would remove this categorical label from all points having that tag.

Aggregates concerns methods that change the granularity of visualization elements. As such, we also consider its converse, segregate, as being associated with this family of methods. For example, a user may adjust the granularity of a continuous time scale in a line graph, aggregating daily values into monthly values, or segregating annual values into quarterly values. Alternatively, a user may aggregate a clique within a node-link diagram into a representative glyph, or segregate clique glyphs into their component nodes.

Introduce: While manipulate methods alter existing elements of the visualization, introduce methods add new elements. Annotate refers to the addition of graphical or textual annotations associated with one or more visualization elements. When an annotation is associated with data elements, an annotation could be thought of as a new attribute for these elements. The earlier example of manually tagging points in a scatterplot with categorical labels is one such instance of annotating data.
In this typology, we have chosen a flexible and agnostic representation of what that accommodates all of these modes of thinking: in short, we have a “bring your own what” mentality. The only absolute requirement is to explicitly distinguish a task’s input and output constraints when describing sequences of interdependent tasks [72]. An extensive discussion of what that dovetails well with this typology appears elsewhere [45], but it cannot be effectively summarized in a few paragraphs; thus, it is beyond the scope of this paper.

3.4 Concise Task Descriptions

Our multi-level topology can be used to concisely describe visualization tasks. Each task is defined by why it is executed, the method(s) by which it is executed (how), and by what the task pertains to. Single tasks may involve multiple nodes from each part of the topology, as shown in Figure 2.

Fig. 2. Task descriptions for Example #1 (left): casually encountering an choropleth electoral map and comparing election results for two regions; and Example #2 (right): presenting a path between two nodes in a large tree using SpaceTree [20] and TreeJuxtaposer [46].

We have chosen to present these descriptions using a simple and flexible visual notation, rather than with a formal grammar [5, 35, 72, 80]: in doing so, creating and iterating on task descriptions can be easily integrated into existing collaborative design and ideation activities, making use of materials such as coloured sticky notes and whiteboards. A crucial aspect of these descriptions is that sequences of interdependent tasks can be chained together, such that the output from earlier tasks forms the input to later tasks, as discussed in the following case study and as represented in Figure 3.

4 Case Study: A Sequence of Interdependent Tasks

Visualization tasks are seldom executed in isolation, and the output of one task may serve as input to a subsequent task. To illustrate this type of dependency, we present a case study in which our typology is used to describe a sequence of interdependent tasks.

Overview [25] is a visual analytics tool for exploring large text document collections. Overview supports the task of producing a set of tags for semantically related documents as a new categorical attribute, using clusters of documents as scaffolding for the discovery process. Tags are assigned to clusters by means of annotation. However, one must first explore visualizations of the document space and identify clusters of interest. The user is presented both with two linked visualizations, one in which documents are encoded as points in a scatterplot, and another in which clusters of documents are encoded in a hierarchical tree layout. Identifying clusters is facilitated by navigating and selecting clusters in both visualizations; upon selection, frequent document or cluster terms and raw individual document text are shown in secondary displays. This task too has a dependency on the result of an earlier task. Before the the data is encoded in the two visualizations, a set of two-dimensional distances between documents must be produced: they are derived and aggregated from the original high-dimensional distance matrix using a multidimensional scaling algorithm.

Using our typology, we can express dependencies in which the output of one task serves as the input of another, such as the relationship between how data is derived and the choice of visual encoding technique [43]. Such dependencies are represented in Figure 3.

As in the examples of Section 2, we can compare our description to those generated by extant classification systems. Consider Chuah and Roth’s classification of basic visualization tasks [14], which distinguishes between three categories of operations. Using this classification, this sequence of tasks could be described as having data operations (derived attributes), graphical operations (encode data), and set operations (create set, express membership). This description does specify how and what, but it does not express the interdependencies within a sequence of tasks, nor does it tell us why these tasks were performed. Neither can we easily distinguish when sets (or clusters) are created, as this might occur before the data is encoded, such as in the case of the cluster hierarchy, but not in the case of the scatterplot.

While the description based on Chuah and Roth’s [14] classification is atemporal, Chi and Riedl’s operator interaction framework [13] defines stage- and transformation-based operators occurring along the visualization pipeline. Their framework does not contain a comprehensive list of operators, so we draw from the example operators cited in their paper to describe this sequence of tasks as follows:

1. visualization transformation operators: dimension reduction, cluster
2. visual mapping transformation operators: tree layout, scatterplot
3. view stage operators: zoom, focus, brush, details-on-demand, pick

This description does capture the interdependencies for this sequence of tasks, though it mischaracterizes the processes of dimension reduction and clustering as transformations on a visualization, rather than transformations on data, a distinction central to our definition of derive. While this description captures up until the second task in our description, it does not capture the final task of producing cluster tags by means of annotation.

The description based on our typology retains the separability of these tasks, ensuring the distinction between interim inputs and outputs. Another problem with descriptions generated by extant classification systems was that of coverage; the how part of our typology includes both derive and annotate, while descriptions generated by other classification systems could not account for the latter [2, 13], or both [50, 61, 73, 78, 82, 83]. Finally, our description also accounts for both why data is derived and why clusters are annotated with tags, whereas descriptions generated using extant classification systems mention how a task is performed in relation only to when it is performed [13] or to what it is performed on [14]. We maintain that a task description requires why, how, and what; the question of when for a sequence of interdependent tasks is best served by denoting task input and output.

5 Connections to Previous Work

Our typology was informed in part by related work, including extant classification systems and established theoretical models, and in part by new thinking with many points of divergence from previous work. We surveyed work relating to tasks and models of user behaviour, spanning the research literature in visualization, visual analytics, human-computer interaction, cartography, and information retrieval. We focus on two subsets that informed the configuration of our typology: thirty works that explicitly contribute a taxonomy, typology, characterization, framework, or model of tasks, goals, objectives, intentions, activities, or interactions [2, 3, 5, 10, 11, 12, 13, 14, 15, 19, 23, 29, 31, 37, 39, 42, 50, 51, 56, 57, 59, 61, 64, 66, 72, 73, 75, 78, 82, 83], along with twenty other references that make compelling or noteworthy assertions about user behaviour [1, 4, 16, 17, 18, 27, 40, 44, 52, 54, 58, 65, 63, 67, 70, 69, 71, 79, 76, 80]. The similarities between the individual nodes of our typology and those of extant classification systems and other related work are presented in detail in Table 1.
Table 1 serves three purposes: it documents our choices of terms for the purpose of reproducibility, it illustrates the influence of previous work on our thinking, and it summarizes overrepresented and underrepresented areas in the literature, such as consume and enjoy in the why part of our typology. Note that non-leaf nodes in the how part of our typology are poorly represented, serving to indicate the gap between low and high levels of abstraction.

5.1 Extant Classification Systems

The scope of extant classification systems can be categorized in three different ways: level, temporality, and specificity.

**High vs. low:** much of previous work can be divided into those having a low or high level of abstraction, with very little falling in between. Relying solely on either type of classification leads to the aforementioned ends-means confusion, thereby limiting the potential for rigorous analysis. Low-level visualization usage is well represented in related work [2, 5, 10, 12, 13, 14, 15, 19, 20, 37, 38, 59, 61, 72, 73, 75, 78, 82, 83]. Elements common to many of these classification systems include select, filter, and navigate. Following Lee et al. [36], we note that low-level classifications of tasks are often conflated with those of interaction techniques. At the high level, abstract tasks can be found in the context of theoretical models, but without explicit connections to low-level visualization tasks [3, 11, 31, 39, 51]. Examples of these include confirm hypotheses, present, and explore. Low-level classification systems often provide a sense of how a task is performed, but not why; high-level models are the converse. Our focus on multi-level descriptions of visualization tasks is intended to close this gap and resolve the ends-means confusion.

**Sequences vs. constraints:** most classification systems are atemporal in that they do not have any way to express sequences or dependencies between different stages. A few classification systems are explicitly temporal in that they divide user behaviour into larger stages that occur in specific sequences or cycles. Examples include pipeline models for visualization construction [13] or data analysis [27], or cyclic models such as knowledge crystallization [11] or information foraging and sensemaking [51]. However, empirical observations of visualization usage have indicated a mismatch between the specific cyclical or sequential patterns proposed by these models and actual user behaviour [26]. Vicente argues that sequence-based approaches to task analysis are overly rigid and thus inappropriate for describing such open-ended tasks [74], and that constraint-based approaches to task analysis allow for more flexibility in terms of how a task is performed. Descriptions based on our typology do not force any strict global temporal orderings, as imposed by sequence- or cycle-based models; instead, they accommodate local interdependencies within sequences of tasks by way of constraints on task input and output.

**General vs. specific:** many visualization classification systems represented in our survey are generally applicable, though specifically-targeted classification systems and models do exist. Examples include Lee et al.’s task taxonomy for graph visualization [37] and Lam-marsch et al.’s task framework for time-oriented data [35]. Our typology encompasses and complements these specific classification systems, and we encourage further development of more like these.

We also identify five general classification systems that span low-level and high-level tasks. These classification systems had the highest contributions to the organization of our own typology:

**Springmeyer et al. (1992) [66]:** This classification of scientific data analysis covers both how and why, but these aspects are not clearly distinguished within its hierarchical structure, that which begins with a high-level distinction between investigation and integration of insight.

**Mullins and Treu (1993) [42]:** This extensive taxonomy contains over 150 elements: an exhaustive list of high-level tasks, and we encourage further development of more like these.

**Roth (2012) [57]:** Roth’s taxonomy, based on Norman’s Stages of Action model [47], classifies cartographic interaction primitives as objectives, operators, and operands. Norman’s model describes a series of translations between the user’s goal, an immediate intention (or objective), and a series of actions (operators) performed on an environment (of operands). Roth’s taxonomy is closely aligned with our notions of why, how, and what, and thus has a high-level structure similar to that of our own typology. However, Roth’s taxonomy imposes a spatial constraint on where operands are located in space, as discussed in Section 2. In contrast, we restrict our classification of what to that of input and output; the location of operands is represented by the search node in the why part of our typology.

Fig. 3. Use of Overview [25], a visual analytics tool for exploring large text document collections, entails a sequence of interdependent tasks. Note that task 1a and task 1b can occur in either order, and that task 2 has a dependency on both of these tasks.
<table>
<thead>
<tr>
<th>WHY?</th>
<th>consume –</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ present</td>
<td>present, [63, 79], author, compose [11]<em>, build (case), tell (story) [51]</em>, depict [50]<em>, express (ideas), describe [66]</em>, guide, share [23]* inform, elaborate [83]*, report [27],</td>
</tr>
<tr>
<td>→ discover</td>
<td>discover, [40], explore [83]* [79], verify [12]* [40], synthesize [42]* [40], investigate, integration (of insight) [66]* [40],</td>
</tr>
<tr>
<td>(generate, verify, hypotheses)</td>
<td>frame operations: constitute, elaborate, question, reframe [31]<em>, assimilate, assess, understand [50]</em>, infer [73]*,</td>
</tr>
<tr>
<td>analyze [42, 50]* [40], support, reevaluate (hypotheses) [51]<em>, monitoring [76], confirm (hypotheses), expose (uncertainty), formulate (cause and effect), concretize (relationships), learn (domain parameters), multivariate explanation [3]</em>, evaluate, learn, investigate [40], open-ended exploration, diagnosis [52], abduction, deduction, induction [50], generate, confirm (hypotheses) [4, 18], integrate, interpret [18], exploratory and confirmatory data analysis [71]</td>
<td></td>
</tr>
<tr>
<td>→ enjoy</td>
<td>visualization use in casual contexts [54, 65], strolling [17]</td>
</tr>
<tr>
<td>produce</td>
<td>export [57]<em>, store [42]</em>, save [39, 57]<em>, extract [11, 61]</em>, generating (images) [52], (a classification) [11, 66]<em>, (a categorization) [42, 78, 83]</em>, (a record of one’s history / process) [23, 61, 66]*</td>
</tr>
<tr>
<td>search</td>
<td>search [11, 51, 83]<em>, acquire [42]</em>, visual queries [76]</td>
</tr>
<tr>
<td>lookup</td>
<td>lookup [12]* [40], identify: lookup (value) [78]<em>, (value) lookup [59]</em>, retrieve (value) [2, 37, 50, 57]* [70], procure [57]*</td>
</tr>
<tr>
<td>→ browse</td>
<td>browse [11, 42, 50, 64]* [17, 69], search [57]<em>, finding (gestalt) [10]</em>, browsing tasks: follow (path) [37]*</td>
</tr>
<tr>
<td>→ locate</td>
<td>locate [39, 42, 73, 78, 83]* [1, 18], search [12]* [17], search (for known item) [40], seek [64]*, pathfinding [76]</td>
</tr>
<tr>
<td>→ explore</td>
<td>explore [39, 50, 82]* [76, 80], forage [11, 39, 51]<em>, finding (gestalt) [10]</em>, (overview) tasks [37]<em>, find (clusters, correlations, extremum, anomalies) [2, 37, 50]</em>, determine (correlations) [59]<em>, determine (clusters) [78]</em></td>
</tr>
<tr>
<td>query</td>
<td>query [56]<em>, posing queries [10]</em>, elementary and syntactic tasks [5]<em>, levels of questions [72]</em>, question answering [40]</td>
</tr>
<tr>
<td>→ identify</td>
<td>identify [37, 42, 50, 57, 73, 78, 83]* [1, 58], reading (the data) [18], read (fact, pattern) [11]<em>, examine [66]</em>, determine (range) [2, 37, 50]<em>, determine / characterize (distribution) [2, 37, 50, 78]</em>, recognize [31]*</td>
</tr>
<tr>
<td>→ compare</td>
<td>compare [5, 31, 42, 50, 57, 66, 72, 73, 83]* [40], compare (within a relation vs. across / between relations) [59, 78]<em>, relation seeking [5]</em>, read comparison [11]<em>, making comparisons [10]</em>, [76], discriminate [42]<em>, associate [57]</em></td>
</tr>
<tr>
<td>→ summarize</td>
<td>summarize [83]<em>, summarize (set), enumerate (set objects) [14]</em>, overview [11, 15, 61]<em>, (overview) tasks [37]</em>, scan [37, 42]<em>, connectional tasks [5]</em>, count [37, 61]*, visualization [17], review [63]</td>
</tr>
<tr>
<td>HOW?</td>
<td>encode encode [14, 50, 82, 83]<em>, create mapping [14]</em>, visualize [23, 73]<em>, generate [66]</em>, transform (visual mapping) [13]*</td>
</tr>
<tr>
<td>manipulate</td>
<td>manipulate [80], (object) manipulation [42]<em>, modify [56]</em>, (data) manipulation loop [76]</td>
</tr>
<tr>
<td>→ select</td>
<td>select [23, 42, 50, 56, 72, 75, 82]<em>, brush [19, 29, 50]</em> [13, 76, 80], distinguish [78, 83]<em>, emphasize [83]</em>, differentiate [50]<em>, highlight [15, 23, 56]</em> [76], identify: portray, individualize, profile [83]<em>, indicate [42, 56]</em>, mark [42, 82]<em>, reference [42]</em>, outline (clusters) [83]<em>, promote [11]</em>, track [82]<em>, pick [42]</em> [13], express (set membership) [14]* connect [50, 82]*</td>
</tr>
<tr>
<td>→ navigate</td>
<td>navigate [23, 64, 75]* [40, 44, 52, 76, 80], focus [10, 15]* [13], details-on-demand [11, 61]<em>, [13], flip through [13] zoom [10, 11, 15, 19, 29, 42, 50, 57, 62, 82]</em> [13, 44, 80], pan [10, 19, 42, 50, 57, 82]<em>, [80], elaborate [50, 82]</em>, abstract [50, 82]<em>, change (range) [19]</em>, drill down [15]<em>, maneuver / navigate [66]</em>, rotate [13, 80]* revisit [19, 37]*</td>
</tr>
<tr>
<td>→ arrange</td>
<td>arrange [10, 57]<em>, sort [2, 19, 23, 37, 50]</em> [44], rank [57, 78, 83]<em>, coordinate [23]</em>, delineate, sequence [57]<em>, index [59]</em>, move [42, 56]<em>, edit [42]</em>, organize [23]* [63], orient, permute, position, translate [13], reorder [11, 80], configure [73]<em>, reconfigure [50, 82]</em>, restructure [39]*</td>
</tr>
<tr>
<td>→ change</td>
<td>change (parameters) [15]* [13], change (metaphor) [19]<em>, change (representation) [15]</em>, change (vis. encoding) [44], transform [56]* [40, 80], transform (mapping), shift, scale, set (graphical value) [14]<em>, rotate, scale [13], configure [73]</em>, animate [13, 80], distort [29, 75]* [13], orient / transform [66]<em>, (object) manipulation: transform, stretch, shape [42]</em>, re-express, re-symbolize, re-project [57]<em>, edit [42, 57]</em>, activate [56]*</td>
</tr>
<tr>
<td>→ filter</td>
<td>filter [2, 11, 19, 23, 31, 37, 42, 50, 51, 57, 61, 82]* [44, 70, 80], subsetting, (value) filtering, (view) filtering [13], exclude [40, 70], screen: filter, suppress, conceal [42]<em>, maneuver: (data) management / culling [66]</em>, configure [73]<em>, delete (objects, sets, graphical objects) [14]</em>, delete [11, 19, 56]<em>, overlay [57]</em>, restore [19, 42]*</td>
</tr>
<tr>
<td>→ aggregate</td>
<td>aggregate [42]* [13, 44], cluster [11]* [13], aggregate, [42, 78, 83]<em>, simplify [13], link [10, 15, 29, 42]</em> [63, 80], merge [19]<em>, generalize / merge [83]</em>, assemble [42]<em>, create (set) [14]</em>, split [19]<em>, disassemble [42]</em>, dissociate [42]<em>, reveal: itemize, separate [83]</em>, segregate: ungroup, unlink [42]<em>, rewrite, overlay [42]</em></td>
</tr>
<tr>
<td>introduce</td>
<td>introduce [42]*</td>
</tr>
<tr>
<td>→ annotate</td>
<td>annotate [19, 23, 57]<em>, add placemark [82], create (anchors) [39]</em>, create / copy (graphical objects) [14]<em>, create / modify (note) [19]</em> externalize (analysis artefacts) [63], give a meaningful name to (groups / clusters) [37]*</td>
</tr>
<tr>
<td>→ import</td>
<td>import [57]<em>, add (objects) [14]</em>, create [11, 42]<em>, generate [56]</em>, (data) entry [42]*, load [39]</td>
</tr>
<tr>
<td>→ derive</td>
<td>derive [23]<em>, derived (attributes) [14]</em>, derive (new conditions) [66]<em>, compute (derived value) [2, 37, 50]</em>, copy [56]<em>, compute [83]</em>, calculate [42, 57, 66]<em>, configure, determine [73]</em>, average [11]<em>, computation operators [12]</em> transform (data) [13]<em>, estimate, generate (statistics) [66]</em>, extrapolate [42]* [18], interpolate [42]* [18]</td>
</tr>
<tr>
<td>→ record</td>
<td>record [23, 42, 66]<em>, bookmark [19]</em>, history [61]<em>, redo, undo [19, 82]</em></td>
</tr>
</tbody>
</table>

Table 1. Nodes in our typology of abstract visualization tasks and their relation to the vocabulary used in previous work. The table serves to document our typology in detail, to illustrate the influence of previous work, and to show overrepresented and underrepresented areas in the literature. Underlining is used where a term used in our typology appears in related work. Terms in parentheses are encompassed by the what part of our typology. Previous work that explicitly contributes a classification system is denoted by *; otherwise sources incidentally make compelling or noteworthy assertions about visualization usage.
What these five classification systems have in common is that they are atemporal, and most span our characterization of why and how. Our typology integrates and extends this work, adding a specification of what, the input and output tasks. As a result, our typology can be used to describe sequences of tasks, in that the output of one task may serve as the input of another.

5.2 Theoretical Foundations

Our typology was also informed by four theoretical frameworks:

Distributed cognition: The distributed cognition literature offers us a useful distinction between pragmatic and epistemic actions [30, 38]. Pragmatic actions are explicitly and consciously goal-directed, while epistemic actions refer to processes of coordination between internal mental models and external representations of information [38], often performed in support of pragmatic actions. This distinction is lost in low-level classification systems; in isolation from higher-level goals we are unable to discern between pragmatic and epistemic actions. Our typology accommodates this distinction. External representations are the graphical elements of visualizations displayed to or created by the user. Pragmatic actions are correspond to the how part of our typology, while epistemic actions are captured by the why part of our typology. The set of manipulate methods are particularly well-suited for the purpose of describing epistemic actions and their role in coordinating between internal and external representations.

Stages of action: Norman’s Stages of action model [47] and its influence on Roth’s objective-operand-operator classification of extant classification systems [58] helped shape the why-what-how organization of our typology. In the process of evaluating visualizations, we can discuss the Norman’s gulf of execution with respect to the how part of the typology, asking if the user has the means to execute the task. Also central to Norman’s model is the gulf of evaluation, useful for reasoning about whether the output of a task matches the user’s expectation. However, this gulf is more applicable when reasoning about specific interaction techniques, which are not directly addressed by our typology. More recently, Lam extended the model with a gulf of goal formation [33], relevant whenever a user articulates their own questions pertaining to a visualization, thereby specifying the ends of a task. This gulf corresponds to the why part of our typology, which allows us to abstractly describe these questions.

Sensemaking: The why part of our typology overlaps with and bridges to high-level processes of decision making and prediction described in theories of information foraging and sensemaking, both temporal stage-based models [11, 51] and atemporal data-frame models [31]. In particular, sensemaking models connect at the levels of discover, denoting hypothesis generation and formation, present, and the types of search: lookup, locate, browse, and explore.

Play theory: Casual interactions with visualization pose another set of problems for many extant classification systems, in that task specifications for these contexts are not easy to motivate by a need to present, discover, or produce [65]. We included enjoy in the why part of the typology to encompass casual consumption of information, curiosity-driven tasks without expectations or predicted outcomes [16, 69]. The choropleth map example from the paper’s introduction is an instance of this type of task. As visualization becomes increasingly pervasive in casual contexts, we may turn to theories of casual information seeking and newsreading behaviour, such as Stephenson’s Play theory [67] to motivate visualization tasks in these contexts. This theory accounts for media consumption activities that bring no material gain, serving no “work” functions, but instead induce moments of absorption and self-enchantment. Casual media consumption relies upon serendipitous apperception, a readiness to interact with information relating to existing interests. Studies of newsreading behaviour indicated that people read most avidly what they already know about [67], a seemingly irrational activity that cannot be described as an explicit need to discover new information. We posited that this behaviour is also true of some visualization consumption, particularly in non-work contexts [54, 65].

6.1 Using the Typology to Describe

The typology’s descriptive power is in its provision of a consistent lexicon for tasks in terms of why, how, and what in a way that supports precise task comparisons across different visualization tools and application domains. This lexicon can be used to describe and compare user tasks as they occur in situ, of particular use to those analyzing current work practices and use of visualization tools “in the wild”. This form of inquiry is often performed within a single domain, such as enterprise data analysis [27] or intelligence analysis [28], wherein tasks are described in a domain-specific way. An interface- and domain-independent vocabulary for multi-level tasks allows practitioners to perform comparative analyses of tasks involving different visualization tools occurring in different disciplines.

Only the descriptive aspect of the typology has been directly validated in this paper; we used our typology to describe several empirical cases including single tasks and a sequence of interdependent tasks. We also demonstrated its ability to facilitate the comparison of tasks as they are performed using different visualization tools. Future work includes a further examination of its descriptive power by analyzing whether it covers the full set of abstract tasks described in previous published design studies. In addition, we acknowledge that the typology does not at present explicitly address collaborative use of visualization, although we did consider some of the issues involved during its development [26]. Future work will verify if the typology can sufficiently describe collaborative tasks, or if extensions are needed.

6.2 Using the Typology to Generate

The typology’s generative power stems from its ability to prescribe and inform design. In particular, the typology is well-suited to support task analysis occurring throughout the formative discover and design stages of the nine-stage design study framework [60]. In the discover stage, the practitioner must transform a domain problem into an abstract task description; the typology provides an explicit set of choices for why the task is performed, possibly making this difficult aspect of design studies more tractable [41, 43]. In the design stage, the practitioner then chooses how the task will be supported, calling upon the existing repertoire of encoding and interaction techniques or inventing new ones [44]. During both stages, the practitioner must consider what comprises the inputs and outputs of these tasks, remaining aware that these tasks may have interdependencies. We aim to validate the typology’s generative power in our future design studies.

6.3 Using the Typology to Evaluate

Part of the motivation for developing this typology arose from our struggle to characterize and compare the tasks of different users in an ongoing post-deployment qualitative user study evaluating...
Overview [25], the system described in Section 4. The typology is intended to facilitate the evaluation of user experiences as well as field studies examining visualization usage [34]. We plan to validate the typology’s evaluative power by assessing its utility a multi-level code set for labelling user behaviour, a common practice in observational studies of open-ended visualization use; these include longitudinal insight-based studies [48] and multidimensional in-depth long-term case studies (MILCs) [62]. A MILC study of SocialAction [49], a social network visualization tool, incorporated Yi et al.’s [82] categorization of interaction techniques into the analysis of how users’ tasks were performed; we envision a similar use for our task typology, in which the scope of analysis will broaden to include why and what. Mixed-method qualitative evaluation studies allow practitioners to identify how a task is performed along with its inputs and outputs via interaction logs and observational analysis; we can also identify why a task is performed via interviews, think-aloud protocols, and artefact analysis.

Task descriptions generated by our typology could be also used to better understanding users’ individual analytical strategies and the context-dependent variability with regards to how a task is performed [74, 84]. Understanding individual problem solving strategies in terms of mental model formation and coordination is also an ongoing and well-distributed cognition research [24, 39]. Our typology and its accommodation of pragmatic and epistemic actions may serve to further this research in the context of visualization usage.

Finally, while our typology does not provide the low level of specification required for empirical experiments aimed at evaluating user performance with respect to specific interaction and visual encoding techniques [34], it may be useful in suggesting external validity for these low-level benchmark experimental tasks.

7 Conclusion

Our primary contribution is a multi-level task typology that relates both why and how a task is performed to what the task pertains to in terms of inputs and outputs. The typology allows for the precise description of complex tasks as sequences of simpler tasks, with their interdependencies made explicit. One major advance of the new typology is that it bridges the gap between the low-level and high-level tasks of previous work by providing linkages between them, distinguishing the ends and means of a task. Our typology integrates new thinking with extant systems for classifying tasks, and with previously tasks of previous work by providing linkages between them, distinguishing the ends and means of a task. Our typology integrates new thinking with extant systems for classifying tasks, and with previously established theoretical frameworks spanning multiple literatures. The multi-level task typology presented here is another step towards a systematic theoretical framework for visualization, helping us to describe existing user experiences, evaluate them, and generate new ones.

Acknowledgements

We thank NSERC for funding. We thank Jessica Dawson, Joel Ferstay, Stephen Ingram, Joanna McGrenere, Miriah Meyer, Michael Seldmair, and Colin Ware for their feedback. We also thank François Guimbretière for very early discussions on the topic. We acknowledging Ron Rensink for the original idea of classifying lookup, browse, locate, and explore according to target identity and location.

References
