

Visualization Onboarding: Learning How to Read and Use Visualizations

Christina Stoiber*
University of Applied Sciences
St. Pölten

Florian Grassinger†
University of Applied Sciences
St. Pölten

Margin Pohl‡
TU Wien

Holger Stitz§
Johannes Kepler University Linz

Marc Streit¶
Johannes Kepler University Linz

Wolfgang Aigner||
University of Applied Sciences
St. Pölten

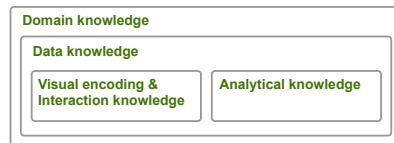
1 WHY

Why is visualization onboarding needed?

The aim of visualization is to support humans in dealing with large and complex information structures, to make them more comprehensible, facilitate exploration, and enable knowledge discovery. But, users often have problems in reading and interpreting data from visualizations, in particular when they experience them for the first time.

3 WHO

Who is the user? Which knowledge gap does the user have?



2 WHAT

What is visualization onboarding?

Visualization onboarding is the process of supporting users on how to read, interpret, and extract information of visual representations of data.

4 HOW

How is visualization onboarding provided?

Type	Context-sensitivity	Interaction
including for example:	- context-sensitive	- passive
- guided tours	- context-free	- active
- step-by-step wizards	- embedded	- reactive
- video-based tutorials		
- tooltips		
- walkthroughs		
- help center		
- mouseover popup		
- overlays		
- learning environment		
- etc.		

VISUALIZATION ONBOARDING



5 WHERE

Where is visualization onboarding provided?

- external
- internal
- learning environment

6 WHEN

When is visualization onboarding used?

- before
- during
- first use

Figure 1: A visual overview of the onboarding design space and of how all six questions, Why, What, Who, How, Where, and When relate to one another. Each question corresponds to one paper section as indicated by the numbered tag near each question title.

ABSTRACT

The aim of visualization is to support humans in dealing with large and complex information structures, to make these structures more comprehensible, facilitate exploration, and enable knowledge discovery. However, users often have problems reading and interpreting data from visualizations, in particular when they experience them for the first time. A lack of visualization literacy, i.e., knowledge in terms of domain, data, visual encoding, interaction, and also analytical methods can be observed. To support users in learning how to use new digital technologies, the concept of onboarding has been successfully applied in other domains. However, it has not received much attention from the visualization community so far. With our position paper, we aim to work towards filling this gap by proposing a design space of onboarding in the context of visualization.

Index Terms: Visualization literacy—learning—user onboarding—visual analytics

1 INTRODUCTION

Visualization has become more important and more widespread; not only in the context of science and business, but also in everyday contexts such as data stories in newspapers, in books, or on the internet. The size and complexity of today’s datasets overwhelm traditional business charts such as bar charts, line charts, or pie charts. As a fact, more advanced visual representations are necessary to capture more complex data structures and larger amounts of data.

Visualization can be seen as a process that transforms data into a visual form [11, 17]. As a user, this transformation needs to be

transparent in order to be able to decode the visual representation and correctly reason about the data. Even though humans are visual beings and visual representations are easier to understand than other forms of data representations, users still need to learn how to read and comprehend them – a certain level of visualization literacy is essential. Especially for novice users, this task is often difficult and leads to wrong conclusions and insights concerning the data. In that case, visualization onboarding methods can support users in understanding data visualizations and take full advantage of the power of visual representations. However, very little is known about onboarding methods for visualizations.

In general, the onboarding process is a key aspect of user experience (UX) that allows them to discover functionality of applications and achieve their goals. Onboarding flow pattern such as instructional text, just-in-time hints, or interactive tutorials can be found to support new users in learning the applications key benefits and features. Industry practitioners [4, 27, 28, 30] have started to create guidelines for the design of onboarding. Yet, there is scarce methodological and theoretical guidance to inform onboarding design decisions.

In the context of education some studies have been carried out to find out how novices read and create charts [1, 3, 14, 44]. However, it is questionable whether these concepts can be applied to onboarding concepts for visualizations. In this paper, we introduce a design space for visualization onboarding. Currently, we are working on a systematic literature review of existing onboarding concepts in scientific publications and commercial visualization tools. We present first results and insights in this field. In our work on visualization onboarding we consider interactive visual analytic (VA) systems. Firstly, we introduce the topic of visualization onboarding in Section 2 and present related areas in Section 3. As a result of our literature search and prior experience we derived an abstract, descriptive design space that brings together the different aspects of onboarding concepts for visualization (see Section 4). We structured our design space of visualization onboarding along the *Five W’s and How* tool [25, 26]: **WHY** is visualization onboarding needed? **WHAT** is visualization onboarding? **WHO** is the user? Which knowledge gap does the user have? **HOW** is visualization onboarding provided? **WHERE** is visualization onboarding provided? **WHEN**

*e-mail: christina.stoiber@fhstp.ac.at

†e-mail: florian.grassinger@fhstp.ac.at

‡e-mail: margit@igw.tuwien.ac.at

§e-mail: holger.stitz@jku.at

¶e-mail: marc.streit@jku.at

||e-mail: wolfgang.aigner@fhstp.ac.at

In Section 5, we discuss gaps and challenges. In the next section we present the background on user onboarding including the grounding of this area in visualization.

2 WHAT IS VISUALIZATION ONBOARDING?

We start by explaining onboarding from the perspective of the Human-Computer-Interaction community, followed by a discussion of onboarding from the visualization point of view.

2.1 Onboarding in HCI

In general, the term onboarding originated from organizational theory where it is widely used and refers to the step taken by an organization to facilitate and socialize newcomer adjustment [36]. Faced with a similar challenge in classical user interface design, the Human-Computer Interaction (HCI) research community quickly recognized the importance of providing effective help systems for software applications [12]. Different approaches have since been investigated in depth [45], such as *step-by-step wizards* [6, 21, 34], *overlays* [21], *Q&A forums* [42], or *video-based tutorials* [5, 23, 47]. Moreover, context-sensitive and adaptive help approaches have been discussed. Context-sensitive help has largely been about attaching help options to specific user interface controls [15], e.g., tooltips, “?”-icon, dialog boxes, or even choosing a command to see animated steps [52] or videos [23]. One drawback is that designers have to anticipate where users might seek help, so that they can develop these tools beforehand.

More recently, the topic of onboarding has received interest by the UX practitioner community. There are several blogs and articles presenting guidelines and design inspirations for onboarding concepts [4, 27, 28, 30]. UX practitioner S. Hulik¹ defined *user onboarding* as “the process of increasing the likelihood that new users become successful when adopting your product.” [37]. Examples of onboarding design patterns include: instructional text, tours, progress bars, just-in-time hints, tips placed in feeds of user-generated content, and interactive tutorials [4]. Hulik and Higgins [27, 28, 30] provide a set of guidelines and processes to engage with the challenge of onboarding design and integration. Higgins advocates three principles of well designed onboarding [27]: 1) Facilitate exploration in an authentic space, 2) Gradually engage, 3) Provide clear next steps.

2.2 Towards Visualization Onboarding

Visual mapping is the process of assigning data variables to visual channels, which results in either a static or an interactive visual representation. This process is the central component of virtually all known conceptual models of visualization, such as [11, 54]. Understanding visual mapping is key for the correct decoding of both the visual representation and the underlying information. In addition to that, also the data analysis, filtering, and rendering steps of the visualization process influence the appearance of a visualization idiom and need to be transparent to a certain extent. Particularly for visualization-illiterate users in data analysis and visualization, however, this task is often difficult and prone to the risk of drawing wrong conclusions or insights regarding the data. Onboarding concepts can help users comprehend the visualization process and support learning. We define visualization onboarding as the following: *Visualization onboarding is the process of supporting users in reading, interpreting, and extracting information from visual representations of data.*

To ground our work in VA theory we use the “Simple Visualization Model” by Van Wijk [54] together with extensions by Federico and Wagner et al. [20] in order to capture automated data analysis components. The choice of this model follows a recent uptake of it in current VA research, such as guidance [13] or knowledge-assisted

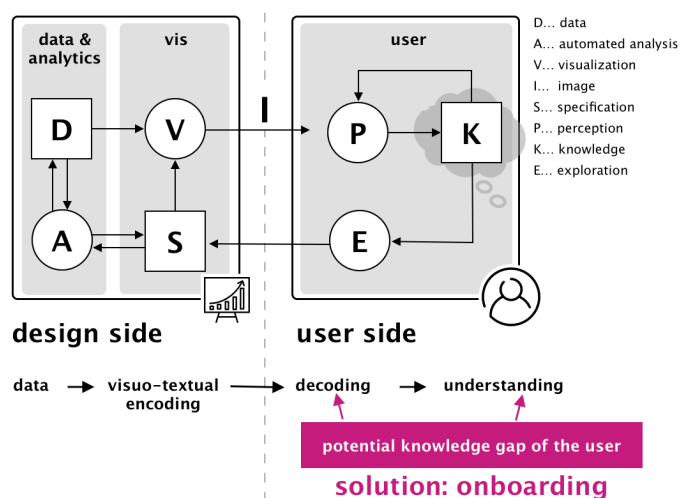


Figure 2: Grounding of visualization onboarding in VA theory [10, 54]

VA [20]. Figure 2 shows an overview of the model along with its relationship to visualization onboarding. *Top*: Conceptual model of visual analytics [20, 54] with the three areas of data & analytics, vis, and user. The first two (left) comprise the machine side where data (D) is transformed by analytics methods (A) and visualized (V) based on some specification (S). The output image (I) is then communicated to the user side (right). The human user perceives and cognitively processes (P) the visual input which leads to new knowledge (K). Using interactive exploration (E) the user can influence both, representation and analytical methods with interaction. *Bottom*: Two sides of the visualization process [10]. The visuo-textual encoding and transformations applied are mainly decided on by the designer of a VA system (left side). At the user side, the image needs to be decoded and understood (right side). This requires knowledge about the topic as well as familiarity with graphic forms (visualization literacy). Without this (meta-)knowledge, a user might not be able to interpret a visualization and draw correct conclusions. In case of such a knowledge gap, VA onboarding can be applied to aid the user in closing the knowledge gap.

Visualization literacy: Recently, an increasing interest in visualization literacy can be observed. For instance, Chevalier et al. [14] reflect on visualization literacy in early education and provide lessons learned and directions for future research. Börner et al. [7] introduced a data visualization literacy framework developed to define, teach, and assess data visualization literacy. However, although the topics of onboarding and visualization literacy have been identified as future challenges by the visual analytics community (e.g., in [33]), not much work has been conducted so far [35]. Visualization literacy has been defined as the ability and skill to read, comprehend, and interpret data from visualization tools [40]. Boy et al. [9] conducted one of the pioneering studies on visualization literacy in the community. They defined the term visualization literacy and proposed a method to assess an individual’s level of visualization literacy by using a set of test items. The definition of Boy et al. describes “the ability to use well-established data visualizations, such as line graphs, to handle information in an effective, efficient, and confident manner” [9]. Börner et al. [8] explored the general public’s current level of visualization literacy. They showed 20 printed, static visualizations to visitors of museums. The results of this study have shown that participants with low visualization literacy lack the ability to read network visualizations. In their work, the authors defined visualization literacy as “the ability to make meaning from the interpret patterns, trends, and correlations in visual representations of data”. Lee et al. [40] have further refined the definition of visualization literacy based on the ones of Börner et

¹<https://useronboard.com>

al. [8] and Boy et al. [9]: “Visualization literacy is the ability and skill to read and interpret visually represented data in and to extract information from data visualizations” [40].

Visualization Onboarding Approaches: Tanahashi et al. [53] investigated top-down and bottom-up teaching methods, as well as active or passive learning types for scatter plot, graph, storyline, and treemap representations. Their analysis indicates that top-down exercises were more effective than bottom-up. Additionally, their study shows that utilizing the active learning type with top-down tasks were the most effective. Kwon and Lee [38] explored the effectiveness of active learning strategies. Three tutorial types (static, video-based, and interactive) are used to support the learning of scatterplot visualizations. They observed that participants using the interactive and video tutorials outperformed participants with static or no tutorials. Ruchikachorn and Mueller [48] proposed a concept of teaching by analogy, i.e., demonstrating an unfamiliar visualization method by linking it to a more familiar one. They provide demonstrations of various visualization techniques, e.g., data tables and parallel coordinates; scatterplot matrix and hyperbox; linear chart and spiral charts; hierarchical charts and treemaps. Yalçin [55] presented a design of a contextual in-situ help system for the visual data interfaces Keshif², called HelpIn, to explain the features of this tool. Nowadays most commercial visualization tools already integrate onboarding concepts focusing on the explanation of features. *IBM Cognos Analytics*³, for example, uses step-by-step tours with tooltips and overlays for onboarding new users. A more traditional approach is used by the commercial visualization tool *Advizor*⁴ which uses textual descriptions to explain the visual mapping for various visualization techniques, for example “*A Heat Map summarizes data in a digestible graphical format. It uses both color and size to relay information about data values.*”

3 RELATED AREAS

The field of visualization onboarding covers different disciplines, such as educational theories and cognitive science. Moreover, we elaborate on how onboarding is related to the concept of guidance in visual analytics.

3.1 Educational Theories and Cognitive Science

As visualization onboarding aims for filling the knowledge gaps of users by supporting the learning of new concepts, it makes sense to build upon knowledge from the fields of learning theories and cognitive science. We distinguish between three main educational theories: behaviorism, cognitivism, and constructivism [19]. E-learning systems often integrate elements from different educational theories. This also applies to most onboarding systems.

Nevertheless, constructivist theories seem to be the most appropriate in order to explain learning processes with onboarding systems because they reflect the application of learning in a practical context. The concept of cognitive apprenticeship plays an important role in constructivism [18, 49]. *Cognitive apprenticeship* is a kind of guided participation by learners in real processes of knowledge generation. This is related to the concept of scaffolding [29]. *Scaffolding* is a process where teachers gradually reduce their support for the student until the student can work autonomously. Cognitive apprenticeship and scaffolding can explain processes related to onboarding because the goal of the learners is to solve a real task, and they need support to use the visualization which is gradually reduced.

Another theoretical framework relevant for onboarding is *graph comprehension*, a theory that aims to explain how users make sense of graphs. Most of the investigations in this context deal with simple

small graphs [46]. Nevertheless, the findings from graph comprehension yield interesting results that can inform the design of visualizations. This is especially valuable for onboarding systems because investigations in this area often address the issue of how to design graphs that are appropriate for use in educational contexts. One of the most influential models in the context of the theory of graph comprehension describes this activity as consisting of three stages [22]: (1) reading the data (i.e., finding individual data values), (2) reading between the data (i.e., finding relationships between the data), (3) going beyond the data (i.e., interpreting the data, developing hypotheses about the data).

Support can happen at different levels in the visual exploration process. Besides onboarding, there is also the concept of guidance [13, 50] which is described in the next section.

3.2 Distinction between Onboarding and Guidance

Schulz et al. [50] characterized the concept of guidance in visual analytics. The governing aspects of guidance are context, domain, target and degree. The concept of guidance was further developed by Ceneda et al. [13] who defined *guidance in visual analytics* as “... a computer-assisted process that aims to actively resolve a knowledge gap encountered by users during an interactive visual analytics session.” They illustrate guidance using the metaphor of a smart car which supports the driver during a journey to a destination. If drivers are confident about how to get to the planned destination, they will drive the car, while the car provides guidance by showing the names of the traversed streets, highlighting the position of stop or traffic lights, or also streaming the weather conditions for the current day. In cases where drivers do not know how to reach the destination, the car could provide a higher degree of guidance by displaying turn-by-turn navigation instructions. An even higher level of guidance could be to drive to the destination autonomously, taking over every decision, changing paths if needed — but always leaving the driver the freedom of taking over the steering wheel to depart from the route or react to unexpected situations. These different degrees of guidance are all useful to help drivers reach their destination. However, what if drivers would like to reach a certain destination but do not know how to drive this specific type of car in the first place? Or, what if they do not know about smart cars in general and what they can actually accomplish? This more basic understanding is not covered by the concept of guidance. Therefore, the concept of onboarding aims to fill this gap by supporting the driver in learning how to use this smart car and explaining its main features and capabilities. Both, onboarding and guidance methods can be applied to visual analytics tools: Onboarding helps users to learn, interpret and use the applied VA methods correctly [8]. Guidance supports analysts during an interactive VA session with concrete tasks and data [13].

4 DESIGN SPACE

We describe the aspects of visualization onboarding along the following questions: **WHO** is the user? Which knowledge gap does the user have? **HOW** is visualization onboarding provided? **WHERE** is visualization onboarding provided? **WHEN** is visualization onboarding used?

4.1 WHO is the user?

As users, we need to understand the process and reasoning that lead to the visual appearance, interactive behavior, and findings. Hence, we need to make the process transparent to the users to a certain degree. For conceptualizing this aspect, we adapt the nested model by Munzner and colleagues [43] as the guiding framework for presenting different levels of knowledge. The nested model is a unified approach that splits the design into four levels and combines these with appropriate evaluation methods to mitigate threats to validity at each level. In order to be able to cover visual analytics approaches

²<https://www.keshif.me/>

³<https://www.ibm.com/at-en/marketplace/business-intelligence>

⁴<https://www.advizorsolutions.com/software/interactive-data-visualization/>

and include automated data analysis components, we expand the original model by adding analytical methods alongside visual encoding/interaction idioms. Analytical knowledge—such as different automated data analysis approaches, machine learning methods, or statistical methods applied to the data—is necessary to understand complex visualization interfaces and data. Figure 1 (3) keeps the nesting but shows an altered representation of the different levels. The model components represent the different levels of knowledge that (a) visualization users need in order to correctly interpret (interactive) visualization artifacts and (b) visualization designers have to consider when developing onboarding concepts.

The levels consider the users' prior knowledge such as domain knowledge, data knowledge, knowledge of visual encoding & interaction concepts, and analytical knowledge.

Domain knowledge: A specific domain is a particular field of interest by target users of a visualization tool (e.g., medicine, data journalism, bioinformatics). Each domain has its own vocabulary for describing the data and problems, workflows, and how data can be used to solve a problem. Domain knowledge is also an ensemble of concepts, intellectual tools, and informational resources that a user can draw upon to put the visualized data into context.

Data knowledge: Many visualization tools are specific to a particular *type of data*, such as multivariate data, hierarchical data, network data, time-oriented data.

Data knowledge refers to the necessary knowledge for understanding the data types and structures, or statistical properties of the data. In many cases, users need to know how to get their data into a specific visualization tool as a first step. This relates to a more technical level of knowledge about a particular file format (e.g., CSV, JSON) or structure of the data – *data format*– (e.g., order and data types of individual variables).

Visual encoding knowledge: This type of knowledge is the most obvious one in the context of visualization, as it concerns the visual appearance of the data. Data elements are mapped to visual channels to form visualizations. Understanding this mapping is the basis for being able to correctly interpret the visualization.

Interaction knowledge: Interactivity is crucial for exploratory visualization tools. An interactive visualization tool can support the investigation at multiple levels of detail, such as either a high-level overview or fully detailed views that show a small data subset only [43]. Understanding the interaction concepts used in a visualization tool is important for users to perform tasks and understand connections and relationships in the data.

Analytical knowledge is defined as the knowledge of different automated data analysis methods, for example clustering (e.g., k-means) or data aggregation (e.g., dimensionality reduction). In certain cases, users need to have at least a basic understanding of their characteristics in order to choose or parameterize them correctly.

4.2 HOW is visualization onboarding provided?

Type, context sensitivity, and interaction are relevant aspects of the question of how visualization onboarding is provided. The *type*, as described in Section 4.2, captures the design and style of onboarding concepts. The form of *contextual aid* is extremely important for applications [24]. The help system should be designed to guide users by demonstration in the *context* of their own interface. Chilana et al. [15] developed an approach to provide a new framework for integrating crowd-sourced contextual help into web applications. In their work, they also discussed the importance of contextual help and beyond *adaptive help* systems. Based on these results, we also integrate the aspect of *context sensitivity* into our framework for visualization onboarding. Fernquist et al. [21] introduced a set of the most relevant aspects for interactive tutorials for a sketching software. Based on their design space for sketching software, we adopted the aspect of *interactivity*.

Type: Onboarding can be provided in different types, such as

guided tours, step-by-step wizards, video-based tutorials, help centers, and overlays. We derived this terminology from our literature review and Pronovic's blog article about context-sensitive and embedded help formats [16].

Context Sensitivity: Context-sensitive help provides assistance at a specific point in the current state of the tool. It is the smallest possible chunk of information the user needs to understand at that point. Examples are help centers, guided tours, or mouseover pop-ups. A type of context-sensitive help is **embedded help** which goes beyond basic information and explanations by either detecting the user's need for help or offering a guided tour right on the interface. Examples are tooltips, instructions on the interface, or walkthroughs.

Context-free help can be called at any state of usage and does not relate to the current state of help-seeking. Examples are online documentations and help videos.

Interaction: Interaction is applied within the onboarding process itself. We refer to Fernquist [21] for defining the degree of interactivity in onboarding concepts. Help systems can be **passive** if the user only consumes the learning material, such as reading an article or viewing a video. If users can try out the concepts, the onboarding concept is defined as **active**. Active tutorials that are aware of the users' interactions and can respond to these are referred to as **reactive**.

4.3 WHERE is visualization onboarding provided?

Based on Fernquist et al. [21] who introduced a set of the most relevant aspects, we also adopted the aspect of the integration of onboarding concepts by asking *Where is visualization onboarding provided?* - externally, internally, or as a learning environment. An onboarding system that is integrated **internally** into the visualization can be more helpful for users because they do not have to jump back and forth between two different systems. **External** sources for onboarding concepts can be defined as sources which can be reached independently of the current state of the tool. At the tightest level of integration, help systems can be provided **internally**. Special learning environments which are not internal or external are defined as **learning environments**, e.g., small games that users may play before using an application in order to become familiar with the basic concepts used. It should be pointed out, however, that integrating onboarding systems into the visualization or visual analytics tools is challenging and requires a considerable amount of effort.

4.4 WHEN is visualization onboarding used?

The aspect of *WHEN* describes the temporal aspect of intended onboarding use (see Figure 1 (6)). Onboarding concepts can be integrated **before** using the actual visualization tool or called up **during** the use of a certain tool, e.g., when support regarding a particular feature is needed. In addition, some onboarding concepts are designed and integrated for the **first use**.

5 DISCUSSION

Based on the results gathered in the course of this work, we have identified a number of challenges.

5.1 Educational Theory

In our literature review we also recognized a theoretical grounding of design decisions based on educational theories, for example concreteness fading by Alper et al. [1] or Kwon and Lee [38] using a learning by doing approach. Ruchikachorn et al. [48] integrated a morphing approach to explain the visual encoding of visualization techniques by analogy.

Ways to effectively support the learning process of users with different knowledge gaps can be considered by using educational theories. However, the literature lacks educational theories with a special focus on onboarding concepts. Nevertheless, existing theories and results of educational research can be used to inform the

design of onboarding systems. Onboarding systems can either be designed like help systems, which implies a cognitivist approach, or they might use a scaffolding approach, applying features such as prompts, tools to structure information or higher-order questions. Constructivist theory supports the assumption that especially higher-order reasoning processes and the ability to make inferences and draw conclusions from the data are supported by cognitive apprenticeship or scaffolding in particular. Higher-order reasoning is the last stage in the model suggested by graph comprehension but also the ultimate goal of most visualization systems. Based on educational theories and graph comprehension, we assume that approaches such as scaffolding are the most appropriate for onboarding. Nevertheless, such assumptions have to be tested by empirical research.

5.2 Constructive Visualization

Nolan [44] published an article discussing how to make statistical graphics a more prominent element of undergraduate statistics curricula. They developed graphics assignments for teaching data visualization and computing. Their teaching concept includes having students deconstruct and reconstruct plots, copy high-quality graphs, create one-minute visual revelations, convert tables into “pictures”, and develop interactive visualizations. They [44] find out that “*The process of identifying and deconstructing a problematic graph and then reconstructing it in a more appropriate form can help students to better interpret, critique, and construct meaningful graphics.*” [44].

The study of Nolan [44] showed that the concept of constructing visualizations can be beneficial for a better understanding and interpretation of the given visualization. In addition, other researchers are beginning to explore the phenomenon of constructive visualizations besides education. Huron et al. [31] have presented a constructive visualization paradigm which offers the possibility of providing people with the means to construct visualizations of their own in a simple, expressive and flexible way. Based on this paradigm, Huron et al. [32] conducted a study and asked people to create, update, and explain their own visualizations using only tangible building blocks. They found out that the visual mapping process is deconstructed into three high-level activities: construction, computation, and storytelling. These were composed of several logical tasks: loading data, building constructs, combining constructs, extending, correcting, aggregating, categorizing, computing new values, and unitizing. The storytelling component was composed of highlighting and marking.

The results and observations of these studies [31, 32, 44] should be considered when designing onboarding concepts for visualization tools. A constructing and deconstructing approach seems promising.

5.3 Storytelling

Storytelling is an effective way of conveying information and knowledge [41]. In the field of visualization, storytelling is rapidly developing technique that enhances understanding [2, 39, 51]. Therefore, this approach sounds promising for the design of visualization onboarding methods. Bach et al. [2] introduced data comics for data-driven storytelling to communicate insights of the data through visualizations. Especially, design pattern for the visual encoding (e.g. build-up, legends, and annotated transitions) as they describe explanations helping readers to understand the visualization can be used as a basis for our design of visualization onboarding flows. Additionally, we need to think about ways to integrate onboarding concepts in visualization tools by using different types of narrative structures, such as the “martini-glass structure” [51]. For example, first the user can be introduced to the visual encoding, data, interaction concepts, and the key insights to be drawn from the data (the stem of the glass), and then opening up for free exploration of the onboarding elements (the body of the glass).

6 CONCLUSION

In this work, we proposed the design space of visualization onboarding. This conceptual framework contains the six aspects: **WHY** is visualization onboarding needed? **WHAT** is visualization onboarding? **WHO** is the user? Which knowledge gap does the user have? **HOW** is visualization onboarding provided? **WHERE** is visualization onboarding provided? **WHEN** is visualization onboarding used? The results of our first literature review show that no common concepts are available for an effective onboarding process for interactive VA tools at the moment. We found very little literature concerning the question of appropriate educational theories for visualization onboarding, therefore more research is needed in order to identify useful educational theories. The concept of visualization onboarding is increasingly relevant and should be emphasized more in the future.

REFERENCES

- [1] B. Alper, N. H. Riche, F. Chevalier, J. Boy, and M. Sezgin. Visualization literacy at elementary school. In *Visualization Literacy at Elementary School*, pp. 5485–5497. ACM, 2017. doi: 10.1145/3025453.3025877
- [2] B. Bach, Z. Wang, M. Farinella, D. Murray-Rust, and N. Henry Riche. Design patterns for data comics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 38:1–38:12. ACM, event-place: Montreal QC, Canada. doi: 10.1145/3173574.3173612
- [3] R. S. Baker, A. T. Corebett, and K. R. Koedinger. Toward a model of learning data representations - eScholarship. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 23(23), 2001.
- [4] K. Balboni. We categorized 327 user onboarding experiences into 9 UI patterns (Find out which you should use) - www.appcues.com. <https://www.appcues.com/blog/user-onboarding-ui-ux-patterns>, 2016. Accessed: 2019-07-09.
- [5] N. Banovic, T. Grossman, J. Matejka, and G. Fitzmaurice. Waken: Reverse engineering usage information and interface structure from software videos. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, UIST '12, pp. 83–92. ACM, 2012. doi: 10.1145/2380116.2380129
- [6] L. Bergman, V. Castelli, T. Lau, and D. Oblinger. DocWizards: A system for authoring follow-me documentation wizards. pp. 191–200. ACM, 2005. doi: 10.1145/1095034.1095067
- [7] K. Börner, A. Bueckle, and M. Ginda. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, 116(6):1857–1864, 2019. doi: 10.1073/pnas.1807180116
- [8] K. Börner, A. Maltese, R. N. Balliet, and J. Heimlich. Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. *Information Visualization*, 15(3):198–213, 2016. doi: 10.1177/1473871615594652
- [9] J. Boy, R. A. Rensink, E. Bertini, and J. D. Fekete. A principled way of assessing visualization literacy. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1963–1972, 2014. doi: 10.1109/TVCG.2014.2346984
- [10] A. Cairo. *The Functional Art: An introduction to information graphics and visualization*. New Riders, 2013.
- [11] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc., 1999.
- [12] J. M. Carroll. *Interfacing thought: Cognitive aspects of human-computer interaction*. The MIT Press, 1987.
- [13] D. Ceneda, T. Gschwandtner, T. May, S. Miksch, H.-J. Schulz, M. Streit, and C. Tominski. Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics (VAST'16)*, 23(1):111–120, 2017. doi: 10.1109/TVCG.2016.2598468
- [14] F. Chevalier, N. H. Riche, B. Alper, C. Plaisant, J. Boy, and N. Elmqvist. Observations and reflections on visualization literacy in elementary school. *IEEE Computer Graphics and Applications*, 38(3):21–29, 2018. doi: 10.1109/MCG.2018.032421650

- [15] P. K. Chilana, A. J. Ko, and J. O. Wobbrock. LemonAid: selection-based crowdsourced contextual help for web applications. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, p. 1549. ACM Press, 2012. doi: 10.1145/2207676.2208620
- [16] An overview of context-sensitive and embedded help formats. <https://pronovix.com/blog/overview-context-sensitive-and-embedded-help-formats>, 2019. Accessed: 2019-03-14.
- [17] S. dos Santos and K. Brodlie. Gaining understanding of multivariate and multidimensional data through visualization. *Computers & Graphics*, 28(3):311–325, June 2004. doi: 10.1016/j.cag.2004.03.013
- [18] T. M. Duffy, J. Lowyck, and D. H. Jonassen. Designing environments for constructive learning. Nato ASI Subseries F.: Springer-Verlag, 1993.
- [19] P. A. Ertmer and T. J. Newby. Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 26(2):43–71, 2013. doi: 10.1002/piq.21143
- [20] P. Federico, M. Wagner, A. Rind, A. Amor-Amors, S. Miksch, and W. Aigner. The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics. In *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 92–103. IEEE, 2017. Projekt: KAVA-Time. doi: 10.1109/VAST.2017.8585498
- [21] J. Fernquist, T. Grossman, and G. Fitzmaurice. Sketch-sketch revolution: An engaging tutorial system for guided sketching and application learning. pp. 373–382. ACM, 2011. doi: 10.1145/2047196.2047245
- [22] S. N. Friel, F. R. Curcio, and G. W. Bright. Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2):124–158, 2001. doi: 10.2307/749671
- [23] T. Grossman and G. Fitzmaurice. ToolClips: An investigation of contextual video assistance for functionality understanding. pp. 1515–1524. ACM, 2010. doi: 10.1145/1753326.1753552
- [24] T. Grossman, G. Fitzmaurice, and R. Attar. A survey of software learnability: Metrics, methodologies and guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '09*, pp. 649–658. ACM, 2009. doi: 10.1145/1518701.1518803
- [25] G. Hart. The five W's: An old tool for the new task of task analysis. *Technical communication*, 43(2):139–145, 1996.
- [26] G. Hart. The five ws of online help systems, 2002. <http://www.geoffhart.com/articles/2002/fivew.htm>.
- [27] K. Higgins. Engaging new users: Guided Interaction. <http://www.kryshiggins.com/guided-interaction/>, 2015. Accessed: 2019-07-09.
- [28] K. Higgins. Evaluating onboarding experiences. <http://www.kryshiggins.com/evaluating-your-new-user-experience/>, 2016. Accessed: 2019-07-09.
- [29] D. Holton and D. Clarke. Scaffolding and metacognition. *International Journal of Mathematical Education in Science and Technology*, 37(2):127–143, 2006. doi: 10.1080/00207390500285818
- [30] S. Hulick. The elements of user onboarding. <https://www.useronboard.com/user-onboarding-ux-design/training/>, 2019. Accessed: 2019-07-09.
- [31] S. Huron, S. Carpendale, A. Thudt, A. Tang, and M. Mauerer. Constructive visualization. In *Proceedings of the 2014 Conference on Designing Interactive Systems, DIS '14*, pp. 433–442. ACM, 2014. doi: 10.1145/2598510.2598566
- [32] S. Huron, Y. Jansen, and S. Carpendale. Constructing visual representations: Investigating the use of tangible tokens. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2102–2111, 2014. doi: 10.1109/TVCG.2014.2346292
- [33] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann. *Mastering the information age: solving problems with visual analytics*. Eurographics Association, 2010. OCLC: 835305616.
- [34] C. Kelleher and R. Pausch. Stencils-based tutorials: Design and evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05*, pp. 541–550. ACM, 2005. doi: 10.1145/1054972.1055047
- [35] A. Kirk. *Data Visualisation: A Handbook for Data Driven Design*. Sage Publications Ltd, 2016.
- [36] H. J. Klein, B. Polin, and K. L. Sutton. Specific onboarding practices for the socialization of new employees. *International Journal of Selection and Assessment*, 23(3):263–283, 2015. doi: 10.1111/ijsa.12113
- [37] A. Kumar. Why user onboarding is essential for your application?, 2017. Accessed: 2018-08-21.
- [38] B. C. Kwon and B. Lee. A comparative evaluation on online learning approaches using parallel coordinate visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '16*, pp. 993–997. ACM Press, 2016. doi: 10.1145/2858036.2858101
- [39] B. Lee, N. H. Riche, P. Isenberg, and S. Carpendale. More than telling a story: Transforming data into visually shared stories. *IEEE computer graphics and applications*, 35(5):84–90. doi: 10.1109/MCG.2015.99
- [40] S. Lee, S.-H. Kim, and B. C. Kwon. VLAT: Development of a visualization literacy assessment test. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):551–560, 2017. doi: 10.1109/TVCG.2016.2598920
- [41] E. M. Lidal, M. Natali, D. Patel, H. Hauser, and I. Viola. Geological storytelling. *Computers & Graphics*, 37(5):445–459. doi: 10.1016/j.cag.2013.01.010
- [42] J. Matejka, T. Grossman, and G. Fitzmaurice. Ambient help. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11*, pp. 2751–2760. ACM, 2011. doi: 10.1145/1978942.1979349
- [43] T. Munzner. A nested process model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics (InfoVis '09)*, 15(6):921–928, 2009. doi: 10.1109/TVCG.2009.111
- [44] D. Nolan and J. Perrett. Teaching and learning data visualization: Ideas and assignments. *The American Statistician*, 70(3):260–269, 2016. doi: 10.1080/00031305.2015.1123651
- [45] S. Palminter and J. Elkerton. An evaluation of animated demonstrations of learning computer-based tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '91*, pp. 257–263. ACM, 1991. doi: 10.1145/108844.108906
- [46] M. Pohl, M. Smuc, and E. Mayr. The user puzzle - explaining the interaction with visual analytics systems. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2908–2916, 2012. doi: 10.1109/TVCG.2012.273
- [47] S. Pongnumkul, M. Dontcheva, W. Li, J. Wang, L. Bourdev, S. Avidan, and M. F. Cohen. Pause-and-play: Automatically linking screencast video tutorials with applications. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*, pp. 135–144. ACM, 2011. doi: 10.1145/2047196.2047213
- [48] P. Ruchikachorn and K. Mueller. Learning visualizations by analogy: Promoting visual literacy through visualization morphing. *IEEE Transactions on Visualization and Computer Graphics*, 21(9):1028–1044, 2015. doi: 10.1109/TVCG.2015.2413786
- [49] R. K. Sawyer and J. G. Greeno. Situativity and learning. In M. Aydede and P. Robbins, eds., *The Cambridge Handbook of Situated Cognition*, pp. 347–367. Cambridge: Cambridge University Press, 2009.
- [50] H.-J. Schulz, M. Streit, T. May, and C. Tominski. Towards a characterization of guidance in visualization. In *Poster Compendium of the IEEE Conference on Information Visualization (InfoVis '13)*. IEEE, 2013.
- [51] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148. doi: 10.1109/TVCG.2010.179
- [52] P. Sukaviriya and J. D. Foley. Coupling a UI framework with automatic generation of context-sensitive animated help. In *Proceedings of the 3rd Annual ACM SIGGRAPH Symposium on User Interface Software and Technology, UIST '90*, pp. 152–166. ACM, 1990. event-place: Snowbird, Utah, USA. doi: 10.1145/97924.97942
- [53] Y. Tanahashi, N. Leaf, and K. Ma. A study on designing effective introductory materials for information visualization. *Computer Graphics Forum*, 35(7):117–126, 2016. doi: 10.1111/cgf.13009
- [54] J. van Wijk. Views on visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):421–432, 2006. doi: 10.1109/TVCG.2006.80
- [55] M. A. Yaln. A systematic and minimalist approach to lower barriers in visual data exploration, 2016.