Supporting Domain Characterization in Visualization Design Studies With the Critical Decision Method

Lena Cibulski* Fraunhofer IGD

Evanthia Dimara† Utrecht University

Setia Hermawati‡ The University of Nottingham

Jörn Kohlhammer§ TU Darmstadt, Fraunhofer IGD

Abstract

While domain characterization has become an integral part of visualization design studies, methodological prescriptions are rare. An underrepresented aspect in existing approaches is domain expertise. Knowledge elicitation methods from cognitive science might help but have not yet received much attention for domain characterization. We propose the Critical Decision Method (CDM) to the visualization domain to provide descriptive steps that open up a knowledge-based perspective on domain characterization. The CDM uses retrospective interviews to reveal expert judgment involved in a challenging situation. We apply it to study three domain problems, reflect on our practical experience, and discuss its relevance to domain characterization in visualization research. We found the CDM's realism and subjective nature to be well suited for eliciting cognitive aspects of high-level task performance. Our insights might guide other researchers in conducting domain characterization with a focus on domain knowledge and cognition. With our work, we hope to contribute to the portfolio of meaningful methods used to inform visualization design and to stimulate discussions regarding prescriptive steps for domain characterization.

Index Terms: Human-centered computing—Visualization—Visualization application domains; Human-centered computing—Visualization—Visualization design and evaluation methods

1 Introduction

To design visualizations for solving real-world problems, we typically conduct design studies [44]. In a design study, we engage closely with domain experts and their data analysis problems. The very first step of a design study is the *domain characterization*. Domain characterization aims at a broad examination of the needs, tasks, and goals of domain experts as well as the conditions and constraints that will frame the visualization use [44]. Domain characterization is conducted *before* the design stage, and its outcome is critical for all subsequent layers of the visualization design process [33]. If we want to design useful problem-driven visualizations, we need to know how to conduct domain characterization effectively.

However, structured guidance on how to conduct domain characterization is scarce. Munzner observed that "hardly any papers devoted solely to analysis at this level [domain characterization] have been published in venues explicitly devoted to visualization" [33]. Marai states that "although visualization design models exist [...], these models do not present a clear methodological prescription for the domain characterization step" [23]. She proposed an actionable framework for domain characterization, which centers around activities and tasks. Yet, we still lack explicit guidance on how to extract the expertise and experience involved with problem-solving.

*e-mail: lena.cibulski@igd.fraunhofer.de

†e-mail: evanthia.dimara@gmail.com

‡e-mail: setia.hermawati@nottingham.ac.uk

§e-mail: jörn.kohlhammer@gris.tu-darmstadt.de

To convey the intended message with a visualization, we need to trace how an expert applies domain knowledge to interpret the depicted information [8]. This understanding is difficult to obtain. Domain expertise and experience rely heavily on personal *tacit knowledge* [38, 49], which involves contextual implications, analogies, or judgments of typicality. Unlike explicit knowledge that has been verbalized, written down, or stored in a database, tacit knowledge is hidden in users' minds. It cannot be derived from observable behavior and users find it hard to articulate how they do something that is based on expertise [38]. Thus, tacit knowledge can only be acquired from humans through their cognitive processes [12]. While knowledge elicitation approaches [32] from cognitive science might help in this context, more effort is needed to study their relevance for domain characterization in visualization design studies.

In this paper, we build on knowledge elicitation approaches to contribute to the transition from existing high-level frameworks towards prescriptive steps for domain characterization (Figure 1). Toward this goal, we make the following contributions:

- We propose a knowledge elicitation method called *Critical Decision Method* (CDM) from cognitive science [15] to the visualization domain.
- We reflect on our experience with applying CDM for characterizing domain problems in three different applications.
- We discuss and advocate its relevance for domain characterization in the context of problem-driven visualization research.

2 Related Work: Domain Characterization in Visualization Research

A *design study* is a project where visualization researchers analyze a real-world problem in a target domain, design a validated visualization solution for it, and reflect on lessons learned to refine guidelines [44]. For this, researchers need to cooperate closely with data expert users who have deep knowledge in the target domain [48].

In the initial *domain characterization* stage, also referred to as *context of use* [16], *task elicitation* [34], or *domain problem characterization* [33], visualization researchers aim to understand the data, tasks, problems, and experiences of the domain experts. The outcome of domain characterization is often implicitly defined and includes data-user-task definitions [30], dominant concepts as a result of coding [46], or design implications [20]. Translated to domain-agnostic abstractions, they guide the visual encoding decisions in the subsequent steps of the design study [33].

Such guidelines depend on the quality of the domain characterization *activities* performed by visualization researchers. They typically read about the domain [9, 22, 27, 41], gain experience with domain tools and data themselves [2, 25, 43], or perform a variation of "talking with and observing domain experts" [44]. Engaging with domain experts includes interviews and real-world observations [3, 21], contextual inquiries [22, 41, 42], focus groups [41, 42], or workshops [13]. To understand more concretely what those activities entail, we reviewed around 30 papers on problem-driven visualization research (including the design studies reviewed by Sedlmair et al. [44]).

Most domain characterization reports focus on the outcome of the activities, i.e., data and task abstractions, rather than how they arrived there. Descriptions of the process remain fairly high-level.

They range from ”interviews with application experts” [37] over having ”interviewed experienced assay developers” [39] to having observed ”real model developers” [3], ”the domain expert on a real-world use case” [9], or ”daily work practices” [40]. Some reports are even limited to having collaborated with domain experts over a certain time period [4, 26, 28]. Reports sometimes include the type of interviews (e.g., guided [43], semi-structured [21, 41], or unstructured [47]) as well as the researchers' topics of interest (e.g., workflows [1, 9], data and analysis methods [27], or tools [25, 43]). Few works explain the interview contents, goals, and subsequent data analysis steps in detail [11, 21, 22, 42]. Few also provide rich reports of the observation procedures they followed and the qualitative insights they gained [41–43]. Still, it often remains unclear whether these descriptions reflect the individual choices of the researchers in that specific context or an established methodological protocol, which could be re-used beyond the specific design study.

This limited methodological justification, or under-reporting, of the domain characterization stage can be attributed to the lack of accessible methods on how to study an application domain in the context of visualization design studies. Current practices are grounded on the diversity of methodology used to study people, cultures, and habits in ethnography [33]. However, these methodologies have not been developed against the background of data analysis. Despite recent advances in ethnographic methodologies, task taxonomies [5], and analytic question sets [20], we identify a lack of prescriptive steps for visualization researchers to follow in a design study in order to elicit domain knowledge and derive task abstractions. With this work, we attempt to fill that gap by moving from high-level advice towards prescriptive steps for domain characterization.

3 The CDM for Domain Characterization

What makes a domain expert is the experience and expertise she brings to a particular field of application. This domain knowledge largely consists of concepts, contextual information, typicalities, personal beliefs, learnings, and insights that have been internalized over years of working practice: it is *tacit* knowledge [38]. In this section, we propose a method to elicit and exploit such kind of knowledge from cognitive science to the visualization domain.

3.1 Motivation and Background

Different fields of research have evolved around studying internalized knowledge. *Knowledge externalization* [49] aims to convert it to explicit representations, e.g., protocols, that can be reused or shared. Common applications are collaborative sense-making [51] or knowledge-assisted guidance [31]. Articulation is achieved through direct creation of narratives and diagrams, like causal flow charts [50], or indirect inference from user interaction with tools [12]. Psychologists summarize techniques to capture the unobservable knowledge, mental processes, and goals underlying task performance under the term *Cognitive Task Analysis* (CTA) [32]. Visualization researchers have applied CTA for studying larger groups of domain experts, but not in the context of a design study. Dimara et al. used the critical incident technique to survey the software needs of decision-makers in organizations [11]. Parsons et al. asked participants to retell a past design process to survey the situated knowledge applied by data visualization practitioners [36].

However, it remains unclear how these methods can be applied to domain characterization, because they have been developed for a different context. Knowledge externalization targets tacit knowledge that results from working with data rather than domain knowledge. Cognitive Task Analysis methods center around domain knowledge but have not been applied in data analysis settings. To make them actionable for domain characterization, CTA methods need to be translated to the domain of visualization research. Beyond a first step in the human factors domain [14], the suitability of CTA techniques for domain characterization has not yet been investigated. For this purpose, we explore the methodological issues of applying the *Critical Decision Method* (CDM) [18] as a representative of CTA.

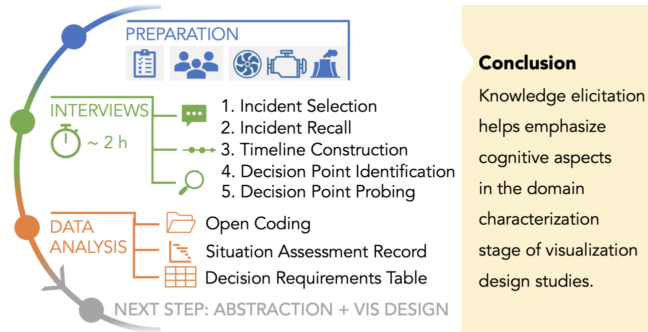

Figure 1: Our CDM procedure for domain characterization. To complete a design study, the data analysis results need to be translated into domain-agnostic abstractions to inform the visualization design.

The Critical Decision Method is a technique to elicit tacit knowledge underlying expert task performance in complex situations [19]. The method grew out of efforts to capture the ”knowledge and experience involved in real-world [...] problem-solving” [15]. It has proven useful to investigate dynamic non-routine situations in diverse domains, such as fire fighting or emergency service. Its effectiveness has been demonstrated for a variety of goals, e.g., to develop support systems, design training material, or establish communication strategies [18]. Although these undertakings did not specifically involve visualization, their variability suggests the CDM's applicability also for domain characterization in visualization research. By providing a step-by-step data collection procedure as well as examples for output representations, the CDM carries the potential to address the lack of formal scripts how to conduct domain characterization and how to represent domain characterization results.

The Critical Decision Method uses semi-structured interviews that are often augmented with observations [18]. Thus, it can be considered another variant of ”talking with and observing domain experts” [44]. In contrast to previous domain characterization practices, however, it offers more prescriptive guidance. The method aims at a systematic *retrospection* of a situation that involved the participant's expert judgment. The CDM is not meant to replace prospective visualization design. Rather, its retrospective nature is well-suited to understand the vocabulary of the target domain [33] without anticipating future design choices. Traditional approaches to domain characterization ask about current problems or envisioned changes or request users to perform artificial tasks. In contrast, walking through a past real-world situation reveals cognitive aspects of the current problem-solving strategies, thus helping assess if and how visualization can involve human expertise to better solve the domain problem. The investigated situation, thus, needs to come from the participant's real-world experience. Different targets like cues, knowledge, options, or experience (see Table 1) are then probed to understand the expert's reasoning during the situation.

3.2 Data Collection Procedure

The preparation phase (Figure 1, blue) is followed by a five-stage data collection (green), where the strategy is to gradually focus on critical cognitive points by sweeping a situation multiple times [18].
1. Incident Selection The goal of the *Incident Selection* stage is to select a task involving competences beyond routine knowledge. The participant should be the primary decision-maker in the situation. To extract true expertise, the task should pose a unique challenge for the participant's competence, i.e., one can expect a difference between the decisions of an expert and those of a novice. It is a pitfall to select a case where participants can rely on formalized procedures.

Table 1: Excerpt of sample questions proposed by Klein et al. [18].

Type	Content
Cues	What were you seeing, hearing, smelling ...?
Knowledge	What information did you use in making this decision, and how was it obtained?
Options	What other courses of action were considered by or available to you?
Experience	What specific training or experience was necessary or helpful [...]?
Basis	How was this option selected [...]? What rule was being followed?
Goals	What were your specific goals at this time?

2. Unstructured Incident Recall The goal of the *Unstructured Incident Recall* stage is to activate the participants' memory and to get a first impression of the scenario. The participant is asked to describe the situation from beginning to end. For example, this might range from loading a data set until an interesting correlation has been found. Interviewers should focus on understanding the story. Interruption for other than minor clarifications is a pitfall.
3. Timeline Construction The goal of the *Timeline Construction* stage is to establish a common understanding among interviewers and participant. From what they heard, the interviewers reconstruct the situation in the form of a timeline. It contains the sequence and duration of events. An event can be an occurrence (like a data point becoming highlighted or a view being switched) or subjective thoughts reported by the participant (e.g., "I would consider this point an outlier"). The timeline is then retold to the participant to identify inconsistencies, add clarifications, and fill in missing details.
4. Decision Point Identification The goal of the *Decision Point Identification* stage is to identify decision points in the timeline for a detailed investigation. The interviewers extract those moments where different ways to understand the situation existed or multiple courses of action were possible. Some are obvious from verbal cues (e.g., "I had to decide whether to include this predictor in the selection"). Others involve taking one of multiple courses of action (e.g., looking at one part of the data first) or making a judgment that affects the action (e.g., "this shape looks like an anomaly but we can safely ignore it"). The granularity of decision points can be adapted.
5. Decision Point Probing The goal of the *Decision Point Probing* stage is to better understand the meaning of information for the participant's assessment of the situation. The interviewers work through the decision points and ask for elaboration. Different probes can be applied for this purpose based on the interviewers' interest. Table 1 lists example questions about how cues, prior knowledge, or different options influenced the participant's course of action.

An interview is expected to last about two hours. This can vary according to the application (timeline construction might be replaced by observation or decision point probing might take place during breaks). Klein et al. recommend to share the interviewing responsibilities among two interviewers and to record the sessions [18].

3.3 Data Analysis and Output

The Critical Decision Method does not prescribe a data analysis approach (Figure 1, orange) because it depends on the research questions motivating the undertaking. In general, coding is used to prepare the ground for converting the interview data to different representations that describe domain knowledge, reasoning, and task activity [15]. Klein et al. present representations that worked well for their applications [18]. We highlight two of these artifacts that we found to particularly match the purpose of domain characterization.

A *situation assessment record* (SAR) reflects the expert's understanding of the dynamic evolution of an incident. It specifies the turning points of the situation together with their underlying cues, experience, knowledge, goals, and actions. Klein et al. propose different formats for SARs [18]. Table 2 shows an SAR for water turbine design: an existing turbine is analyzed to derive potential directions for improved running behavior. The granularity of entries

Table 2: Situation assessment record (SAR) for water turbine design.

Situation Assessment 1	Plausibility check
Cues	Deviation from onsite or testbed measurements
Goals	Validate simulation model
Decision Point 1	Calibrate model or directly proceed
Situation Assessment 2	Analysis of current setup (shift)
Cues	Pressure distribution on blades (heatmap), key performance indicators (e.g., torque and power output)
Experience	Problems in previous operation of the turbine (e.g., cavitation at leading edge)
Goals	Understand strengths and weaknesses of existing turbine, identify potential directions of improvement
Decision Point 2	Address cavitation on current turbine blades
Situation Assessment 3	Optimize blade geometry (shift)
Cues	High curvature in pressure line
Knowledge	Correlates with high blade angle change
Experience	Avoid by shifting camber towards leading edge
Options	Change blade angles or meridional length
Goals	Achieve constant pressure change
Decision Point 3	Shift leading edge towards inlet
Situation Assessment 4	Further optimization (elaboration)
Cues	Pressure and velocity distribution (heatmap, streamlines, sweeping plane), key performance indicators, pressure and blade angle across blade length (line chart)
Basis	Trial-and-error exploration of design space, at each step analyze what went right/wrong and how to improve
Knowledge	Flow behavior, relationships between parameters and side effects, hard constraints (e.g., construction volume)
Experience	Dependencies in previous projects, operating permit requires trading 2-3% less efficiency for fish-friendliness, operating conditions might change during the project
Goals	Understand how geometry affects water flow, trade off efficiency and fish-friendliness
Decision Point 4	Proceed with most-preferred turbine design
Situation Assessment 5	Improvement potential (shift)
Cues	Efficiency curves of new design vs. existing turbine
Experience	Desired water flow and pressure lead to high efficiency
Goals	Predict savings/earnings for customer
Decision Point 5	Implement chosen design

can be adapted to the researchers' needs. New events or insights cause the expert to abandon prior goals and prioritize new goals (shift). For example, identifying the cavitation on the turbine blade as the major problem changes the engineer's goal from analyzing the existing turbine towards optimizing the blade geometry (Table 2, SA 3). Sometimes the goals are maintained but new information enhances what was originally known (elaboration). For example, the cavitation happens at the leading edge of the blade (Table 2, SA 4).

A *decision requirements table* contains details on the judgments that were involved in performing the observed task. The columns specify what particular decisions were made, why they were difficult to make, how they were made, and what supporting information was used. The rows correspond to the decision points identified in the situation assessment record. In this way, the decision requirements complement the experience, goals, etc. in the SAR. For example, the difficulty to investigate multiple operating points (Table 3, B2) effectively extends the description of SA 2 in the SAR. The prescribed structure of Table 3 allows for a comparison even across situations.

4 Case Study

This section will present the CDM steps we followed in conducting three domain characterizations to understand how engineers approach optimization problems that are based on simulation data. Some aspects of the CDM study have already been published in a conference paper [14]; however, the data analysis covered a different purpose. While the previous work investigated the feasibility of the CDM from a human factors point of view, this work reflects on the methodology from a visualization design study perspective.

Table 3: Decision requirements table for water turbine design. The rows correspond to the decision points identified in the SAR (Table 2).

	What is the decision? (A)	Why is it difficult? (B)	How is it made? (C)	What is the aid? How does it help? (D)
1	Determine plausibility of simulation results	Simulation model uncertainties, inaccurate real-world measurements, tolerable deviation not known a priori	Compare real-world measurements to simulated values	n.a.
2	Analyze strengths and weaknesses of current turbine	Only one operating point can be investigated at once	Model and simulate the existing turbine and investigate its performance	3D visualization of pressure/velocity/etc. distribution across geometry hints at inconsistencies
3	Optimize blade geometry	Non-linear relationship between geometry and water flow	Identify non-optimal pressure distribution across blade, change blade geometry and observe effect on distribution	2D line charts that convey the pressure distribution and blade angle across the blade length, high curvature hints at non-optimal geometry
4	Choose most-preferred turbine design	Performance increase in one operating point might come with reduced performance in another operating point, manual exploration of design space, intuition difficult to formalize	Trial-and-error approach, iterative manual changes of the geometry and direct observation of performance changes	2D line charts and 3D visualization of output parameters convey distribution across geometry, color encoding draws attention to peaks, instant feedback helps to see the relationship between geometry changes and flow changes, knowledge of turbine designer hints at correct direction to improve
5	Assess improvement potential of chosen option	Affected by entire operating range, hard to predict savings	Compare chosen option to existing design, critical cue is efficiency	Superimposed efficiency curves of both designs enable direct comparison

4.1 Application Background

We studied the current domain practices in three different applications from the field of engineering design: optimizing a water turbine, an electric drive, and the operation modes of a power plant. The expectations regarding engineered systems are constantly rising: customers ask for high-quality products that are available at little cost and in a short time frame. As such, all three applications deal with multi-attribute decision-making, a core goal of visualization [10], which is challenging because rationality is often complemented with intuition [17, 36]. We were interested in the experts' mental processes and domain knowledge involved with trading off multiple criteria during optimization. The gained insights might inform the design of a visualization that supports the experts in navigating the design space and applying their expertise for trade-off strategies.

We first provide context for each application. Then, we detail our realization of the CDM. Explanations applying to all studies are accumulated. Where relevant, we explicitly differentiate the studies.

Water Turbine Design The turbine converts water flow into electric energy. An existing turbine is to be optimized regarding its running behavior given dynamic operating conditions like water throughput. A typical problem is cavitation on the turbine blades that is caused by low water pressure and can lead to serious mechanical damage. The optimization is characterised by repeated geometry changes followed by an exploration of their effects until the engineer is satisfied. This process largely relies on the experience of the designer regarding how the turbine geometry interacts with the water flow.

Power Plant Optimization A thermal power plant burns fuel to convert the heat energy into electric energy. Changing the fuel type without further adjustments might lead to fuel remaining unburned. Operators thus perform an iterative design space exploration to find an operation mode (e.g., input temperature, valve and damper positions) that reduces the amount of unburned fuel while maintaining operational characteristics like nitrogen oxide emissions and exit temperature. A challenge are coupled physical effects such that eliminating one problem might cause an unexpected problem elsewhere.

Electric Drive Design Optimizing electric drives means to specify their geometry, material, winding patterns, etc. such that their performance optimizes given requirements like cost-efficiency, durability, or construction volume. For this, the operational behaviors of many different electric drive designs are simulated and genetically optimized. Among the resulting set of solutions, it is the responsibility of the design engineer to choose the most preferred compromise. This is challenging, because the number of Pareto-optimal solutions is often quite large and optimization criteria are typically conflicting.

4.2 Implementation of the CDM Procedure

Following the recommendation by Klein et al. [18], we recorded a 90-minute remote session for each application. A pair of elicitors shared the interview responsibilities. Participant recruitment, ethics approval, and questionnaire preparation have already taken place.

1. Incident Selection All three applications required expertise beyond the general routine knowledge of a competent individual. The participants performed multi-criteria optimization on a daily basis and had ten to fifteen years experience in engineering design. This qualified them as experts [18]. An obvious but critical prerequisite was that participants were willing to share material related to the optimization during the interview. We completed this stage via e-mail to spend the interview time on the actual knowledge elicitation.

2. Unstructured Incident Recall As specified in the original method [18], we requested the participants in all three studies to verbally provide a brief overview of their application. We did not interrupt them and focused on understanding the story. The recounts covered the purpose of the engineered system, the parameters of a design option, and the approach to arrive at a preferred design. Going beyond the original CDM method, we also sorted the reported approaches into a priori, a posteriori, or interactive optimization [29], depending on when in the optimization process the participants articulated their preferences. This helped us anticipate the chronology of the timeline in the next stage. In the water turbine study, the expert made use of her expertise in each iteration of exploring the effects of geometry modification. Thus, it belonged to the interactive methods.

3. Timeline Construction + 4. Decision Point Identification The CDM procedure provides that both stages are performed in parallel [18]. In all three studies, we reconstructed the process of deciding for a preferred design option by having the participants walk us through material related to their applications. This slightly deviated from the original method, where the timeline construction is based on the *Unstructured Incident Recall*. Table 2 shows the evolution of the water turbine design. It included modeling an existing turbine (SA 1), analyzing its performance (SA 2), iteratively modifying the blade geometry and observing its effects (SA 3 + SA 4), and comparing the optimized geometry to the initial one (SA 5). The decision points are also highlighted. Some were obvious from verbalization, e.g., *"now we [...] want to start the simulation"* or *"I can now change the operating point"*. Others included subjective assessments of the simulation results, e.g., *"I see that the flow below the runner is good"*. While the CDM recommends to capture both the sequence and duration of events, we omitted the duration in all three studies, because the optimizations were not time-critical.

5. Decision Point Probing In line with the CDM method, we asked for additional details on some decision points. In all studies, we browsed the sample questions [18] for inspiration (see also Table 1). Given our focus on choosing the most preferred design, we particularly asked about available options and how an option was selected, e.g., "Based on what constraints did you exclude these options?". We also probed for visual cues, e.g., "Where do you see that in the visualization?", and prior knowledge, e.g., "How do you know what parameter to adjust next?", that helped the experts gain insights at each optimization step. For the water turbine study, Table 2 contains the utilized question types for each decision point.

Data Analysis and Output Upon completion of all three interview sessions, we analyzed our protocols and recordings. The CDM recommends any form of coding to get started [15]. In the *Decision Point Identification* and *Decision Point Probing* stages, we already tagged the decision points with their underlying cues, experience, goals, etc. After all sessions had been completed, we coded the common needs, approaches, and difficulties that emerged from the responses of the participants. From the protocols, recordings, and code set, each interviewer further derived two of the proposed CDM outputs [18] per session: a situation assessment record describing the situation as a series of decision points (Table 2) and a decision requirements table specifying the what, how, why, and aid of the particular decision points (Table 3). We performed the qualitative coding [45] and artifact generation independently. We then discussed the results and jointly refined the codes and artifact representations.

4.3 Practical Considerations

This section presents the methodological issues we observed when applying the CDM for the purpose of domain characterization. We illustrate method properties we recognized, domain characteristics that the method helped elicit, and things that worked (less) well.

Qualitative methods are time-consuming and the CDM is no exception. The 90-minute sessions with the experts were followed by a time-intensive data analysis by us as interviewers. Given the contextual richness of the CDM responses, we consider the efforts justified. Our impression is that domain problems that are expected to be cognition-sensitive and could only be described by a series of low-level tasks benefit most from a characterization using the CDM.

We originally performed another CDM session on a signal filter optimization. It failed, because the optimization relied on a routine weighting strategy rather than expertise. We did not realize this until the *Unstructured Incident Recall* stage. The participant was also in an early PhD stage and thus did not have enough expertise. We eventually decided to discard this session from our collection.

In each of the remaining studies, the *Unstructured Incident Recall* stage conveyed an initial idea of the domain problem as intended. The participants could give an overview of their situations on an appropriate level of abstraction, which kept the entry barrier for the subsequent stages low without too many details. Where needed, we adjusted the level of abstraction by asking clarifying questions.

The *Timeline Construction* walk-through, in contrast, contained a lot of meaningful details. Remembering a past situation in detail is difficult and may provoke a mismatch between user recollections and their actual actions [44]. While we did not evaluate this, our participants did not seem to have difficulties with remembering the past incidents. It might have helped their memory that we, aside from the original CDM method, encouraged them to bring documentation material. The water turbine study, in particular, involved technical details, e.g., *"this is a hand-coded mesh generator"*, that sometimes distracted us from the actual reasoning process.

The CDM centers around an incident from the real-world practice of experts, which helps avoid the domain threat of mischaracterizing the problem (cf. the top level of the nested model [33]). An immediate validation of that threat is naturally incorporated in the *Timeline Construction* stage of the CDM procedure: retelling the constructed timeline to the participant. Deviating from the original procedure, we skipped this step, thus missing the chance to validate the constructed timeline and decision points during the interview. In retrospect, we should have followed the original procedure or included a post-interview validation of our results.

How much the CDM can teach researchers about a target domain might depend on the knowledge gap commonly associated with application-driven visualization research [48]. It is not necessarily a drawback: interviewers who know little about a domain tend to probe more. Still, we underestimated the mental effort for decision point identification and probing without much prior knowledge about an incident. We found it difficult to reconstruct the timeline and choose appropriate probes on the fly during the interview session. Thus, we largely relied on questions that we prepared beforehand without knowing the incidents in detail. We also put together the sequence of decision points in retrospective, i.e., upon completion of all three sessions. While it might help to familiarize with the particular incident prior to an interview, it remains an open question how to succeed in spontaneous timeline construction.

Our independent data analysis results showed a broad consensus regarding the content, especially with respect to those moments of an observed situation that we considered decision points. This concurs with Klein et al., who found that inter-observer variability refers to the significance of a decision point rather than its presence [18]. For a subsequent task abstraction this suggests that disagreements between visualization researchers might mainly relate to *why* a task is performed [5]. Although both interviewers based their analysis on the same situation assessment example, we found the resulting records to significantly deviate in their format, i.e. the mapping between decision points and situation assessments as well as their granularity. These deviations propagated to the decision requirements table, because we transferred the decision points from the situation assessment record to the table rows. Explicitly agreeing on a template beforehand might further reduce the risk of discrepancies.

The systematic CDM procedure revealed what types of domain knowledge and expertise the users carry, e.g., *"the designer knows about the parameter options and side effects"* or *"operators usually trade 2% to 3% less efficiency for fish-friendliness"*. We further learned what cues steer their attention, e.g., *"co-occurrence of high oxygen and high temperature"* or *"too much curvature in the pressure lines"* (compare Table 2, SA 3, Cues). The CDM also made explicit how the experts' goals varied with the situation focus, e.g., from understanding the status quo over reducing unburned fuel to maintaining a reasonable cost-benefit ratio. To conclude, by revealing the role of user expertise in task performance, the CDM has the potential to effectively foster the *appropriateness* of a visualization [30], i.e., its benefit for supporting a given task.

5 Implications for Visualization Research

This section reflects on our experiences with the Critical Decision Method more generically against the background of visualization research. We also highlight questions that remain unanswered and how future research can address these open issues.

5.1 Reflections

Explicitly describing tacit knowledge has been an ongoing challenge. We experienced that the CDM comes quite close by producing artefacts that explicitly describe applied expertise, subjective judgments, and contextual effects. This is particularly relevant for visualizations, because they are highly dependent on the goal, task, and context of their usage. The decision points in the situation assessment record (Table 2) translate to individual contexts that might require different visualization designs. The timeline might inspire narratives for downstream validation of a visualization. For example, in a field experiment [6], where realism is manipulated by asking participants to perform specific tasks, the timeline can frame a particular setting.

We experienced that probes like ”what rules did you follow to make this decision?”, or ”what were your specific goals at this time?” helped experts concentrate on what they want to do rather than what a visualization solution might look like. On the other hand, we identified probes for perceptual cues, i.e., ”what were you seeing?” or ”what caught your attention?”, as one possible starting point for deriving design requirements. They revealed characteristics in the data to be emphasized by a future visualization design. Answers to the role of (visual) aids in the decision requirements table (Table 3) also pointed towards potential entry points for visualization support. Similarly, responses to experience and knowledge probes could help design interaction techniques. In combination with perceptual cues, they might also inform the integration of guidance into a visualization system [7]. We did not use probes for hypotheticals (e.g., ”what difference would it make if ...?”) ourselves but we expect them to help anticipate the consequences of different design choices and raise particular awareness for potential pitfalls.

In prior projects where we did not apply the CDM, we asked experts about a visualization’s rather general context of use. In contrast, the comprehensive, yet systematic, procedure of the CDM helped us stay focused on decision points that notably revealed expert knowledge. Its output can pave the way for turning incidents into abstractions and subsequent design choices. More precisely, the CDM helped us identify and describe critical decision points that can be used in subsequent layers of the design process.

Rather than a strict recipe, the CDM can be seen as a framework where implementations can be chosen according to the research objective. It is open to being combined with dedicated requirements engineering techniques from visualization research. Up to now, domain characterization focuses on tasks [33], but less on domain knowledge elicitation. Complementing existing approaches with the CDM allows for a complete picture of the target domain, including both the task-based and the knowledge-based perspectives.

For example, the CDM is one of many ways to implement the *discover* stage of the nine-stage framework [44]. It contributes to preventing pitfalls PF-15 (”ignoring successful aspects”) and PF-17 (”focusing on visualization solutions”). The *Unstructured Incident Recall* with passive interviewers provokes PF-16 (”expecting talking and passive observation alone to work”) at first sight, but additional think-aloud sweeps of the incident compensate this. As a qualitative approach, the CDM also seems to provoke PF-5 (”insufficient time from collaborators”), but it actually makes efficient use of expert time, leaving the time-consuming part to the researchers.

The CDM builds upon a holistic consideration of probe types like cues, goals, or knowledge. They convey a comprehensive picture of cognitive turning points in a domain problem. Although targeted at decision points in the first place, insights gained through these probe types also hold the potential to advance existing task descriptions. Furthermore, in analogy to the multi-level typology of abstract visualization tasks [5], the *what*, *how*, and *why* classification in the decision requirements table (Table 3) can inform a visualization-oriented characterization scheme for decision points. The CDM can then serve as a systematic data collection method to inform the creation of a taxonomy of decision tasks. In a similar way, it could help identify different domain knowledge types and their representations to inform endeavors in knowledge-assisted visualization.

To conclude, any domain characterization technique will highlight some aspects of the problem domain and de-emphasize others. Depending on the research objectives, multiple methods can be combined to arrive at a concise understanding of a target domain. In this sense, the CDM is a valuable addition to the portfolio. Its output in the form of open coding, situation assessment record and decision requirements table provides a good basis for a) discussions and reflections among visualization researchers and domain experts and b) the subsequent definition of abstract tasks, requirements, and mental models to inform the visualization and interaction design.

5.2 Open Issues and Future Research Paths

This work is a first step towards integrating knowledge elicitation from cognitive science into visualization research. Further research is needed to back up our experiences and turn the *Critical Decision Method* into an actionable model for visualization researchers. We raise open issues that can be points of discussion for the workshop.

A first open issue, beyond looking for a connection to abstract visualization tasks [5] discussed in Section 5.1 and generic reasoning aspects, is how the CDM can be concretely applied to problem-driven visualization design. This includes an adaptation of the terminology (e.g. from *incident* to *analysis*), particular probing of aspects related to analysis (e.g. data quality or correlations) and visualization (e.g. correlation or visual representations), and more concrete dedication of each CDM step to the goals of domain characterization. To arrive at a complete domain characterization model, the CDM procedure also needs to be linked more tightly to the subsequent layers of visualization design.

The previous point is also connected to finding the best path to evaluate the effectiveness of the CDM framework. Among others, Marai and Möller propose *significance* and *pragmatic adequacy* as evaluation criteria for theoretic contributions to visualization research [24]. While we hope to have motivated the significance of the CDM to the visualization field, its usefulness for visualization practice is yet to be confirmed. Reporting on a complete design study clarifies whether the knowledge elicited by the CDM actually helps improve the results of subsequent visualization design layers. This might also target professional practitioners as opposed to design studies conducted in visualization research [35]. With a number of observations collected, we might also be able to identify meaningful practices that can serve as guidelines on how to design visualizations that foster the exploitation of human cognition and knowledge for analytic tasks. Validating practical experiences with the CDM should also include data analyses with different high-level tasks. This can even be extended to a set of real-world case studies from different domains that discuss the CDM procedure step by step and compare it to alternative approaches. By replicating domain characterizations from existing design study papers, the previous approaches could be compared to a CDM domain characterization, highlighting the differences that stem from using the CDM. We note that this would mean to compare empirical methodologies rather than techniques, which is not commonly done in the visualization community.

6 Conclusion

Accessible methods on how to conduct domain characterization for visualization design studies are scarce. We reflected on the *Critical Decision Method* (CDM), an interview technique for knowledge elicitation in cognitive science, as a means to study an application domain. We illustrated its systematic procedure, how we implemented it in three domain characterizations, and what methodological issues we encountered when applying it in visualization research.

The CDM provides an alternative way of learning about domain experts and the conditions framing their task performance. Its focus on real-world incidents aligns well with the required realism in tasks, data, and users for understanding work practices [20]. The CDM particularly encourages participants to reflect on their own cognitive processes. It suggests a novel perspective on domain characterization by favoring decision points over tasks. We found it to be a promising way to emphasize cognitive aspects and we hope to have raised interest in devoting closer attention to knowledge elicitation in domain characterization. With our work, we aim to encourage other visualization researchers to use the CDM for their design studies and share their experience.

Acknowledgments

The authors wish to thank the participants of the study. This work was supported by the EC H2020 programme under Grant 768892.

REFERENCES

[1] A. Barsky, T. Munzner, J. Gardy, and R. Kincaid. Cerebral: Visualizing multiple experimental conditions on a graph with biological context. *IEEE transactions on visualization and computer graphics*, 14(6):1253–1260, 2008.

[2] D. Baur, F. Seiffert, M. Sedlmair, and S. Boring. The streams of our lives: Visualizing listening histories in context. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1119–1128, 2010.

[3] S. Bergner, M. Sedlmair, T. Moller, S. N. Abdolyousefi, and A. Saad. Paraglide: Interactive parameter space partitioning for computer simulations. *IEEE Transactions on Visualization and Computer Graphics*, 19(9):1499–1512, 2013.

[4] M. Booshehrian, T. Möller, R. M. Peterman, and T. Munzner. Vismon: Facilitating analysis of trade-offs, uncertainty, and sensitivity in fisheries management decision making. In *Computer Graphics Forum*, vol. 31, pp. 1235–1244. Wiley Online Library, 2012.

[5] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013.

[6] S. Carpendale. Evaluating information visualizations. In *Information Visualization*, pp. 19–45. Springer, 2008.

[7] 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*, 23(1):111–120, 2016.

[8] C. Chen. Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications*, 25(4):12–16, 2005.

[9] L. Cibulski, H. Mitterhofer, T. May, and J. Kohlhammer. Paved: Pareto front visualization for engineering design. In *Computer Graphics Forum*, vol. 39, pp. 405–416. Wiley Online Library, 2020.

[10] E. Dimara and J. Stasko. A critical reflection on visualization research: Where do decision making tasks hide? *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[11] E. Dimara, H. Zhang, M. Tory, and S. Franconeri. The unmet data visualization needs of decision makers within organizations. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[12] P. Federico, M. Wagner, A. Rind, A. Amor-Amorós, S. Miksch, and W. Aigner. The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics. In *IEEE Conference on Visual Analytics Science and Technology*, pp. 92–103. IEEE, 2017.

[13] S. Goodwin, J. Dykes, S. Jones, I. Dillingham, G. Dove, A. Duffy, A. Kachkaev, A. Slingsby, and J. Wood. Creative user-centered visualization design for energy analysts and modelers. *IEEE transactions on visualization and computer graphics*, 19(12):2516–2525, 2013.

[14] S. Hermawati, L. Cibulski, and G. Lawson. Towards using the critical decision method for studying visualisation-based decision-making. In *Contemporary Ergonomics & Human Factors*, pp. 130–136. 2021.

[15] R. R. Hoffman, B. Crandall, and N. Shadbolt. Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis. *Human Factors*, 40(2):254–276, 1998.

[16] ISO. ISO 9241-210:2019 Human-centered design for interactive systems, 2019.

[17] G. A. Klein. *Sources of Power: How People Make Decisions*. MIT Press, 2017.

[18] G. A. Klein, R. Calderwood, and D. MacGregor. Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(3):462–472, 1989.

[19] D. S. M. KOH, H. G. TEE, B. K. SOH, and A. L. S. TAN. Tools for facilitating critical decision method during tacit knowledge elicitation. In *International Conference on Naturalistic Decision Making*, 2015.

[20] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, 2011.

[21] H. Lam, D. Russell, D. Tang, and T. Munzner. Session viewer: Visual exploratory analysis of web session logs. In *2007 IEEE Symposium on Visual Analytics Science and Technology*, pp. 147–154. IEEE, 2007.

[22] D. Lloyd and J. Dykes. Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study. *IEEE transactions on visualization and computer graphics*, 17(12):2498–2507, 2011.

[23] G. E. Marai. Activity-centered domain characterization for problem-driven scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):913–922, 2017.

[24] G. E. Marai and T. Möller. The fabric of visualization. In *Foundations of Data Visualization*, pp. 5–14. Springer, 2020.

[25] P. McLachlan, T. Munzner, E. Koutsofios, and S. North. Liverac: Interactive visual exploration of system management time-series data. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1483–1492, 2008.

[26] M. Meyer, T. Munzner, A. dePace, and H. Pfister. Multeesum: A tool for exploring space-time expression data. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2010)*, 16(6):908–917, 2010.

[27] M. Meyer, T. Munzner, and H. Pfister. Mizbee: a multiscale synteny browser. *IEEE transactions on visualization and computer graphics*, 15(6):897–904, 2009.

[28] M. Meyer, B. Wong, M. Styczynski, T. Munzner, and H. Pfister. Pathline: A tool for comparative functional genomics. In *Computer Graphics Forum*, vol. 29, pp. 1043–1052. Wiley Online Library, 2010.

[29] K. Miettinen. Introduction to multiobjective optimization: Noninteractive approaches. In *Multiobjective Optimization*, pp. 1–26. Springer, 2008.

[30] S. Miksch and W. Aigner. A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics*, 38:286–290, 2014.

[31] S. Miksch, H. Leitte, and M. Chen. Knowledge-assisted visualization and guidance. In *Foundations of Data Visualization*, pp. 61–85. Springer, 2020.

[32] L. G. Militello and R. J. Hutton. Applied cognitive task analysis (acta): a practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, 41(11):1618–1641, 1998.

[33] T. Munzner. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928, 2009.

[34] M. Oppermann and T. Munzner. Data-first visualization design studies. In *2020 IEEE Workshop on Evaluation and Beyond-Methodological Approaches to Visualization (BELIV)*, pp. 74–80. IEEE, 2020.

[35] P. Parsons. Understanding data visualization design practice. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):665–675, 2021.

[36] P. Parsons, C. M. Gray, A. Baigelenov, and I. Carr. Design judgment in data visualization practice. In *IEEE Visualization Conference*, pp. 176–180. IEEE, 2020.

[37] H. Piringer, W. Berger, and J. Krasser. Hypermoval: Interactive visual validation of regression models for real-time simulation. In *Computer Graphics Forum*, vol. 29, pp. 983–992. Wiley Online Library, 2010.

[38] M. Polanyi. The tacit dimension. In *Knowledge in Organizations*, pp. 135–146. Routledge, 2009.

[39] A. J. Pretorius, M.-A. Bray, A. E. Carpenter, and R. A. Ruddle. Visualization of parameter space for image analysis. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2402–2411, 2011.

[40] M. Sedlmair, C. Bernhold, D. Herrscher, S. Boring, and A. Butz. Mostvis: An interactive visualization supporting automotive engineers in most catalog exploration. In *2009 13th International Conference Information Visualisation*, pp. 173–182. IEEE, 2009.

[41] M. Sedlmair, A. Frank, T. Munzner, and A. Butz. Relex: Visualization for actively changing overlay network specifications. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2729–2738, 2012.

[42] M. Sedlmair, P. Isenberg, D. Baur, M. Mauerer, C. Pigorsch, and A. Butz. Cardiogram: visual analytics for automotive engineers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1727–1736, 2011.

[43] M. Sedlmair, B. Kunze, W. Hintermaier, and A. Butz. User-centered development of a visual exploration system for in-car communication. In *International Symposium on Smart Graphics*, pp. 105–116. Springer, 2009.

[44] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on*

Visualization and Computer Graphics, 18(12):2431–2440, 2012.

[45] A. Strauss and J. Corbin. *Basics of qualitative research*. Sage publications, 1990.

[46] M. Tory and S. Staub-French. Qualitative analysis of visualization: A building design field study. In *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, pp. 1–8, 2008.

[47] A. Unger, S. Schulte, V. Klemann, and D. Dransch. A visual analysis concept for the validation of geoscientific simulation models. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2216–2225, 2012.

[48] J. J. Van Wijk. Bridging the gaps. *IEEE Computer Graphics and Applications*, 26(6):6–9, 2006.

[49] X. Wang, D. H. Jeong, W. Dou, S.-w. Lee, W. Ribarsky, and R. Chang. Defining and applying knowledge conversion processes to a visual analytics system. *Computers & Graphics*, 33(5):616–623, 2009.

[50] C.-H. Yen, H. Cheng, G. Y.-C. Yen, B. P. Bailey, and Y. Huang. Narratives+ diagrams: An integrated approach for externalizing and sharing people's causal beliefs. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–27, 2021.

[51] J. Zhao, M. Glueck, P. Isenberg, F. Chevalier, and A. Khan. Supporting handoff in asynchronous collaborative sensemaking using knowledge-transfer graphs. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):340–350, 2017.