How can heuristics be communicated?

Isabel Evans[0000−0003−3913−4463], Chris Porter[0000−0003−3813−2941], and Mark Micallef[0000−0003−4159−6923]

University of Malta, Msida, Malta

Abstract. This position paper proposes a model for choosing how to communicate heuristics in a meaningful, usable and accessible manner. The model is a matrix of nine shapes, where we define 'shape' as a combination of the heuristic's format and its directiveness. Examining heuristics used in UX and software testing practices, we found a variety of formats and levels of directiveness. We discuss these shapes with example heuristics from UX, software testing, and everyday life. We present the outcome from a pilot of the model in one specific context. The shape of a heuristic may contribute to its understandability and usefulness in specific contexts, while different contexts may necessitate different ways of communicating heuristics. This includes considering the effect that the shape of a heuristic has on its accessibility and inclusion, suggesting that shape may be one important aspect of heuristics' design and evaluation.

Keywords: Heuristics · Human computer interaction · Design methodology · User-centred design · Accessibility · Inclusion.

1 Introduction

Heuristics are used to describe ways of thinking, as well as to aid in problem-solving, by providing rules of thumb that, while fallible, provoke thought and as aide-mémoires [13].

Our HCI research, [7–10], into the design of tools to aid software testing has led us to develop guidance for test tool designers[1]. We want to communicate our findings back to industry practitioners in ways that are useful, understandable, flexible and easily applied to their work. As heuristics are already used in a variety of ways within the software industry, it made sense to develop heuristics to capture and communicate our main research findings. In this paper, we describe the work we are doing to clarify how best to *communicate* the heuristics we are developing. Those heuristics are themselves out of this paper's scope.

Before developing our heuristics, we needed to know what would provide a good experience of the heuristics for their specific audience and context, and whether the format and directive level made a difference to how well they were

[1] See Evans, I., Porter, C., I & Micallef, M. (2024). Heuristics-for-test-tool-design (Version 1.0.0) [Computer software]. https://github.com/hci-lab-um/heuristics-for-test-tool-design/blob/main/README.md

understood. This led to our research question: **How can heuristics be communicated?** Through a review of academic, industry and grey literature, we found that heuristics are expressed in different ways. There is variation in the level of direction given by a heuristic and the format in which it is displayed. It was unclear as we reviewed the heuristics whether the choices made for the format or level of direction were based on custom, aesthetics, or other factors.

We set out the background for our work on the shape of heuristics, with a brief history of heuristics and their use in HCI and software testing, as well as other areas. We describe the methods we use, and in our findings, we provide an overview of how heuristics can be developed and evaluated and explain our choice of methodology for designing our own heuristics. We describe example heuristics for the directive levels and formats we identified. We show a matrix of nine possible 'shapes' for a heuristic, based on how directive they are (**Telling**, **Prompting** and **Asking**) and their format (**Textual**, **Pictorial** and **Conceptual**), with examples from software testing, UX practice, road signage and other fields. In a pilot of applying the matrix, we trialled the nine heuristic shapes with industry practitioners in a small specific context, to find the shape most suitable for a specific set of heuristics, and then reviewed with industry experts. We discuss why this matters: the Shape of Heuristics Model helps to improve communication and use of heuristics in general. Our literature review, both academic and industrial, suggests that this has not previously been an overt consideration in developing heuristics, and is worthy of further research. We therefore suggest future work for development and evaluation of the Shape of Heuristics Model, and how we are using it in our work.

2 Background

2.1 Heuristics: History and Meaning

The Oxford English Dictionary [22] notes the first use of the word 'heuristic' in 1770. [13] and [22] show the definition and usages of 'heuristic' have changed over time. Usage though the late 1700's and early 1800's in logic and education was in reference to educational methods that enable students to learn by making discoveries for themselves. Over time, heuristic comes to mean methods in problem solving and decision making [22].

It remains a relatively infrequent word; [22] notes that it occurs about 3 times per million words, making it recognisable but unusual [23]. Its use has increased dramatically since the 1950's with the start of its use in computing and AI. Figure 1 shows the frequency of use of the word heuristic, from nearly zero per million words prior to 1940, rising through the 1940's to 1960's to 0.2 per million words, and rising steeply to 1.2 per million words by 2010. It is now used in disciplines including sociology, computing and artificial intelligence.

Academic work about heuristics has increased over the last 50 years, [13] noted an increase in Scopus database entries with 'heuristic' in the title from 70 papers in 1970 to 3783 in 2021. We made a Google Scholar search for 'heuristic

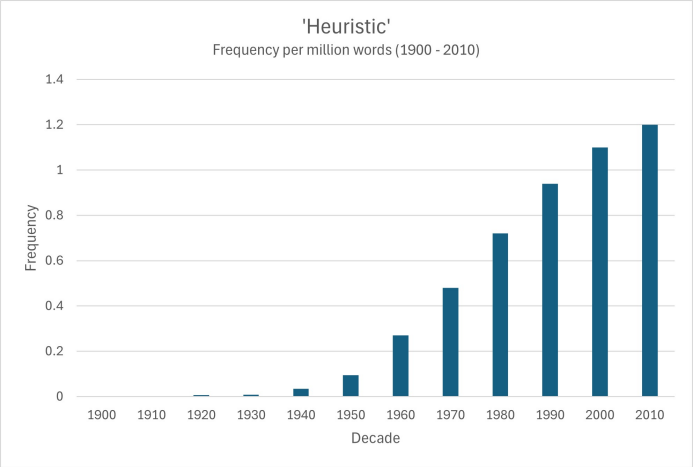

Fig. 1. 'Heuristic' frequency 1900-2010 after [22]

evaluation' showing 482 results between 1950 and 1989 from AI, programming, maths, statistics and psychology. These are papers both about the use of heuristics to evaluate situations and artefacts, and also about the evaluation of heuristics. Between 1989 and 2023 there were 1370 results for 'heuristic + usability' - the majority of these are either in fields outside HCI or are about usability of heuristics. Rather than providing a definition of 'heuristic', [13] discusses their *'defining characteristics'* as *'problem-solving methods that do not guarantee an optimal solution'* which are *'practically applicable in at least a number of interesting cases'*, contrasting a heuristic with what they refer to as *'allegedly rational decision making'* which may take longer and while potentially reaching a more optimal solution, may be impractical.

In this paper we are interested in heuristics used in usability, HCI, UX and software testing, which we discuss next.

2.2 Heuristics in UX for Interface Evaluation

Within the HCI, UX and testing communities. the set of '10 Usability Heuristics for User Interface Design' [20] are known as a method for designing and evaluating user interfaces. These were first presented by [21], following their first mention of Heuristic Evaluation (HE) for usability is in a paper the previous year [19]. Since that paper, these heuristics have been widely adopted, being taught to UX and testing practitioners, for example, [15, 35]. These Ten Heuristics from Nielsen, and variations have been developed and updated to meet both new technologies and specific domains, for example, by [18, 20, 28, 29, 36].

2.3 Heuristics in Software Testing

Heuristics are used by software testers to inform their exploratory and test design work. For example, [5] writes that heuristics are not only our models for thinking but also provide a way of probing and challenging our models and ways of thinking when testing software: *"heuristics are essential tools for thinking test practitioners. ... if you consciously use a model as a heuristic, then you are in a better position to see its weaknesses and potential failure points in a given situation"* [5] Many testing heuristics in a variety of format are available; [4] provides a summary of many heuristics commonly used by testing practitioners to inform their work, including heuristics for test planning, test design and assessing risks.

2.4 Heuristics in Day to Day Life

Heuristics, in a variety of formats, are used in general life. These may be instinctive and based on senses, such as a 'sniff test' for whether food is edible. Additionally, in day-to-day life, we use proverbial rules of thumb which may or may not be up-to-date and well understood, for example 'red sky at night, shepherd's delight'. These may be language, location or culturally specific. Heuristics in everyday life are also designed for specific purposes; examples include information and warning signage, such as no entry signs or hazard signs.

3 Motivation - Research Gap

This position paper is motivated by our current work, which includes the development of evidence-based heuristics that provide design guidance to test tool vendors. The actual guidance is out of this paper's scope. We wanted to design and evaluate evidence-based heuristics through a robust process to evaluate their content. We also wanted to understand how to communicate these heuristics in a meaningful, usable and accessible manner to people who make (design) decisions.

The literature we examined (See Section 5.1) enabled us to apply an iterative methodology for our own heuristics, with design and review/evaluation cycles, and more than one evaluation method. The methodologies showed how to review a heuristic's content, but we could not find work on evaluating the format by which content is communicated. There seemed to be an assumption of heuristics for UX evaluations being communicated as text statements that prompt or direct action, for example [20]. We needed a way to understand what options there are for the way heuristics are communicated and to understand which option is best for our specific heuristics. Our heuristics are intended to aid UX design rather than UX evaluation, and we wanted to see if the text statement was still the best format. We were aware of other heuristic formats in the testing industry, for example mnemonics [4] and of other formats outside IT/software, such as pictorial road signs [31].

We asked the research question: **How can heuristics be communicated?**

4 Methodology

A mixed methods approach was taken for this position paper, including literature and grey literature reviews, thematic and gap analysis of the heuristic formats found, and an expert review in a pilot study, including semi-structured interviews with industry experts and practitioners.

The purpose of the research leading to this position paper is hypothesis building: we examined data in the literature to look for patterns. These patterns led to our proposal of the "Shape of Heuristics" model. We adopted a qualitative approach during a pilot using expert reviews to obtain viewpoints while observing for saturation with a small sample after [12].

Figure 2 maps the route through the methods used to build the hypothesis. Initial data collection was through academic and grey literature reviews, with categorizing and grouping of the heuristics found and a synthesis into a matrix of groups. We searched Google Scholar for papers mentioning heuristics, and from those extracted ones concerned with HCI/UX usability heuristics and software testing heuristics, starting from the current year and working backwards in time to reach Nielsen's first paper 1990 on heuristics [21], as these are significant in both HCI and software testing industrial practice.

Initially, we examined papers on the design of heuristics. We then looked for examples of heuristics formats. We looked at papers about the history, usage, design and evaluation of heuristics. In the grey literature review, we search online for examples of UX and software testing heuristics in industry. These gave us examples of usage and adaptations, as well as of heuristics based on experience rather than a defined development process. The information from all these sources was analysed to identify heuristics development methods (see Section 5.1) and validating heuristics (see Section 5.2).

While doing this analysis we noticed that the heuristics varied both in format and directiveness, especially in the software testing industry sources. We, therefore, reanalysed the data to identify and group the formats and directiveness of the heuristics. We used the outcome of this analysis to build a matrix of format against directiveness and populated each cell with one example.

We then hypothesised that each cell provided a "Shape of Heuristic" that might be suitable for specific contexts, improving their understandability and usefulness.

We piloted that matrix with practitioners in a specific context and also reviewed it with experts from accessibility, UX and software testing industry backgrounds. Pilot methods are covered in Section 6.

5 Findings - Developing the 'Shape of Heuristics' Model

5.1 Review of Designing and Developing Heuristics

We examined literature (academic and industry) in HCI/UX and Software Testing to look at how heuristics are developed. One researcher attended a BCS

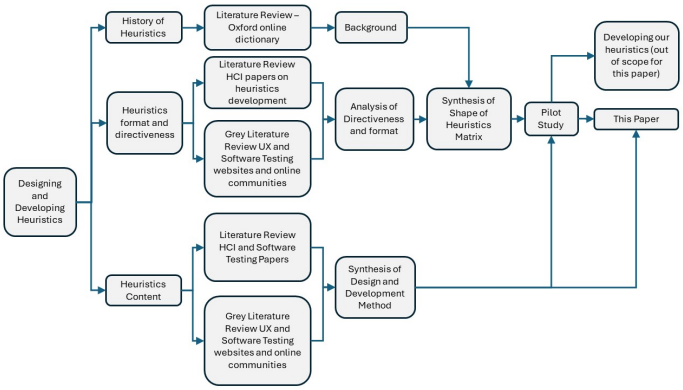

Fig. 2. Methodology for this Paper

seminar on developing heuristics. Within the HCI academic community, several updates to Nielsen's Heuristics set have been published, together with critiques of the original and updated set, and also methodologies for developing and validating HCI heuristics.

Different methodologies for developing heuristics have been published; [29] remarks: *'The relevant literature contains no consensus over the most effective approach to developing domain specific heuristics.'* We identified several ways that heuristics could be developed, with both the design and the evaluation having different levels of rigour and formality.

Instinctively or Intuitively Arrived At: These heuristics are intuited by an individual and used by that individual because 'these work for me'; these heuristics are part of natural world and evolution [13]; they are used by humans and other animals. These heuristics are informal, and not specifically evaluated.

Experience-based Design: These heuristics capture an individual or group's experience in context and may be built and reviewed formally or informally. They may be published informally, perhaps on a blog. They may be changed through usage including but not exclusively the author. For example, [1, 4, 5] all recommend that anyone using testing heuristics also redesign, modify and replace to make their own heuristics to reflect their local context.

Evidence-based Design: Data to build evidence-based heuristics are collected via research and the heuristics themselves are evaluated formally before use. Heuristics may be evaluated by comparison with a known and trusted set such as Nielsen's 10 Heuristics, for example [20] made updates to the Ten Heuristics. Usability heuristics designed to cover changes to technology and specific domains include mobile learning application heuristics, by [18] evaluated against

the Nielsen heuristics. Validation of heuristics for touch screen usability, by [14] used comparison of results with one group of evaluators using Nielsen's Ten Heuristics and another group using their new heuristics.

At a BCS Interaction Group on Research Methods for Human-Computer Interaction[2], Petrie and Sim showed how to design and evaluate heuristics, emphasising the need for a review and evaluation of the heuristics' efficacy. This leads to a methodology where heuristics are designed based on data collected in the field, and then iteratively reviewed and refined through evaluation with experts and practitioners.

5.2 Examples of Validation of Heuristics

In [29] there is a criticism of the approaches used to create heuristics: *'the method used for validating the heuristics. The raw count of the number of usability problems identified may not be an appropriate indicator of the effectiveness of a set of heuristics'*. They suggest criteria to use include thoroughness, correctness, coverage, effectiveness and ease of use of the heuristics, measured in review and usage. This includes a survey of domain users to identify problems, then using existing heuristics, a literature review and card sorts to identify potential new heuristics, cross-checking new heuristics against the problem database.

When comparing user evaluation and expert evaluation of websites, the latter using Nielsen's heuristics and a walkthrough, [24], used a grounded theory approach to group the problems found, plus analysis to see which problems were reported more by users and which were reported more by experts. Results were used to produce an enhanced set of heuristics, which were compared with Nielsen's. This paper ends by noting more evaluation of the heuristics is needed.

A literature review by [25] tabulated methods for developing usability heuristics and provided a methodology suggestion. The literature review covered nine methodologies, seven of which had validation steps and three of which went on to refine the heuristics post-validation. Four methodologies had validation and no refinement, which implies they expect not to find flaws during the evaluation. Further, [26] experimented with the steps in a proposed methodology with flow diagrams for various contexts. They suggest expert validation against four criteria: utility, clarity, ease of use and necessity of an extra checklist.

Heuristics for evaluating speech-based smart devices were suggested by [30] and validated by trial in use: usability experts applied the heuristics and fed back their view of the heuristics in an interview. Evaluators also filled out a usability report, but that seems to be of the devices, not of the heuristics.

An iterative approach is taken by [36], with a three-phase iterative process to refine the heuristics, looping back from validation design and from application back to design before moving on to a published set for use.

[2] 25 November 2022: Research Methods for Human-Computer Interaction, BCS Interaction Group One Day Course for PhD Students - presented by Helen Petrie and Gavin Sim.

Although understandability was mentioned as an evaluation criterion in several methods, the design and evaluation methods seemed to focus on the content of the heuristics rather than how they were communicated in terms of format. We believe that format contributes to understandability.

5.3 Analysis: Reviewing and Categorising Heuristics' Formats

To understand how people express and communicate heuristics, we examined grey literature from the software testing industry, and also heuristics from academic papers uncovered during the literature review. Across HCI, UX and software testing heuristics, we saw a variety of ways of communicating heuristics, and so we cast our net wider to cover heuristics outside these fields.

We sorted and categorised the heuristics looking for similarities, distinct groups, and distinguishing factors. We found we were able to group the heuristics in two ways: by their level of directiveness (telling, prompting, asking) and by their format (textual, conceptual, pictorial). These groups emerged as we examined different heuristics.

We realised that these characteristics were paired so that we could sort the heuristics into a nine-cell grid. We refer to the nine cells in the grid as the **'shapes of the heuristics'**, - a combination of their format and their directiveness. For example **'format: textual'** and **'directiveness: prompt'** becomes the **'Textual Prompting Shape'**.

The analysis that led to the Shape of Heuristics Model is described with examples in Section 5.4.

5.4 Building the 'Shape of Heuristics' Model

Shape of UX Heuristics Heuristics used for UX evaluations can be expressed as short textual statements to prompt thought, such as [20]'s *'User Control and Freedom'*. Sometimes longer more directive text is used, for example one heuristic by [28] for computer aided assessment in education is: *'9. Use clear language and grammar within questions and ensure the score is clearly displayed.'* with the explanation: *'Text should be grammatically correct and make sense. It should be obvious to students what the score is for a particular question and the scoring algorithm applied (eg, if negative marking is used). Question feedback should assist the learning process.'* This is both directive, and provides a full explanation of what to evaluate.

In this group we found heuristics to be **textual**, and **prompting** or **telling**.

Shape of Software Testing Heuristics Many forms of heuristic are in use in software testing. We found textual heuristics as shown in [1]: *'Overwhelm the product'* with a set of explanatory directives such as *'Look for sub-systems and functions that are vulnerable to being overloaded or "broken" in the presence of challenging data or constrained resources.'*. Bach's set of heuristics also includes

questions: *'Charisma. How appealing is the product?'* Conceptual heuristics include Elizabeth Hendriksen's *'Goldilocks testing' (do just enough...)* listed by [4], and [3] discusses the *'FEW HICCUPS'* mnemonic. There are also pictorial heuristics, such as [6]'s mind map of mnemonics.

In this group we found heuristics to be either **textual**, **pictorial** or **conceptual**, and we found heuristics that are either **telling**, **prompting** or **asking**.

Shape of Day to Day Heuristics Day-to-day heuristics include warnings, commands, information-providing pictures, phrases, and sensory perceptions. We examined this group to help us understand when different levels of directiveness and format are appropriate. We looked at road signs, proverbs and food.

Although there are similarities in road signs across the world, there are differences both across languages and across English-speaking territories [31, 32] (Examples from English-speaking countries are shown in Figure 3). An example of a heuristic warning found on road signs in many countries is 'Possible Ice On Road' which does not mean there is always ice on the road; it is a rule of thumb to suggest caution depending on the context (time of year, weather conditions) that indicates a slower speed may be sensible. In the USA this is textual, while in the UK a pictorial symbol is use. These are shown in Figure 3, based on [32]. This Wikipedia page shows a variety of road signs from several countries warning of potential hazards including ice and snow, slippery road surfaces, and loose road surfaces. All of these conditions may cause the car to skid. Some countries, such as Singapore and Uganda, don't have (perhaps don't need) an 'Ice On Road' heuristic, but they all have a 'Slippery Road Surface' sign.

Fig. 3. Example Snow Warning Road Signs(English-speaking Countries) after [32]

The context may dictate whether the heuristic is better formatted as pictorial or textual. Textual messages require learning the language and alphabet but then provide a slower but general purpose mechanism to understand a range of messages, whereas a pictorial heuristic must be learned individually, but is then perhaps faster to interpret. If we compare the UK and USA snow signs, might these reflect context? The UK has a very rigorous driving instruction and

test, which includes learning all the road sign symbols. The USA has a slower maximum driving speed, giving time to process the written word.

Other symbols are difficult to interpret without prior learning, for example in some countries the prohibition sign for 'no motorcycles' has a motorcycle with a line through it, while in other countries it is just a motorcycle. It is worth noting that in UK, some prohibitive signs have lines across and others do not, increasing both the cognitive load to learn and remember, and also reducing their comprehensibility to non-UK drivers.

We saw in Section 2 that heuristics are fallible rules of thumb. However, when used in critical or high-risk situations, heuristics that are built from intuition or experience need to be evaluated and challenged by expertise. For example, in an online article about rules for understanding which fungi are edible and which are poisonous, [2] critiques three heuristics for judging whether fungi are edible, providing the original textual directive and a counter-directive: *'There are some apparent rules for picking safe mushrooms but these are just fanciful if not downright dangerous; 'It's ok if you can peel the cap.' It is easy to peel a Death Cap. 'Mushrooms growing on wood are safe.' No not all of them are and some are deadly, like the Funeral Bell. 'If you see other animals eating them they are ok.' ... is not true, many animals can eat poisonous fungi with no ill effects.'*

Pictorial heuristics include the images on rest room signs. There are many variations on these pictorial prompts. A challenge to the usual image for the women's rest room is described by [17], who for an Axosoft 'Girls in Technology' challenge devised the 'It was never a dress' campaign, an online phenomenon. The pictorial prompt, shown in Figure 4, changes the outline of the woman in a dress to a woman in a superhero cape, providing a pictorial questioning heuristic that provokes the challenge: 'Can you (re)imagine women's roles in tech?'.[3]

In this group we found **textual, conceptual** and **pictorial** heuristics, communicated as **asking, prompting** or **telling**. We found the format and the level of directiveness may need to change, depending on context of use, changing the shape of heuristic that best communicates the heuristic.

Having described some of the example heuristics we discovered in UX, software testing and day to day life, and identified examples of pictorial, textual, conceptual, asking, prompting and telling metaphors, we can see that some of the shapes are easier to find examples than for others.

The Shape of Heuristics Model is shown in Table 1 including examples for each of the heuristics shapes, based on the discussion in this section. For the pilot study we describe in the next section, we chose to attempt to build heuristics in each shape.

[3] This version from https://www.flickr.com/photos/staciebee/39802849515/in/photostream/ where it is displayed under a Creative Commons NonCommercial-ShareAlike (CC BY-NC-SA 2.0) licence.

Table 1. Shape of Heuristics Model (Matrix)

	Textual	**Conceptual**	**Pictorial**
Asking	Textual Asking example: *Security: How well is the product protected against unauthorised use or intrusion?* [1]	Conceptual Asking example: *'If winter's here, can spring be far behind?'* [27]	Pictorial Asking example:  *It was never a dress: What (else) can women do (in tech)?* [16, 17]
Prompting	Prompting Telling example: *User Control and Freedom* [20]	Prompting Conceptual example:: *FEW HICCUPS (Familiar, Explainability, World, History, ... [etc.].* [3]	Prompting Pictorial example:  *Sign for slip and trip warning* (author photograph)
Telling	Telling Textual example: *Testing - Do one thing after another -* *1. Perform multiple activities connected end-to-end ... [etc.]* [1]	Telling Conceptual example: *Goldilocks Testing (do just enough...)* [4]	Telling Pictorial example:  *Road Sign - Stop Command* [32]

Fig. 4. Pictorial Heuristic: It Was Never a Dress [16]

6 Piloting in the Software Testing Context

As the Shape of Heuristics Model emerged, we wanted to understand if it would help us understand what was appropriate for the audience for our work. Who might need different formats or different levels of directiveness, and when? We piloted the Model with potential users of our heuristics to help them choose a preferred heuristic format, and provide reasons for their choices.

The pilot group of participants represented the people who would use our heuristics. As a result of this pilot evaluation, we made refinements in the Shape of Heuristics Model, discussed in Section 7.

6.1 Pilot Method

Our aim in this pilot study was to evaluate which of the shapes was most suitable for the specific heuristics we wanted to build. These were for UX designers to use when designing a testing tool, so a small, specific context.

Based on the content of Table 1, we designed shapes for a subset of three of our potential heuristics. Each of the three heuristics was a rule of thumb for one software quality attribute:

- Learnability: solo versus team learning;
- Maintainability: how tests and related artefacts can be updated, replaced, removed, etc.;
- Interoperability: how this tool works with other tools across a workflow.

We built three slide packs, each with 10 slides, the first showing the nine shapes in a matrix, and the other nine slides each showing the **same heuristic** in a different shape, one per slide. For example, one **pictorial prompting** heuristic for test maintainability included a picture of honey pouring to encourage thinking about viscosity (Figure 5), while a **conceptual asking** heuristic

was *'How easy is it to change the way data in this tool is modelled, visualised? Unmixing an omelette? Or sorting Lego bricks?'*

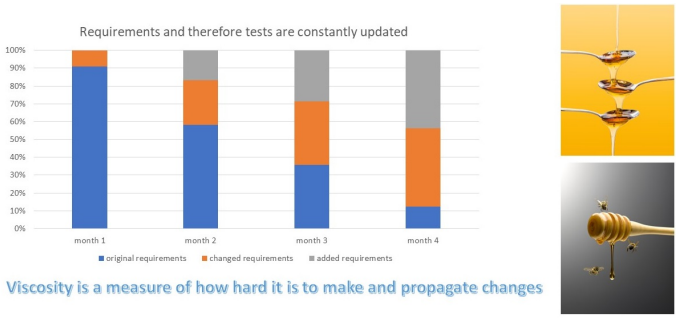

Fig. 5. Example Heuristic - Pictorial Prompt

We initially selected five participants, using convenience sampling. The participants included two UX experts with more than 10 years of experience in the industry field, two practitioners with 3-10 years of experience, and one newer UX professional with under 3 years of experience.

Our intention was to evaluate with this sample, make changes, and iterate the review with another group, until saturation was reached in terms of information, or agreement about which shape was most suitable for our specific work.

Participants received an information and consent form, and their responses were anonymised. Each participant was given a slide set, one or two participants to each attribute. They were given a few days to review the slides and think about them. Participants were individually interviewed, all via Zoom, and the interviews recorded to transcribe later for quotes and coding. Interviews were semi-structured. Initially, each was asked if they were familiar with the quality attribute in their slide set, if they were familiar with Nielsen's heuristics, and if they used Nielsen's heuristics. All the participants had heard of Nielsen's heuristics, and all but the least experienced had used them in projects. They all were familiar with the quality attribute used as an example in their own slide pack. The matrix of shapes was used to introduce the start of the evaluation, and they were asked if there were other formats or levels of directiveness they would add, or other dimensions. Each heuristic was then viewed and the participant asked if they found it understandable and useful, whether it was appropriate at a UX design stage.

6.2 Pilot Results

We transcribed the interviews, and also notes taken during the interviews, arriving at a score for each heuristic shape based on positive, negative and neutral comments; we coded +1 for each positive reaction, -1 for each negative reaction,

and 0 for a neutral reaction. We totalled the scores across participants for each shape. We tabulated those scores (see Table 2) to compare the reactions to the shapes.

Table 2. Heuristic Shapes Review: UX Design of Test Tools (Experts/Practitioners)

	Asking	**Prompting**	**Telling**	**Reviewer Comments:**
Pictorial	–3	–1	–2	accessibility issues
Textual	5	0	–2	preferred format
Conceptual	–2	–5	–3	inclusivity issues
Reviewer Comments:	Encourages divergent thinking: use at design stage	Better as second level of detail: use to provide supporting checklist	Encourages task completion: use at testing and UI evalution stage	

The scores in the table reflect the preferred option for the specific context of a design phase for a testing tool, and show a unanimous preference for a test question at the design stage for this type of tool. When asked the reason, the responses were that this would open wider thinking about the design.

The participants offered comments not just about the specific context, but also about what they would use in other contexts, in particular that they would favour a textual telling heuristic for UX testing, and would like textual prompts to make a check list underlying either the design stage or the test stage heuristics. No participant wanted to add other formats, levels of directiveness or dimensions to the table.

Participants noted both accessibility and inclusivity flaws in some of the shapes (Table 2), and this is one reason the pictorial and conceptual heuristics scored lower than textual. In the Discussion we will consider how the accessibility and inclusivity challenges may be overcome.

Heuristics presented pictorially seemed to be a less familiar format for this group. With some of the diagramatic representations, the participants preferred to build their own; they would rather have a prompt suggesting that a mind map or journey map would be useful to do, unless the diagram was providing specific information about a specific situation, such as a heat map of frequency of system changes for a specific system. The less experienced UXers were also less familiar with some of the diagrams and commented that they could not read them. This fits with [11] who found that experts find visual representations in programming environments easier to understand than novices do.

Following analysis, the results were shared with the participants for member checking, and additionally discussed with an accessibility expert and seven software testing experts, allowing a sense-check of the hypothesis and the pilot outcome.

6.3 Pilot Conclusions

There was unanimous consensus from these participants in this context. All participants independently choose a **textual question** heuristic to encourage divergent thinking during the design stage of a test tool. In comments, participants noted for the UX evaluation stage they preferred a **textual directive** heuristic. They wanted **textual prompt** heuristics in checklists to accompany the top level question or directive, as explanatory information.

At this point, unexpectedly, we had reached a consensus and had enough information to suggest that the Shape of Heuristics Model could be worth evaluating further. This provides a small sample saturation, which we considered sufficient to build our hypothesis, that the Shape of Heuristics affect their communicability, accessibility and inclusiveness in a particular context. This is not of course a sufficient study to make generalised comments about the best shape for heuristics in general, but it does show that reaching a consensus in a very specific context is possible. Our iterative design process for our own evidence-based heuristics continues.

In the next section we discuss the Shape of Heuristics Model and why it might help in heuristic design. We consider how to address some of the accessibility, inclusivity, and understandability issues raised in the pilot study.

7 Discussion

Our hypothesis that different shapes of heuristics may be appropriate in different contexts is supported by the pilot; We see from the pilot study that UX practitioners preferred different shape of heuristics in different circumstances, in this case the UX design versus UX testing stages of a project.

We have seen that pictorial, conceptual and textual heuristics are used both by IT professionals and in everyday contexts. Each of the formats has particular benefits and challenges.

In this section we discuss the formats and levels of directiveness, when they might be appropriate to use, and when the paired approach in the Shape of Heuristics Model may need augmenting.

7.1 Accessibility

The pilot study participants raised accessibility as a potential issue with non-textual heuristics. We considered accessibility, and also ease of learning, interpreting and using heuristics. In Table 3, we show how accessibility might be addressed in textual and pictorial heuristics, for permanent sight impairment.

Different approaches would be needed for temporary or situational sight impairment because in those contexts, the investment in learning Braille, raised diagrams, or the use of a screen reader is not applicable. Additionally, for other disabilities, whether sensory, intellectual, emotional, etc., it is possible to design heuristics that are accessible: using simple clear language, and providing multiple formats in different media.

Table 3. Addressing Accessibility Challenges in Text and Pictorial Heuristics

Format	Media	Disability	Accessibility Solution 1	Accessibility Solution 2
Text	Print	Sight impairment (permanent)	Screen Reader replaces sight by hearing	Braille replaces sight by touch
Pictorial	e-document	Sight impairment (permanent)	Alt Text for screen reader replaces sight by hearing	Raised diagram replaces sight by touch

Using more than one sense or ability to understand a heuristic is used in everyday life: foghorns and alarms provide a heuristic warning in sound to supplement lights and signs when vision is temporarily disabled in fog or smoke. Old pharmacy poison bottles use ridges on the sides of the bottle to indicate poison, a heuristic that allows a double-check of sight of the label and feel of the bottle to assist the pharmacist. A sense of smell is used to check food freshness, and is supplemented by sight (the clearness of a fish's eye on the fishmonger's slab) or in text, in a 'use by' date on packaging in Braille and print. We also use 'smell' as a conceptual heuristic when we talk about code smells or test smells.

7.2 Inclusion

The pilot study participants raised (lack of) inclusion particularly for the conceptual heuristics. These relied on cultural references, some of which were less understandable to participants in different age groups, or from different countries. Removing cultural, generational and contextual blockers to a heuristic's interpretation, depending on its context of use, is an important consideration. While emoji interpretation is a closed book to one researcher, rotary-dial telephones were an equally unsuitable metaphor to the youngest of the participants.

When looking at the collection of road signs on the Wikipedia pages [31, 32] it was clear that some of the iconography was symbolically meaningful but out of date, notably steam trains as the symbol for rail crossing in several countries. We see this also in user interfaces, where the symbol for saving a document was, up till very recently, a 3.5-inch floppy disk, which, although like the steam train, is still available, is not mainstream and may not have been used or seen by the majority of potential users. The context for a heuristic may outlive the symbolism used in the heuristic, leaving it open to misinterpretation, or reinterpretation. Whether the heuristic is conceptual, pictorial or textual, the way it is expressed

will provide information about the context within which it was developed, its age, and its authors' preconceptions.

Some heuristics have been designed to last for a very long time, and be understandable both in much later times, and by quite different cultures. Two examples are long-term signage for nuclear industry waste sites, and the Pioneer plaque. The iconography providing heuristics on the Pioneer plaque are critiqued [34] for lack of racial and gender inclusivity, and potentially over-reliance on specific cultural references. This is a lesson generally applicable to heuristic design. Similarly, the need to provide heuristics that last 10000 years and still convey 'Danger - Keep out!' for nuclear industry waste sites [33] is a perplexing, interesting, and important problem for heuristics design. These two examples may provide a lesson in communicating complex concepts in a heuristic and still making it understandable to all.

Initial indications from the pilot support our hypothesis, it appears that the shape of a heuristic affects how it is understood and used in context. The Shape of Heuristics Model could provide a starting point for designing and evaluating understandability, accessibility and inclusion of heuristics. Note that to achieve this more than one shape may need to be employed simultaneously. An example is the 'Slip and Trip Hazard' sign in Figure 6: a pictorial prompting heuristic plus two texts in different languages.

Fig. 6. Pictorial and Text Heuristic: Slip & Trip (author photo)

8 Conclusion and Next Steps

Heuristics, whether intuitive, experience-based, or evidence-based, may be communicated in different formats and with different levels of directiveness in different contexts. We introduce the first iteration of a model for understanding choices for the shape of a heuristic based on its format and level of directiveness. We present this in nine-cell matrix, which we call the Shape of Heuristics Model. The Model includes textual, conceptual and pictorial formats that can be combined with the telling, prompting and asking levels of directiveness, allowing someone designing or evaluating a heuristic to consider which shape is the most appropriate in their specific context, whether industry or academic.

Through a pilot study, we showed it is possible to review the potential shape of a heuristic with its audience independently of the heuristic's content. The pilot study also raised challenges of accessibility and inclusion. This means the same heuristic may need to be communicated in more than one shape, dependent on context and audience. We showed that context affects inclusion, considering the history, geography and societal meaning of a heuristic. How understandable a heuristic is, may be generational, depending on its age, and the age of those using it. Language and cultural meanings may be inadvertently embedded in a heuristic, by its designers and those using it, causing ambiguity, and exclusion.

8.1 Next Steps

Further work is required to evaluate the hypothesis and Model in different usage scenarios and contexts, to study the impact of the level of directiveness and format and its affect on how participants interpret and use heuristics. How to make heuristics both accessible and inclusive also requires further study.

We propose a future HCI study that trials the Shape of Heuristics Model to identify contexts where different levels of directiveness and different formats are appropriate. The pilot we ran could be a template, with an example heuristic being expressed in multiple ways to examine the effect of directiveness level on the actions people take, and the effect of both directiveness and format on accessibility, inclusion and understandabilty. Studies using heuristics in different domains could be an area of interest. These could then be added to the validation factors of utility, clarity, and ease of use suggested by [26]. In our own work, we will continue to evaluate the Shape of Heuristics Model and use it to inform the understandability, accessibility and inclusivity of the heuristics we develop for software testing and UX applications. Once fully evaluated the Shape of Heuristics Model could be used both in academic and industry settings for designing, evaluating and communicating heuristics.

Acknowledgments. The authors thanks the industry and academic participants in the pilot and discussion of the Shape of Heuristics Model.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bach, J.: Heuristic test strategy model (2024), https://www.satisfice.com/download/heuristic-test-strategy-model
2. Biggane, E.: How to tell the difference between poisonous and edible mushrooms (2014), https://www.wildfooduk.com/articles/how-to-tell-the-difference-between-poisonous-and-edible-mushrooms/
3. Bolton, M.: FEW HICCUPS (2012), https://developsense.com/blog/2012/07/few-hiccupps
4. Bradshaw, R., Deery, S.: Software testing heuristics: Mind the gap! (2019), https://www.ministryoftesting.com/articles/ce0dc29c?s_id=1474072
5. Charles, F.: Workshops: Mastering the power of heuristics in testing (2024), https://www.quality-intelligence.com/workshops.htm
6. Dewar, D.: Testing mnemonics desktop (2015), https://findingdeefex.com/
7. Evans, I., Porter, C., Micallef, M.: Scared, frustrated and quietly proud: Testers' lived experience of tools and automation. In: European Conference on Cognitive Ergonomics 2021. pp. 1–7. ACM, New York, NY, USA (2021)
8. Evans, I., Porter, C., Micallef, M.: Breaking testing stereotypes: Who is testing and why it matters. In: To be published in Proceedings of BCS HCI Conference 2024. BCS HCI (2024)
9. Evans, I., Porter, C., Micallef, M., Harty, J.: Stuck in limbo with magical solutions: The testers' lived experiences of tools and automation. In: Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. pp. 195–202. SciTePress-Science and Technology Publications, SciTePress, Portugal (2020)
10. Evans, I., Porter, C., Micallef, M., Harty, J.: Test tools: an illusion of usability? In: 2020 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW). pp. 392–397. IEEE, IEEE Computer Society, Washington, DC, USA (2020)
11. Green, T.R.G., Petre, M.: Usability analysis of visual programming environments: a 'cognitive dimensions' framework. Journal of Visual Languages & Computing **7**(2), 131–174 (1996)
12. Hennink, M., Kaiser, B.N.: Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science I& Medicine **292**, 114523 (2022). https://doi.org/https://doi.org/10.1016/j.socscimed.2021.114523, https://www.sciencedirect.com/science/article/pii/S0277953621008558
13. Hjeij, M., Vilks, A.: A brief history of heuristics: how did research on heuristics evolve? Humanities and Social Sciences Communications **10**(1), 1–15 (2023)
14. Inostroza, R., Rusu, C., Roncagliolo, S., Jimenez, C., Rusu, V.: Usability heuristics for touchscreen-based mobile devices. In: 2012 Ninth International Conference on Information Technology-New Generations. pp. 662–667. IEEE (2012)
15. ISTQB: ISTQB CTFL Syllabus v4.0 (2023), https://www.istqb.org/certifications/certification-list
16. ItWasNeverADress.org: It was never a dress (2015), https://itwasneveradress.org/about/
17. Katan, T.: Creative Trespassing. Penguin Random House (2019)
18. Kumar, B.A., Goundar, M.S., Chand, S.S.: A framework for heuristic evaluation of mobile learning applications. Education and Information Technologies **25**, 3189–3204 (2020)

19. Nielsen, J., Molich, R.: Teaching user interface design based on usability engineering. SIGCHI Bull. **21**(1), 45–48 (Aug 1989). https://doi.org/10.1145/67880.67885, https://doi.org/10.1145/67880.67885
20. Nielsen, J.: Ten usability heuristics (2024), https://www.nngroup.com/articles/ten-usability-heuristics/
21. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: International Conference on Human Factors in Computing Systems (1990), https://api.semanticscholar.org/CorpusID:17451097
22. OED: Oxford English Dictionary (2024), https://www.oed.com/
23. OED, O.E.D.: Understanding frequency of use per million words (2023), https://www.oed.com/information/understanding-entries/frequency/
24. Petrie, H., Power, C.: What do users really care about? A comparison of usability problems found by users and experts on highly interactive websites. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2107–2116 (2012)
25. Quiñones, D., Rusu, C.: How to develop usability heuristics: A systematic literature review. Computer Standards & Interfaces **53**, 89–122 (2017)
26. Quiñones, D., Rusu, C., Rusu, V.: A methodology to develop usability/user experience heuristics. Computer Standards & Interfaces **59**, 109–129 (2018)
27. Shelley, P.B.: Ode to the west wind (1820), https://www.poetryfoundation.org/poems/45134/ode-to-the-west-wind
28. Sim, G., Read, J.C.: Using computer-assisted assessment heuristics for usability evaluations. British Journal of Educational Technology **47**(4), 694–709 (2016)
29. Sim, G., Read, J.C., Cockton, G.: Evidence based design of heuristics for computer assisted assessment. In: Human-Computer Interaction–INTERACT 2009: 12th IFIP TC 13 International Conference, Uppsala, Sweden, August 24-28, 2009, Proceedings, Part I 12. pp. 204–216. Springer (2009)
30. Wei, Z., Landay, J.A.: Evaluating speech-based smart devices using new usability heuristics. IEEE Pervasive Computing **17**(2), 84–96 (2018)
31. Wikipedia contributors: Comparison of european road signs — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Comparison_of_European_road_signs (2024), [Online; accessed 10-April-2024]
32. Wikipedia contributors: Comparison of traffic signs in english-speaking territories — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Comparison_of_traffic_signs_in_English-speaking_territories (2024), [Online; accessed 10-April-2024]
33. Wikipedia contributors: Long-term nuclear waste warning messages — Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Long-term_nuclear_waste_warning_messages (2024), [Online; accessed 10-April-2024]
34. Wikipedia contributors: Pioneer plaque — Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Pioneer_plaque (2024), [Online; accessed 10-April-2024]
35. Wong, E.: Heuristic evaluation: How to conduct a heuristic evaluation (2022), https://www.interaction-design.org/literature/article/heuristic-evaluation-how-to-conduct-a-heuristic-evaluation
36. Yeratziotis, A., Pottas, D., Van Greunen, D.: A usable security heuristic evaluation for the online health social networking paradigm. International Journal of Human-Computer Interaction **28**(10), 678–694 (2012)