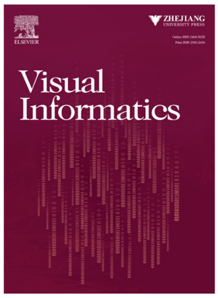
Journal Pre-proof



Incidental visualizations: How complexity factors influence task performance

João Moreira, Daniel Mendes, Daniel Gonçalves

PII: S2468-502X(24)00065-2
DOI: https://doi.org/10.1016/j.visinf.2024.10.005
Reference: VISINF 221

To appear in: *Visual Informatics*

Received date : 15 November 2023
Revised date : 9 September 2024
Accepted date : 27 October 2024

Please cite this article as: J. Moreira, D. Mendes and D. Gonçalves, Incidental visualizations: How complexity factors influence task performance. *Visual Informatics* (2024), doi: https://doi.org/10.1016/j.visinf.2024.10.005.

Incidental Visualizations:
How Complexity Factors Influence Task Performance

João Moreira[a], Daniel Mendes[b], Daniel Gonçalves[a]

[a] *INESC-ID, Rua Alves Redol, 9, 1000-029 Lisboa, Portugal*
[b] *INESC TEC, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*

Abstract

Incidental visualizations convey information to a person during an ongoing primary task, without the person consciously searching for or requesting that information. They differ from glanceable visualizations by not being people's main focus, and from ambient visualizations by not being embedded in the environment. Instead, they are presented as secondary information that can be observed without a person losing focus on their current task. However, despite extensive research on glanceable and ambient visualizations, the topic of incidental visualizations is yet a novel topic in current research. To bridge this gap, we conducted an empirical user study presenting participants with an incidental visualization while performing a primary task. We aimed to understand how complexity contributory factors—task complexity, output complexity, and pressure—affected primary task performance and incidental visualization accuracy. Our findings showed that incidental visualizations effectively conveyed information without disrupting the primary task, but working memory limitations should be considered. Additionally, output and pressure significantly influenced the primary task's results. In conclusion, our study provides insights into the perception accuracy and performance impact of incidental visualizations in relation to complexity factors.

Keywords: Incidental Visualization, Information Visualization, User Study, Task Complexity.

Email address: joaopedro.moreira@inesc-id.pt (João Moreira)

Incidental Visualizations: How Complexity Factors Influence Task Performance

João Moreira[a], Daniel Mendes[b], Daniel Gonçalves[a]

[a]*INESC-ID, Rua Alves Redol, 9, 1000-029 Lisboa, Portugal*
[b]*INESC TEC, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*

Abstract

Incidental visualizations convey information to a person during an ongoing primary task, without the person consciously searching for or requesting that information. They differ from glanceable visualizations by not being people's main focus, and from ambient visualizations by not being embedded in the environment. Instead, they are presented as secondary information that can be observed without a person losing focus on their current task. However, despite extensive research on glanceable and ambient visualizations, the topic of incidental visualizations is yet a novel topic in current research. To bridge this gap, we conducted an empirical user study presenting participants with an incidental visualization while performing a primary task. We aimed to understand how complexity contributory factors—task complexity, output complexity, and pressure—affected primary task performance and incidental visualization accuracy. Our findings showed that incidental visualizations effectively conveyed information without disrupting the primary task, but working memory limitations should be considered. Additionally, output and pressure significantly influenced the primary task's results. In conclusion, our study provides insights into the perception accuracy and performance impact of incidental visualizations in relation to complexity factors.

Keywords: Incidental Visualization, Information Visualization, User Study, Task Complexity.

Email address: `joaopedro.moreira@inesc-id.pt` (João Moreira)

1. Introduction

In the course of people's daily lives, instances arise when multiple sources of information are present, yet it becomes challenging for individuals to effectively monitor all ongoing events simultaneously. Consider a scenario with four appliances functioning, each equipped with a timer for a specific task. However, if an individual were to gain access to these timers – through a bar chart displaying four bars within their field of vision while they engage in a primary task – this access could enable that person to stay informed about various details in real-time and with a quick glance.

Although there are currently ways to get data in real-time, presenting it effectively and quickly, while not disrupting too much, is yet an unsolved issue. Even with ubiquitous devices such as smartphones or smartwatches, people would still need to consciously remember they can access the information. However, by the time they looked for the information, time was already lost, and the task could already be compromised. If people have to choose and search for the information for themselves, it may induce unwanted additional cognitive load, even with useful information. For example, augmented reality interfaces have been explored (Lu and Bowman, 2021) on how they might apply to everyday uses to convey information, and two of the problems they reported are distraction and occlusion. Even though they did not explore information visualization, we argue that distraction and occlusion might also happen if we try to present several graphs at the same time.

For instance, let's imagine people living in a fully intelligent house where every resource, such as energy, water, gas, and food, is measured in real-time, and assuming the existence of technology that can determine when to present visualizations. The primary tasks here refer to any task a person is engaged in that doesn't require them to explicitly focus on a visualization. By displaying graphs incidentally, people could gain a better understanding of how these resources are utilized. For example, a line chart can show energy consumption over time, or a bar chart can display ingredient stock levels, all without the need for users to actively search for and explore this data. The question at hand is: How can visualizations effectively convey information to users during their primary tasks without causing distraction or excessive obscuring of the real-world environment? Additionally, how can users receive this information quickly and effortlessly, all without requiring them to consciously search for or request it? These are essential considerations in designing user-friendly visualizations that seamlessly integrate with users' primary activities and provide the necessary data without disrupting their workflow.

Calm Technology has been defined (Weiser and Brown, 1996) as any device that interacts with people via auditory or visual channels while allowing information to be conveyed from the periphery to the center of human attention and back. In other words, calm technology informs but does not demand our focus or attention. In information visualization, calm technology has been explored as ambient information systems such as ambient (Skog, 2004), glanceable (Blascheck and Isenberg, 2021), and incidental (Moreira et al., 2020, 2023a,b) visualizations. These last two types share a core concept: they are seen at a glance for very short exposure times. However, they differ in their use cases. In some glanceable visualizations people actively search for the depicted information on an easily accessible device, (Gouveia et al., 2016) such as a smartwatch. In contrast, incidental visualizations depict information within one's field of view without the need for conscious searching, thereby ensuring that the focus of attention remains on the primary task. Incidental visualizations are precisely what people would want for the mentioned resource management example. Incidental visualization revolves around graphs that have the potential to manifest at any location and specific instances without requiring explicit user initiation. In that house, people could be presented with incidental visualizations through some technology like augmented reality without people searching for the graphs, allowing them to be **aware in real-time** of contextualized information displayed in their field of view, without stopping their current tasks.

The scenario we envision, assumes the existence of theoretical background on incidental visualizations, for which design guidelines are not yet available. Additionally, there is a lack of studies exploring the hardware required for implementing such visualizations, although augmented reality is speculated to be a suitable fit. In two preliminary studies on incidental visualizations (Moreira et al., 2020, 2023b), the perception of pre-attentive tasks was investigated to understand the effectiveness of different marks and channels. The authors concluded that information displayed (up to four marks) for very short durations (up to one second) at specific moments can be effectively perceived. Subsequently, the influence of incidental visualizations, which were presented alongside a primary task (maze game), on users' performance was explored (Moreira et al., 2023a). The study concludes that these visualizations do not disrupt the primary task and, in fact, they may enhance users' ability to respond to related questions. Building upon this research, our study aims to validate the effectiveness of incidental visualizations in real-life scenarios and examine how task properties influence task performance, with the objective, in future work, of exploring the optimal placement and presentation of these visualizations in the real world.

In this paper, we present the results of an empirical online between-subjects user study with 120 participants, where we studied the extent to which people can effectively perceive information from an incidental visualization while performing a primary task. We tested three different task complexity factors and analyzed the influence each one had on participants' performance. The primary task had participants interact with a choropleth map to select regions that satisfied specific criteria. At a certain point, an incidental visualization (a horizontal bar chart), would appear to add another needed piece of information for participants to successfully finish the task. According to our results, the overall primary task performance was normally high, and the horizontal bar chart was mostly perceived accurately, indicating that incidental visualizations should be further explored for future design concepts.

Our major contributions with this paper are a user study to evaluate a primary task with an incidental visualization; how the task performance and visualization accuracy are affected by three complexity factors; key insights for future research with incidental visualizations.

2. Related Work

Glanceable perception is the topic that matches the most with the incidental perception definition, except for some use case scenarios. Therefore, we will start by mentioning some of the research that has been conducted lately. Then, to motivate future implementation for incidental visualizations, we will go through augmented reality (AR), in particular studies that mention glanceable AR. Next, more closely related to our work, we will show some task complexity frameworks that have been defined lately. Finally, to justify some design choices made, we will state the current knowledge regarding graphical perception theory.

2.1. Glanceable Perception

Incidental visualizations must be designed to be seen at-a-glance, and "Glanceability refers to the perception and interpretation of information after the user is paying attention to the interface" (Matthews, 2006). There are several challenges related to glanceable perception, but all of those challenges end up in trying to understand how one person can be aware of information received at-a-glance. For several years, these challenges have been addressed. One good example is the Info-lotus (Zhang et al., 2005), where authors presented a peripheral notification system where people could see several types of personal information at-a-glance, all in one place. Around the same time, visualizations were developed (van Dantzich et al., 2002) designed to be glanceable and reduce the mental effort people had to spend when deciding where to look at. Nowadays, mostly due to their size, smartwatches and fitness bands are usually the choice to convey such information (Gouveia et al., 2016). One recurrent theme is health, due to the capability of these devices of monitoring body metrics. For example, it was demonstrated (Ankrah et al., 2022) that children were fully aware of their health using Apple Watches; A survey was made on sleep data (Islam et al., 2022a) on fitness trackers, and concluded that the visualizations shown on the devices had the potential to be glanceable and were effective at communicating sleep data to wearers ;And a study was conducted (Grioui and Blascheck, 2021) to test three visualizations displayed on a smartwatch, and concluded that participants preferred information displayed in a bar chart. If not with wearables, we have other examples. Health related, glyphs were tested (DEURZEN et al., 2022) by being displayed on smartphones to make workers aware of their current posture. If not regarding health, pervasive visualizations were tested (Wilkinson et al., 2020) to enhance people's awareness of the data-sharing from their smartphone.

Regarding particular details for information awareness, some challenges can be context, body movements, or screen size. The information that is perceived could depend on what the user is currently doing (Islam et al., 2022b). Then, one of the consequences of using devices placed on the body is constant movement. Therefore, information will be read in motion. For example, several visualizations were reviewed (Islam et al., 2022c) in motion and the authors concluded that viewers can only afford to glance at watch faces for very short exposure times, which is particularly relevant for running athletes where movement is much more intense. User acceptance and utility of real-time visualizations on smartwatches were explored during sport activities (Schiewe et al., 2020) and the authors concluded there is not enough evidence that athletes will perform better. However, they still preferred to use these devices for self-monitoring and motivation.

The issue with reading in motion using smartwatches or fitness bands is also related to the small screen sizes. The challenge here is how to create effective visualizations that fit these screens without compromising comprehensibility. A study was conducted to assess how quickly people could read the information in small-scaled visualizations, (Blascheck et al., 2019) and the authors concluded that individual bars and donut sectors could be assessed in a few hundred milliseconds. They later replicated their study on a laptop (Blascheck and Isenberg, 2021) and the overall trends were still the same, proving that perception on larger screens may still apply to small ones. Around the same time, it was investigated how people would perform at seeing information in smartwatches (Neshati et al., 2019), but using high-density continuous time-series data. The authors concluded that graph segments are best interpreted when compressed along the x-axis.

Understanding how people can effectively perceive and comprehend information presented at a glance is a significant challenge. When it comes to incidental visualizations, the primary difference between the current state of the art and our knowledge gaps lies in our lack of advance knowledge regarding the location and timing of the information. Existing solutions primarily concentrate on providing information through easily accessible devices, such as smartwatches, which can be quickly explored in brief intervals.

Glanceable augmented reality offers a potential solution for presenting information at-a-glance, but it also introduces new challenges, such as visual occlusion (Daskalogrigorakis et al., 2022). However, the exploration of information visualization in the context of glanceable augmented reality remains relatively limited. Designers must carefully determine the amount of information conveyed to avoid overwhelming users (Davari et al., 2020). Secondary stimuli, such as notifications, can disrupt attention (Faulhaber et al., 2022). It was found that a circular progress bar was effective in providing progress information without affecting eye contact (Janaka et al., 2022).

Research on everyday tasks in glanceable augmented reality has been conducted (Lu, 2021; Zhang et al., 2022). The importance of real-world visibility and minimizing visual clutter was identified (Davari et al., 2020), and the potential interference caused by peripheral content was highlighted (Lu et al., 2021). However, overall user perception of glanceable augmented reality remains positive (Lu and Bowman, 2021). Context-aware glanceable augmented reality aims to provide relevant information (Davari et al., 2022). Users can perceive information effectively with minimal interference in primary tasks (Davari et al., 2022).

2.2. Task Complexity

Even if an incidental visualization is perfectly designed, the ongoing primary task must be considered carefully. Its complexity might affect how accurately the visualization is perceived. Task complexity is an area of research that has been studied for several years. However, even now, it is hard for authors to find a global consensus. Robert Wood (Wood, 1986; Wood et al., 1987) and Donald Campbell (Campbell, 1988) are known for their research in task complexity, and inspired research until now. Wood defined complexity as a combination of three types of complexity: component, coordinative, and dynamic. Campbell divided task complexity into four categories: decision, judgment, problem, and fuzzy. Although these were the foundations at the time, several frameworks have emerged since then. For example, complexity was divided into Coordinative Complexity and Component Complexity (Lazzara et al., 2010). The former assumes several team members, whereas the latter does not.

A significant contribution to the field was made (Liu and Li, 2012), where authors conducted a comprehensive survey of literature up until 2012 to summarize the research conducted on task complexity. Initially, they categorized papers based on different viewpoints and definitions, leading to the proposal of a task model. In essence, a task comprises six components: goal, input, output, process, presentation, and time, each consisting of various contributing factors. Subsequently, the authors restructured the task model, expanding it to include ten dimensions: size, variety, ambiguity, relationship, variability, unreliability, novelty, incongruity, action complexity, and temporal demand. Each dimension also encompasses multiple contributing factors, with each factor corresponding to one of the six components.

More recently, a task complexity framework was proposed (Efatmaneshnik and Handley, 2021), based on two components, objective view and subjective view, components that were explored in 1997 (Maynard and Hakel, 1997). Efatmaneshnik et al. proposed their framework to enhance human systems' integration processes. The objective view is based on a task model where tasks are defined by input, processing skills, and constraints. The subjective view is based on a task model where complexity depends on several personal metrics such as motivations, experience, etc. Then, completely shifting from traditional frameworks, it was argued that task complexity is, in fact, a social practice, and it is something that dynamically changes (Danner-Schröder and Ostermann, 2022). They concluded that tasks do not possess a complexity, but they become complex in the enactment due to four mechanisms: forming new paths, keeping paths open, enacting interdependent paths in parallel, and dissolving old paths.

Finally, some works have explored task complexity interaction with task performance. For example, it was explored how task complexity affected information seeking (Byström and Järvelin, 1995). One of the conclusions they drew was that, if a person needs information to complete a certain task, if its complexity is high, it will decrease the chances of the person finding the information needed. Then, a framework was proposed that connects several metrics we presented already (Liu and Li, 2011), but also highlighted how they could impact task performance. In the context of incidental visualizations, a study was conducted to assess their impact on primary tasks (Moreira et al., 2023a), and the conclusion was that their presence did not significantly affect performance in most cases. However, the study did not investigate if this holds true for tasks with distinct and specific characteristics.

The shared objective of research on task complexity is to enhance our understanding of how task complexity affects information seeking and overall task completion. As we have observed, a lack of consensus indicates the difficulty in reaching a definitive solution. Nevertheless, in the end of this section, we will delve into a detailed discussion, focusing on the empirical framework proposed by Liu and Li, which we deem suitable for our studies.

2.3. Graphical Perception

To create any visualization, designers make use of research regarding graphical perception. Cleveland and McGill (Cleveland and McGill, 1987, 1986, 1985, 1984) are very well-known for their original work on graphical perception, which is the visual decoding of information encoded on graphs. These results have been used ever since, and even recently it was explored if these results apply to both adults and children (Panavas et al., 2022), and the authors concluded that the theory also applied to these ages. In information visualization, graphical perception theory is used to choose the best marks and channels to convey information effectively when performing specific tasks.

However, graphical perception rankings should not be used blindly (Bertini et al., 2020). It was concluded that rankings do not capture how people use and learn visualizations. Therefore, rankings should be thought regarding the tasks to be performed. For example, a survey was conducted to understand which visual idioms are more effective for low-level tasks (Quadri and Rosen, 2022; Brehmer and Munzner, 2013). Furthermore, a literature review was made to compare visualization designs in terms of visual perception and human performance under different analysis tasks (Zeng and Battle, 2021).

Then, besides the task's nature, individual differences might change how accurately people perceive visualizations. The Cleveland and McGill studies were replicated (Davis et al., 2022), and the authors examined the performance variance between participants themselves. They concluded that a visual idiom that is ranked best for the average participant may not be ranked best for a substantial portion of people, which may comprise current design guidelines. Then, still on individual differences, now regarding cognitive ones, a survey was made on how individual differences have been measured to use that data at accessing visualization performance (Liu et al., 2020).

Regarding incidental graphical perception, two user studies have been conducted (Moreira et al., 2020, 2023b). In the first, the authors tested six combinations of marks and channels: horizontal and vertical positions; length and slope of lines; the size of areas; and color luminance intensity, and concluded that horizontal position identification is the most accurate and fastest task to do. In the second, they tested three combinations of marks and channels: length of lines, horizontal and vertical positions of dots, and angle of lines. However, in this study, each combination was displayed with either one, two, three, or four marks because the authors wanted to measure accuracy at the subitizing range, which is the term given (Kaufman et al., 1949) for the enumeration of small sets precisely and effortlessly.

Up to four objects, people can instantly recognize and accurately enumerate the number of objects without the need for counting, and this range was confirmed also with incidental visualizations. However, even now, whether attention is necessary for subitizing remains debatable. A survey was conducted (Chen et al., 2022) to determine whether manipulations to attention demonstratively affect subitizing, and proposed a framework for the involvement of attention in visual enumeration.

In summary, there are three enumeration mechanisms: estimation, subitizing, and counting. For subitizing, medium attention demands still allow for this mechanism to work effectively, such as divided, selective, spatial, or temporal, attention.

Following this work, the same authors investigated the effects of incidental visualizations on users' performance within a maze game context. Through an empirical study, the research reveals that these visualizations neither hinder the primary task nor adversely affect participants' question-answering abilities. Notably, this positive impact remains consistent, and working memory appears to have no significant influence in this regard.

Finally, selecting appropriate visual marks and channels can have a significant impact. Therefore, it is crucial to comprehend the accuracy rankings based on the specific tasks we aim to accomplish. However, it is important not to blindly rely on rankings alone. Various factors such as age disparities between adults and children, the nature of the tasks influencing visual perception, and individual cognitive differences can all introduce variations and potential interferences.

2.4. Discussion

Incidental visualizations and glanceable visualizations share a common objective: enabling individuals to perceive information at a glance. However, there are fundamental differences between the two. Glanceable visualizations are accessed either on purpose or based on peripheral stimuli (Blascheck et al., 2021), and people consciously will look at it. In contrast, **incidental visualizations are not actively sought out or perceived; instead, they spontaneously appear in a person's field of view during an ongoing primary task without needing conscious focus.** They are, in essence, trully embedded as calm technology (Weiser and Brown, 1996). While incidental visualizations may not encounter challenges related to body movements or screen size, they do face the contextual challenge of providing information based on the user's current task or activity. Body movements are no problem because we argue that an incidental visualization should appear in each person's field of view, not at a specific location such as a body part. Then, due to the study where the authors replicated the smartwatch results (Blascheck and Isenberg, 2021), the screen size should also not be an issue. Therefore, if we are to picture a possible hardware implementation, augmented reality probably fits what we want. Although occlusion, distraction, or even annoyance may become the next big challenges, we hope they will have lesser importance due to two major design choices in incidental visualizations. First, they should appear for short exposure times, which may minimize the occlusion time of the real world. Second, they should be designed to minimize cognitive load to avoid distraction from primary tasks, which then may lead to less annoyance felt. While our current study does not specifically focus on glanceable augmented reality, we recognize its potential as a promising avenue for implementing incidental visualizations in the future.

When it comes to task complexity, as we have observed, there is no clear consensus regarding the preferred framework to be utilized. For our specific study, frameworks that involve multiple individuals or assume complexity as a social construct will not be employed. We intend to initially test incidental visualizations in isolation before considering social scenarios. Furthermore, we aim to avoid subjective perspectives on task complexity as our primary focus is on the inherent nature of tasks. Thus, the framework proposed by Liu and Li (Liu and Li, 2012) proves to be more suitable for our purposes. Their framework offers a level of granularity that enables studies to concentrate on specific factors, three of which we will utilize as independent variables in our research. Specifically, we evaluated the complexity dimensions of size and temporal demand. The former relates to the number of task components, where we examined the complexity contributory factors of output quantity and input quantity. The latter pertains to the pressure-induced task requirements, for which we tested the complexity contributory factor of pressure.

Incidental visualizations, like any other type of visualization, will be created with visual idioms designed with graphical perception theory. However, if people are performing primary tasks, incidental visualizations should demand as little attention as possible without compromising visual accuracy. Therefore, as Moreira et al. (Moreira et al., 2023b,a) used in their studies, we will conduct ours within the subitizing range. Besides, according to Chen et al. (Chen et al., 2022) framework, subitizing works in divided attention situations, which is what happens with incidental visualizations. Furthermore, according to Quadri et al. (Quadri and Rosen, 2022), for retrieving values tasks, which are the tasks we used in our study, bar charts are preferred. Therefore, we used bar charts to convey information, each with four bars due to the subitizing range. To address individual differences in visual perception (Liu et al., 2020), we utilized the OSPAN (Operation Span) test (Turner and Engle, 1989) to assess participants' working memory capacity. The test involves solving mathematical operations while memorizing a sequence of letters and recalling them in the correct order. The OSPAN measure provides insights into individuals' ability to maintain and manipulate information in working memory. In our study, we replaced letters with icons for the test.

3. User Study

Since incidental graphical perception has already been validated, and we are aware that incidental visualizations can coexist with a primary task, we shifted our focus to the primary task itself. Our objective was to examine the extent to having an incidental visualization influences people's performance, response time, and confidence in the primary task, while systematically varying three complexity contributory factors of the primary task: input complexity, output complexity, and pressure. Additionally, we aimed to determine the accuracy of perceiving the information conveyed by the visualization regardless of the primary task performance.

According to Liu and Li framework (Liu and Li, 2012), input and output are related to the size task complexity dimension, and pressure belongs to the temporal demand dimension:

- **Input**: Number of criteria to be satisfied;
- **Output**: Number of values to retrieve;
- **Pressure**: Time available to finish the task.

As the primary task, we asked participants to explore a choropleth map and select regions according to a set of creteria: selecting a region with a high population density, selecting a region with a high hotel count, selecting a region near the sea, and selecting a region with warm weather. This last criterion was the only where the information needed to satisfy it was conveyed by the incidental visualization. All the other could be satisfied by exploring the choropleth map.

We included the incidental visualization in the task to evaluate how participants would perceive it. While participants knew that the visualization would appear at some point, they were unaware of the exact timing, ensuring that their perception of it would be incidental. As a result, they began to meet the other criteria, thus avoiding the need to wait for the incidental visualization. Further details can be found in Section 3.1.

There are two insights to discover: how input, output, and pressure influence performance, and how they influence the perception of the incidental visualization. Although the task performance depends on the perception accuracy of the visualization, we still wanted to isolate the visualization score. For example, if the task performance was low, we would still want to know if people correctly perceived the visualization. Therefore, our research questions are the following:

- **RQ1**: How much input, output, and pressure complexity contributory factors impact the accuracy, response time, and confidence of the primary task?
- **RQ2**: How much input, output, and pressure complexity contributory factors impact the perception accuracy of the incidental visualization?

Our hypothesis is that modifying task complexity factors will result in noticeable impacts on task performance and the perception of incidental visualizations. In particular, we believe that as we increase the input, output, and pressure, the accuracy, response time, and confidence in performing the primary task will likely decrease. Additionally, we anticipate a decrease in the perception accuracy of incidental visualizations. Nevertheless, we contend that it is imperative to validate these findings in the context of incidental visualizations, given the scarcity of relevant literature. This underscores the need to proceed cautiously, taking incremental steps before delving into a more in-depth exploration of these aspects.

3.1. Study Design

We conducted a crowdsourced online user study. Using a Crowdsourcing platform for information visualization solves one major issue recurrent in laboratory studies: small participant sample sizes. However, several details must be carefully considered (Borgo et al., 2017). To simplify this process, a checklist was proposed (Borgo et al., 2018) for reporting crowdsourcing experiments, that we used for our experiment. This checklist summarizes the six key aspects identified by the authors (Borgo et al., 2018) as critical for the successful explanation of crowdsourcing experiments. By considering all of these aspects, the experiment is likely to be more sound. Our study followed a mixed design with one between-subjects variable and two within, all of them with three levels. It was conducted using Prolific and only required a desktop computer.

Our between-subjects variable was the **input (two, three, or four criteria)**, and our within-subjects variables were the **output (one, two, or three selected regions)** and **pressure (15s, 20s, or 25s to finish the task)**. The pressure values were obtained from a pilot study involving 10 participants. On average, participants took 20 seconds to select three regions and satisfy four criteria. Therefore, we tested slight variations from the average as different levels of pressure time. Therefore, we divided our participants into three groups, and in each one, each participant underwent nine trials. We measured the performance of the primary task (variable named accuracy), which was the percentage of items correctly retrieved, the response time for each trial (variable named response time); the confidence of each trial submission (variable named confidence) measured using a 5-point Likert scale, 1 - Very Low, and 5 - Very High; and the participants' accuracy at perceiving the incidental visualization (variable named vis effectiveness), **which was the percentage of regions selected that only satisfied the criterion related to the visualization**.

To conduct our user study, we first needed to define our **primary task**. We examined a scenario similar to real life where an incidental visualization could enhance the overall performance of the primary task. **Participants used a computer to search for regions on a choropleth map that encoded three pieces of information** via three visual channels: hue, text, and position. However, **a fourth piece was needed**, but there was no optimal way to encode that information on the choropleth map. Therefore, **an incidental visualization was used to convey that additional information** for one second to avoid distracting participants from their main focus. Participants were aware that the visualization would appear but were uncertain about the timing. We chose one second due to the results of a previous study on incidental visualizations (Moreira et al., 2023b) where authors determined that exposure times between 300ms and 1000ms were effective. Also, although participants were unaware of when the visualization would appear, it was not triggered randomly. We ensured that it always appeared at the same moment for participants within a specific testing group. For instance, if a participant had to select three regions, the visualization would appear after they hovered over two regions.

3.2. Visual Design

Maps are used for this type of task (Quadri and Rosen, 2022), retrieving values that satisfy a given set of specific criteria. These criteria corresponded to our **input** factor, and the retrieved values to our **output** factor. Finally, the time to perform this primary task corresponded to our **pressure** factor.

To create the regions of the map, we used the NUTS (EurostatNUTS) classification (Nomenclature of territorial units for statistics), which is a hierarchical system for dividing up the economic territory of the EU and the UK. Our map contained the NUTS regions of Portugal, from which only the 'Continente' NUTS-I region was selected. In total, there are five NUTS-II and 23 NUTS-III regions in 'Continente', and the NUTS-III regions corresponded to the items that people needed to retrieve for completing the primary task. Depending on the trial, each participant either selected one, two, or three NUTS-III regions.

In our visualization, the hue within each region represented the population density, with higher saturation of green indicating higher density. To ensure accessibility, we selected a color scheme with four distinct classes using ColorBrewer (ColorBrewer), taking care to ensure colorblind friendliness. We ensured that only one of these four color classes appeared in each region to prevent interpolated colors, which could introduce additional variability in our results. Additionally, we incorporated text tooltips to provide specific information about each region, which appeared when the mouse cursor was placed over a region. Moreover, the position of the sea was indicated on the map, presented on the left side of Portugal using a blue hue. To maintain clarity and minimize unnecessary elements, we included only essential graphic components required for participants to complete the tasks. Furthermore, since participants received a tutorial beforehand, they were familiar with the search objectives and the meaning of each visual property. As a result, apart from the tooltip text, no additional text labels were included on the map.

Regarding the incidental visualization, due to the effectiveness of bar charts for retrieving values (Quadri and Rosen, 2022), and due to the subitizing mechanism (Chen et al., 2022) studied in incidental visualizations (Moreira et al., 2023b), we encoded the temperature using a horizontal bar chart with four bars, each with a green hue, where each one was aligned with the corresponding NUTS-II region. However, since there are five, we had to discard one. We discarded the 'Área Metropolitana de Lisboa' region (Fig. 1) because there were already several regions aligned with the third bar of the bar chart. Lastly, similar to the choropleth map, the bar chart also lacked labels and conveyed information solely through the length of the bars.

3.3. Tasks

Because of our between-subjects variable, each participant performed all nine trials with two, three, or four criteria. In every case, there was a criterion to **select a region with warm weather**. To satisfy it, participants needed

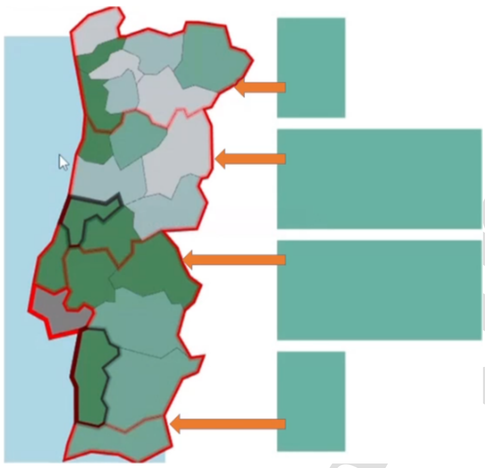


Figure 1: Example to show how each bar from the horizontal bar chart corresponded to each NUTS-II region. Each of these regions' boundaries was highlighted using a red stroke. The arrows were not shown in the study, only in this image. The gray region was the one we discarded to create only four bars. In this example, the NUTS-II regions with warm weather are the ones in the center.

to glance at an incidental visualization that they knew would appear while interacting with the choropleth map, thus ensuring it stayed hidden until it was needed. We wanted to simulate a scenario in which there is no more space in the choropleth map to encode further information, in this case, the temperature of NUTS-II regions.

Participants were only presented with the other criteria depending on the testing group. With two criteria, participants had to satisfy the one about the weather plus **selecting one NUTS-III region with a high population density**, which was encoded using a highly saturated green hue. With three criteria, participants had to satisfy the previous two stated plus **selecting one region with five or more hotels** in it, which was encoded using a text tooltip triggered by a mouse-over event. Finally, with four criteria, participants had to satisfy the previous three stated plus **selecting a region next to the sea**, which was encoded with a blue hue fill.

Before the primary task with the choropleth map, each participant first was presented with a set of instructions explaining how the study would be conducted. The first phase consisted of the OSPAN test. Since we only presented a bar chart with four bars, we designed the OSPAN to test the capability of the working memory for four items. From a set of 16 icons, and using a Latin square distribution, each participant was presented four icons (Fig. 2) and four math operations (like this one $(2 + 3) * 5$).

The workflow (Fig. 3) was, each icon was shown for one second, followed by the math operation shown for five seconds. After each operation, participants had to answer true or false questions regarding the answer to the corresponding operation. This process was repeated three more times, and at the end, each participant had to select the four icons in order of appearance. The OSPAN was calculated by weighting the score of the operations with the order of icons selected. For example, with two operations and two icons correct, the final score would be 0.25 $(0.5 * 0.5)$ out of 1.

The second phase commenced with a video tutorial that explained the process and objectives of the primary task to each participant. After viewing the tutorial, participants proceeded to engage in the primary task. The content of the tutorial and the specified goals varied depending on the between-subjects group to which each participant belonged, which encompassed two, three, or four criteria. Consequently, each participant exclusively performed the task under one of these conditions. Subsequently, participants completed the task under various combinations of within-subjects variables. Selecting one, two, or three regions (output) and completing the trial within 15, 20, or 25 seconds (pressure).

In total, each participant had to complete nine trials, which is the number of possible combinations of within-subjects variables. The order of each combination of trials was generated using a Latin square distribution. Using only the computer mouse, each participant had to explore the map to fulfill all the criteria. During this process, an incidental visualization appeared next to the choropleth map (Fig. 4), for one second, conveying information about the weather in each NUTS-II region. Participants were informed about the location where the visualization would appear and that it would be displayed for a duration of one second. However, they were not aware of the exact timing of its appearance. As the visualization was incidental in nature, it

Figure 2: Four of the 16 icons available.

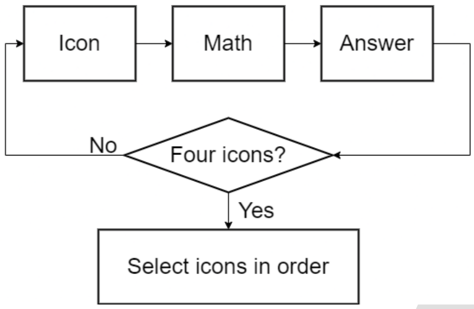


Figure 3: Flow of the OSPAN test.

did not appear immediately upon request. Instead, it appeared a few seconds after participants began hovering over the map. This timing was consistent for all participants. After every region was selected, participants had to submit their selection (Fig. 4), report the confidence they felt at submitting, and move on to the next trial. After nine trials, the study ended.

3.4. Participants

Because we wanted to have 40 participants for each group of our between-subjects variable we recruited 120 participants using a crowdsourcing platform (Prolific), and we used three prescreen filters. First, the sample needed to be gender balanced for each group of our between-subjects variable. Second, each participant needed to have normal or corrected-to-normal vision. Finally, every participant needed to have some video game experience (3-6 hours, 6–9 hours) ensure a diverse pool of individuals who were familiar with gaming. This familiarity would facilitate their interaction with an interface that involved time-restricted tasks. Of the 120 participants, 60 were female and 60 male, the median time of each participant was 9m08s, and each participant was paid £9/hour on average. This value was based on an average calculated using the mean time spent by all participants and the total amount of money paid. Furthermore, 86 participants were between 18 and 28 years old, 24 between 29 and 38, six between 38 and 48, and four between 48 and 58. People participated from 25 different countries, but 52.5% of participants were from Portugal, Poland, and South Africa.

3.5. Quality Assurance

Due to quality assurance, besides the 120 participants, three participations were rejected, 100 were returned, and seven were timed out. The rea-

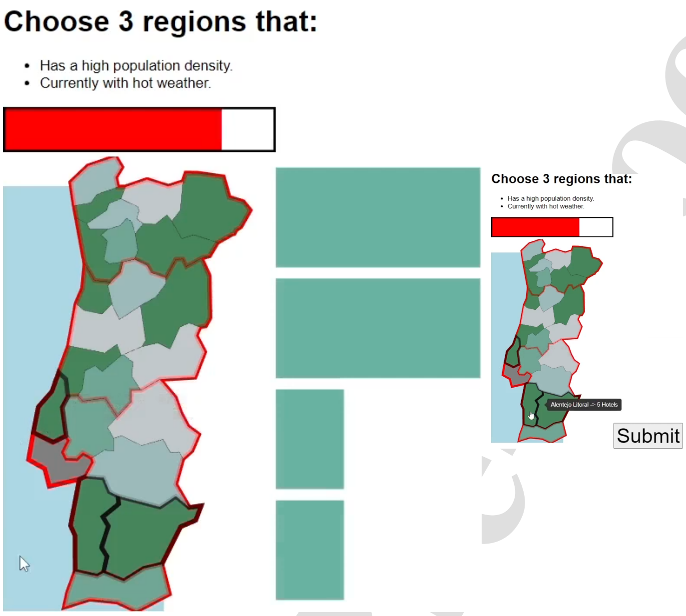


Figure 4: On the left, we can see the moment depicted shows the incidental visualization before disappearing, and the choropleth map with three regions selected, followed by the moment depicted on the right that shows the submit button that only appears after the incidental visualization disappears.

sons for rejection were either because participants became idle, or lost focus on the browser. The former was identified by incorporating timers at each stage of the study on the website. During the primary task, the timer was visually represented as a red bar positioned above the choropleth map. The latter was monitored by continuously checking the cursor's placement on the study website. For instance, if a participant minimized the browser or switched to another tab, their participation would be terminated.

Unfortunately, we had no method to avoid random clickers because we could not incorporate an attention check to our primary task without it being too obvious, but we prevented multiple participation using the ProlificID.

4. Results

The data analyzed in this user study is available online [dataset](Figshare). A three-way mixed ANOVA was conducted to examine the effects of the number of regions selected, time available to answer, and the number of restrictions on four dependent variables: accuracy, response time, confidence, and vis effectiveness. The assumption of normality was violated for all dependent variables. However, ANOVAs are considered robust to deviations from normality (Statistics, 2015). We analyzed all data and did not exclude any outliers as they represented genuine unusual values. The assumption of homogeneity of variances was violated for accuracy, response time, and vis effectiveness, but not for confidence. Nonetheless, the ANOVA was performed due to equal or approximately equal group sample sizes (Statistics, 2015).

The assumption of sphericity was met for some within-subject effects but violated for the within-subject effect of pressure, necessitating a Greenhouse-Geisser correction. The summary of results, including three-way and two-way interactions, and main effects, are in Table 1. The conclusions of each statistical result will be highlighted, and the corresponding consequences will be explained in Section 4.5, where we will explicitly state how much metrics decreased or increased

4.1. Three-Way Interaction

There was no significant three-way interaction between output, pressure, and input. The value of the dependent variables did not significantly depend on any combination of the other two independent variables.

4.2. Two-Way Interaction

There was no significant two-way interaction between output and input. The value of the dependent variables did not significantly depend on the input when choosing the output. Similarly, there was no significant interaction between pressure and input or between output and pressure. When considering the covariate of working memory, there was no significant interaction between output and pressure. The value of the dependent variables did not significantly depend on the input and was not affected by working memory.

4.3. Main Effects

The main effect of output was found to be statistically significant. With a greater number of regions available to select from, participants' performance at the primary tasks noticeably decreased, response time exhibited an increase, confidence levels decreased, and visualization accuracy saw a decrease. The main effect of pressure yielded significance in terms of accuracy, response time, and vis effectiveness. When provided with additional time to

Table 1: Summary of the three-way mixed ANOVA. Our three independent variables were input, output, and pressure, our covariate was the OSPAN, and our dependent variables were accuracy, response time, confidence, and vis effectiveness. The statistically significant differences are in bold.

	Accuracy (0-1)	Response Time (s)
input : output : pressure	F(8, 456) = 0.874, p = 0.538	F(8, 252) = 0.873, p = 0.539
output : input	F(4, 228) = 0.961, p = 0.430	F(4, 126) = 2.105, p = 0.084
pressure : input	F(3.702, 221.021) = 1.197, p = 0.313	F(3.034, 95.579) = 2.062, p = 0.110
output : pressure	F(4, 288) = 1.838, p = 0.120	F(4, 126) = 0.213, p = 0.931
output : pressure : OSPAN	F(4, 288) = 1.416, p = 0.227	F(4, 126) = 0.608, p = 0.657
output	**F(2, 228) = 5.596, p = 0.004**	**F(2, 126) = 13.154, p < .001**
output : OSPAN	F(2, 228) = 0.552, p = 0.557	F(2, 126) = 3.015, p = 0.053
pressure	**F(1.851, 211.021) = 7.177, p = 0.001**	**F(1.517, 95.579) = 5.782, p = 0.008**
pressure : OSPAN	**F(1.851, 211.021) = 4.073, p = 0.021**	F(1.517, 95.579) = 0.054, p = 0.905
input	F(2, 114) = 0.352, p = 0.704	F(2, 63) = 1.415, p = 0.205
	Confidence (1-5)	**Vis Effectiveness (0-1)**
input : output : pressure	F(8, 252) = 0.973, p = 0.458	F(8, 456) = 0.748, p = 0.649
output : input	F(4, 126) = 1.739, p = 0.145	F(4, 228) = 1.482, p = 0.209
pressure : input	F(3.360, 105.855) = 0.461, p = 0.731	F(3.702, 221.003) = 0.962, p = 0.425
output : pressure	F(4, 126) = 0.307, p = 0.873	F(4, 288) = 1.733, p = 0.142
output : pressure : OSPAN	F(4, 126) = 0.254, p = 0.907	F(4, 288) = 1.308, p = 0.266
output	**F(2, 126) = 6.708, p = 0.002**	**F(2, 228) = 7.329, p < .001**
output : OSPAN	**F(2, 126) = 4.448, p = 0.014**	F(2, 228) = 1.426, p = 0.242
pressure	F(1.680, 105.855) = 1.954, p = 0.154	**F(1.851, 211.003) = 10.051, p < .001**
pressure : OSPAN	F(1.680, 105.855) = 1.374, p = 0.256	**F(1.851, 211.003) = 5.603, p = 0.005**
input	F(2, 63) = 0.079, p = 0.924	F(2, 114) = 1.998, p = 0.140

complete the task, participants' performance at the primary tasks showed improvement, response time increased, confidence levels remained relatively unchanged, and visualization accuracy increased. The main effect of input did not reach statistical significance. The performance, response time, confidence, and visualization accuracy remained similar when faced with more criteria to satisfy.

4.4. Main Effects and Working Memory

When considering the covariate, the main effect of output showed no significant difference in accuracy, response time, and vis effectiveness, but was significant for confidence. The primary task performance was affected by working memory when the number of regions to select changed. The main effect of pressure was significant for accuracy, indicating that the primary task performance was affected by working memory when the time available to finish the task changed.

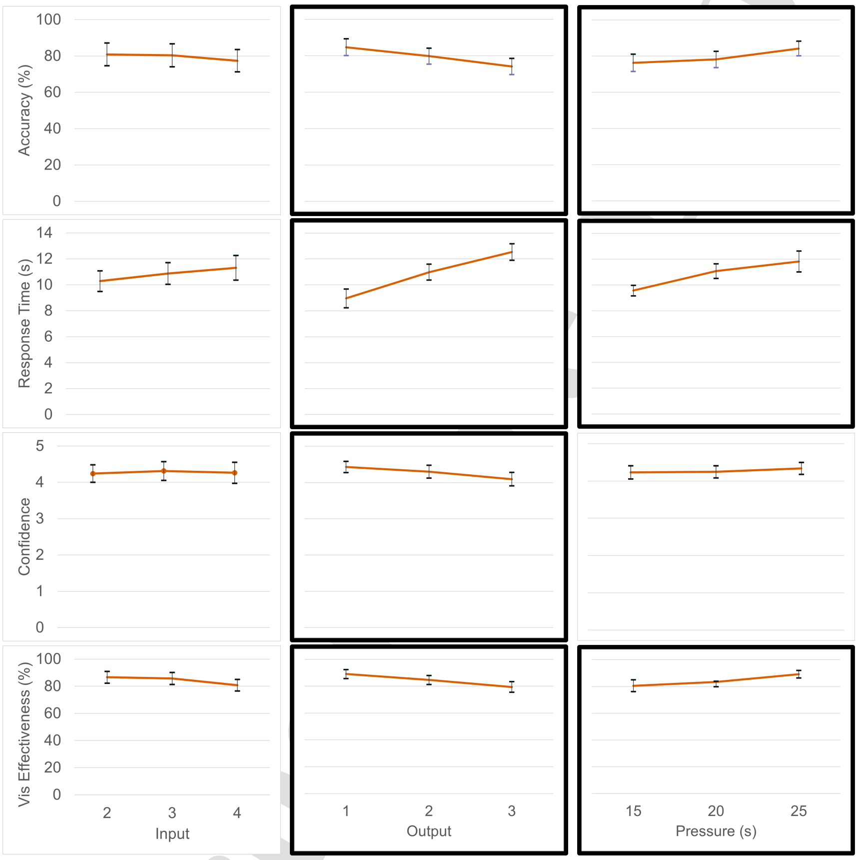


Figure 5: Line charts displaying the average values and corresponding confidence intervals (95%) of our dependent variables across all three levels of each independent variable. The line charts highlighted represent variables that have a significant impact, as indicated in Table 1.

4.5. Means

A summary of all measures is presented in Table 2, while Figure 5 provides a visual representation of all the data. We will present pairwise comparisons

Table 2: A summary of all means. Our three independent variables were input, output, and pressure, and our dependent variables were accuracy, response time, confidence, and vis effectiveness

.

	Output			Pressure			Input		
	1	**2**	**3**	**15s**	**20s**	**25s**	**2**	**3**	**4**
Accuracy (%)	84.70	79.80	74.10	76.30	78.20	84.20	80.90	80.40	77.40
Response Time (s)	8.955	10.983	12.538	9.572	11.08	11.825	10.284	10.875	11.317
Confidence (1-5)	4.427	4.300	4.091	4.232	4.248	4.338	4.242	4.312	4.264
Vis Effectiveness (%)	89.00	84.60	79.40	80.50	83.40	89.10	86.60	85.70	80.80

with statistical significance, employing the Bonferroni adjustment for multiple comparisons. For each variable, we will indicate the extent of variation along with the corresponding error.

Accuracy significantly decreased between output 1 and 3 (-0.106 ± 0.026, $p < 0.001$) and between output 2 and 3 (-0.057 ± 0.022, $p = 0.033$). Response time significantly increased between outputs 1 and 2 ($2.028s \pm 0.399s$, $p < 0.001$), between outputs 1 and 3 ($3.583s \pm 0.392s$, $p < 0.001$), and between outputs 2 and 3 ($1.555s \pm 0.343s$, $p < 0.001$). Confidence significantly decreased between output 1 and 3 (-0.336 ± 0.074, $p < 0.001$) and between output 2 and 3 (-0.208 ± 0.074, $p = 0.019$). The visualization score significantly decreased between output 1 and 3 (-0.096 ± 0.023, $p < 0.001$). Additionally, it increased significantly between pressure 1 and 3 (0.086 ± 0.023, $p < 0.001$) and between pressure 2 and 3 (0.058 ± 0.020, $p = 0.015$).

5. Discussion

We will address each research question we proposed. Then, we will provide some insights into our contribution to incidental visualizations.

RQ1 – How much input, output, and pressure complexity contributory factors impact the accuracy, response time, and confidence of the primary task?: By looking at Table 1, we can see at a glance that there were no three-way nor two-way statistically significant interactions. Therefore, for this primary task, we cannot prove that our complexity contributory factors impacted each other, even when considering each participant's working memory for retaining four items.

Regarding main effects, our results showed statistically significant differences in the output and pressure variables, but not in the input variable, which means that we cannot prove the number of criteria that needed to be satisfied (input) had an impact on accuracy, response time, or confidence. Although this may look promising because it might discard the impact of the input complexity factor, we were a bit skeptical. First, by analyzing the means when the input increased, accuracy decreased and the response time increased. This may indicate a possible trend for these metrics, even though there was no significant difference. Second, increments of one restriction (input) may not be enough to add enough complexity, which means that going from two to three/four criteria, or from three to four may not be enough to induce a significant difference. Third, by restricting the possible regions to be selected, like we did when people had to choose one or more regions next to the sea, we may have facilitated the task. We argue that these issues should be addressed in future work.

In any case, regarding the output and pressure main effects, both presented significant differences. With more regions to select (output), accuracy and confidence decreased, and response time increased. Particularly for confidence, the working memory impacted the results. Then, with less time to complete the task (pressure), all metrics except confidence significantly decreased. Particularly for accuracy, the working memory impacted the results. In summary, output and pressure are two complexity contributory factors that significantly impacted accuracy, response time, and confidence of the primary task. The results aligned with our predictions.

RQ2 – How much input, output, and pressure complexity contributory factors impact the perception accuracy of the incidental visualization? Results were similar to the other previous three dependent variables. No interactions between input, output, and pressure, but there were significant main effects. When the number of regions selected increased, and when pressure increased, the less accurate was the perception of the visualization. Furthermore, working memory impacted the results when pressure differed. In summary, output and pressure are two complexity contributory factors that significantly impacted the perception of the incidental visualization. Again, the results aligned with our predictions, and we can conclude that we proved our hypothesis: **modifying task complexity factors will result in noticeable impacts on task performance and the perception of incidental visualizations.**

Overall Performance: By looking at the charts, the mean accuracy was always above 70%, which we consider still a high value. Therefore, primary task performance was usually high. Confidence was always above 4, which indicated that performing this type of task plus having an incidental visualization did not make participants feel less confident. Finally, the response time was always below 14 seconds, which is below the lowest pressure time we tested (15s). Therefore, although people sometimes had 20 or 25 seconds to complete the task, on average, people took less than 15 seconds. Finally, by looking at the vis effectiveness chart, accuracy was always above 79%. Therefore, regardless of the overall performance of the primary task, the incidental visualization was usually perceived accurately.

5.1. Insights for Incidental Visualizations

Having answered all proposed research questions, we now can discuss these results in the context of the current state of the art. The first two questions were about the impact of three complexity factors, and the last two were about overall results.

At their core, our results proved that the primary task, on average, was performed accurately with some required information being displayed using an incidental visualization. Therefore, we argue that incidental visualizations can convey additional information while people are focusing on performing a primary task without disrupting it. Furthermore, having this information shown only for a specific moment did not affect people's confidence. Moreover, our results match the effectiveness predicted by using the horizontal bar chart because the search tasks people went through were supported by a bar chart whose encodings are technically the most effective. Therefore, the bar chart chosen as an incidental visualization proved to be an effective choice. Finally, we were able to conclude that people's working memory should be taken into consideration. Although our independent variables did not always interact, future work should dive more into detail about this type of individual difference.

Regarding task complexity, although Liu and Li's framework (Liu and Li, 2012) contemplated many complexity contributory factors, the ones we tested were relevant enough to be taken into consideration in this primary task with incidental visualizations. However, many factors were left out that may be relevant for specific primary tasks. Nonetheless, we argue that these three may be easier to use in defining tasks because they relate closely to the task's goals, restrictions, and time available.

In summary, incidental visualizations were effectively perceived during an ongoing task, while allowing people to perform it accurately, quickly, and with high confidence levels. Although our scenario is too specific, the major implications of our study are:

- Information needed for a primary task conveyed using an incidental visualization can be effectively perceived.
- Performance in a primary task will not be disrupted by a horizontal bar chart as an incidental visualization.
- Working memory should be taken into consideration when using incidental visualizations with a primary task.
- Output and pressure are complexity contributory factors that will significantly impact the overall results of the primary task.

5.2. Limitations

Although our results show promise, there are several limitations. The primary one is the empirical nature of our study, which stems from the fact that we have dealt with a toy problem scenario. While we have successfully demonstrated the effectiveness of incidental visualizations in this simplified context, it's crucial to recognize that our investigation has been confined to a specific primary task and visualization. In the future, it will be interesting to test other combinations to better delineate the frontiers of the design space for effective incidental visualization use.

Furthermore, as we stated previously, our input variable should probably be tested with more levels, with a higher complexity gap between, or with a different nature. We cannot conclude if having more criteria facilitated the task, or if three levels are enough to retrieve significant differences. We also acknowledge that having each participant perform only once for each of the nine combinations may have introduced some level of noise in the data. Then, the confidence measure used in our study was treated as a quantitative variable in the analysis, despite being an ordinal variable. This decision was made to include the Likert-format variable in the three-way mixed ANOVA test, although we recognize that this approach may deviate from strict statistical conventions.

Also, each participant only performed each one of the nine trials combination once. Although participants knew beforehand what they needed to do, repeating trials could potentially reduce the noise in the final data collected. Nest, other complexity factors can still be studied. We only tested three that are easily mapped to restrictions, goals, and time, but there are others from the framework we used that might have the same or bigger importance.

Finally, conducting the experiment in a desktop setting was largely due to the nature of the toy problem. Since this is a novel area of research, our goal was to empirically assess first whether incidental visualizations are viable. As mentioned in the Introduction, augmented reality appears to be a promising avenue for further exploration in this field.

6. Conclusion

Incidental visualizations are still a novel topic. Although they share core concepts with glanceable visualizations, they differ in when and how the information depicted is perceived. We conducted an online user study to understand how a specific set of task complexity factors interfere with a particular real-case primary task performance and incidental visualization perception accuracy. Participants had to explore a choropleth map to retrieve regions that satisfied a set of criteria and glance at an incidental visualization to retrieve important information. In the end, we concluded that people usually had high task performance and high perception accuracy of the incidental visualization they were exposed to. For future work, it would be interesting to understand which graph characteristics influence the factors' effect. What happens if another graph is used? Then, the input variable should be explored with different values. For which number of criteria will the results significantly start to differ? Finally, different real-life tasks should be tested because at this point incidental visualizations are not yet restricted to specific contexts.

Acknowledgments

This work was supported by FCT, through grants:

- PTDC/CCI-CIF/28939/2017
- UIDB/50021/2020
- SFRH/BD/143496/2019

I would also like to thank my colleague Tomás Alves for his help in discussing the work, and with some minor revisions.

References

Ankrah, E.A., Cibrian, F.L., Silva, L.M., Tavakoulnia, A., Beltran, J.A., Schuck, S.E., Lakes, K.D., Hayes, G.R., 2022. Me, my health, and my watch: How children with adhd understand smartwatch health data. ACM Trans. Comput.-Hum. Interact. doi:10.1145/3577008. just Accepted.

Bertini, E., Correll, M., Franconeri, S., 2020. Why shouldn't all charts be scatter plots? beyond precision-driven visualizations, in: 2020 IEEE Visualization Conference (VIS), pp. 206–210. doi:10.1109/VIS47514.2020.00048.

Blascheck, T., Bentley, F., Choe, E.K., Horak, T., Isenberg, P., 2021. Characterizing glanceable visualizations: From perception to behavior change, in: Mobile Data Visualization. Chapman and Hall/CRC, pp. 151–176.

Blascheck, T., Besançon, L., Bezerianos, A., Lee, B., Isenberg, P., 2019. Glanceable visualization: Studies of data comparison performance on smartwatches. IEEE Transactions on Visualization and Computer Graphics 25, 630–640. doi:10.1109/TVCG.2018.2865142.

Blascheck, T., Isenberg, P., 2021. A Replication Study on Glanceable Visualizations: Comparing Different Stimulus Sizes on a Laptop Computer, in: IVAPP 2021 - 12th International Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, SCITEPRESS - Science and Technology Publications, Virtual, Austria. pp. 133–143. URL: https://hal.inria.fr/hal-03151924, doi:10.5220/0010328501330143.

Borgo, R., Lee, B., Bach, B., Fabrikant, S., Jianu, R., Kerren, A., Kobourov, S., McGee, F., Micallef, L., von Landesberger, T., Ballweg, K., Diehl, S., Simonetto, P., Zhou, M., 2017. Crowdsourcing for information visualization: Promises and pitfalls, in: Archambault, D., Purchase, H., Hoßfeld, T. (Eds.), Evaluation in the Crowd. Crowdsourcing and Human-Centered Experiments, Springer International Publishing, Cham. pp. 96–138.

Borgo, R., Micallef, L., Bach, B., McGee, F., Lee, B., 2018. Information visualization evaluation using crowdsourcing. Computer Graphics Forum 37, 573–595. doi:10.1111/cgf.13444, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13444.

Brehmer, M., Munzner, T., 2013. A multi-level typology of abstract visualization tasks. IEEE Transactions on Visualization and Computer Graphics 19, 2376–2385. doi:10.1109/TVCG.2013.124.

Byström, K., Järvelin, K., 1995. Task complexity affects information seeking and use. Information Processing & Management 31, 191–213. URL: https://www.sciencedirect.com/science/article/pii/030645739580035R, doi:10.1016/0306-4573(95)80035-R.

Campbell, D.J., 1988. Task complexity: A review and analysis. The Academy of Management Review 13, 40. doi:10.2307/258353.

Chen, J., Paul, J.M., Reeve, R., 2022. Manipulation of attention affects subitizing performance: A systematic review and meta-analysis. Neuroscience & Biobehavioral Reviews 139, 104753. URL: https://www.sciencedirect.com/science/article/pii/S0149763422002421, doi:10.1016/j.neubiorev.2022.104753.

Cleveland, W.S., McGill, R., 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. Journal of the American Statistical Association 79, 531–554. doi:10.1080/01621459.1984.10478080, arXiv:https://www.tandfonline.com/doi/pdf/10.1080/01621459.1984.10478080.

Cleveland, W.S., McGill, R., 1985. Graphical perception and graphical methods for analyzing scientific data. Science 229, 828–833. doi:10.1126/science.229.4716.828, arXiv:https://www.science.org/doi/pdf/10.1126/science.229.4716.828.

Cleveland, W.S., McGill, R., 1986. An experiment in graphical perception. International Journal of Man-Machine Studies 25, 491–500. URL: https://www.sciencedirect.com/science/article/pii/S0020737386800190, doi:10.1016/S0020-7373(86)80019-0.

Cleveland, W.S., McGill, R., 1987. Graphical perception: The visual decoding of quantitative information on graphical displays of data. Journal of the Royal Statistical Society: Series A (General) 150, 192–210. doi:10.2307/2981473, arXiv:https://rss.onlinelibrary.wiley.com/doi/pdf/10.2307/2981473.

ColorBrewer, . ColorBrewer. https://colorbrewer2.org. Accessed on November 14, 2023.

Danner-Schröder, A., Ostermann, S.M., 2022. Towards a processual understanding of task complexity: Constructing task complexity in practice. Organization Studies 43, 437–463. doi:10.1177/0170840620941314, arXiv:https://doi.org/10.1177/0170840620941314.

van Dantzich, M., Robbins, D., Horvitz, E., Czerwinski, M., 2002. Scope: Providing awareness of multiple notifications at a glance, in: Proceedings of the Working Conference on Advanced Visual Interfaces, Association for Computing Machinery, New York, NY, USA. pp. 267–281. doi:10.1145/1556262.1556306.

Daskalogrigorakis, G., McNamara, A., Marinakis, A., Antoniadis, A., Mania, K., 2022. Glance-box: Multi-lod glanceable interfaces for machine shop guidance in augmented reality using blink and hand interaction, in: 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 315–321. doi:10.1109/ISMAR-Adjunct57072.2022.00070.

Davari, S., Lu, F., Bowman, D.A., 2020. Occlusion management techniques for everyday glanceable ar interfaces, in: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pp. 324–330. doi:10.1109/VRW50115.2020.00072.

Davari, S., Lu, F., Bowman, D.A., 2022. Validating the benefits of glanceable and context-aware augmented reality for everyday information access tasks, in: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 436–444. doi:10.1109/VR51125.2022.00063.

Davis, R., Pu, X., Ding, Y., Hall, B.D., Bonilla, K., Feng, M., Kay, M., Harrison, L., 2022. The risks of ranking: Revisiting graphical perception to model individual differences in visualization performance. IEEE Transactions on Visualization and Computer Graphics , 1–16doi:10.1109/TVCG.2022.3226463.

DEURZEN, B.V., Hendrikx, M., VANACKEN, D., LUYTEN, K., 2022. Work-a-pose: Ergonomic feedback and posture improvement interfaces for long-term sustainable work. International Conference on Advanced Visual Interfaces.

Efatmaneshnik, M., Handley, H.A.H., 2021. Task complexity measurement framework for human systems integration. IEEE Systems Journal 15, 2787–2797. doi:10.1109/JSYST.2020.3000499.

EurostatNUTS, . Eurostat Nomenclature of Territorial Units for Statistics (NUTS) Maps. https://ec.europa.eu/eurostat/web/nuts/nuts-maps. Accessed on November 14, 2023.

Faulhaber, A.K., Hoppe, M., Schmidt, L., 2022. Evaluation of priority-dependent notifications for smart glasses based on peripheral visual cues. i-com 21, 239–252. doi:10.1515/icom-2022-0022.

Figshare, . Dataset. https://figshare.com/s/175d6247034b327fa62e.

Gouveia, R., Pereira, F., Karapanos, E., Munson, S.A., Hassenzahl, M., 2016. Exploring the design space of glanceable feedback for physical activity trackers, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Association for Computing Machinery, New York, NY, USA. pp. 144–155. doi:10.1145/2971648.2971754.

Grioui, F., Blascheck, T., 2021. Study of heart rate visualizations on a virtual smartwatch, in: Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology, Association for Computing Machinery, New York, NY, USA. pp. 1–3. URL: https://doi.org/10.1145/3489849.3489913, doi:10.1145/3489849.3489913.

Islam, A., Aravind, R., Blascheck, T., Bezerianos, A., Isenberg, P., 2022a. Preferences and effectiveness of sleep data visualizations for smartwatches and fitness bands, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. pp. 1–17. URL: https://doi.org/10.1145/3491102.3501921, doi:10.1145/3491102.3501921.

Islam, A., Blascheck, T., Isenberg, P., 2022b. Context Specific Visualizations on Smartwatches. Posters of the European Conference on Visualization (EuroVis). URL: https://hal.inria.fr/hal-03694026, doi:10.2312/evp.20221122. poster.

Islam, A., Yao, L., Bezerianos, A., Blascheck, T., He, T., Lee, B., Vuillemot, R., Isenberg, P., 2022c. Reflections on Visualization in Motion for Fitness Trackers, in: MobileHCI 2022 - 24th International Conference on Mobile Human-Computer Interaction, Vancouver, Canada. URL: https://hal.inria.fr/hal-03775633.

Janaka, N., Haigh, C., Kim, H., Zhang, S., Zhao, S., 2022. Paracentral and near-peripheral visualizations: Towards attention-maintaining secondary information presentation on ohmds during in-person social interactions, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. pp. 1–14. URL: https://doi.org/10.1145/3491102.3502127, doi:10.1145/3491102.3502127.

Kaufman, E.L., Lord, M.W., Reese, T.W., Volkmann, J., 1949. The discrimination of visual number. The American Journal of Psychology 62, 498. doi:10.2307/1418556.

Lazzara, E.H., Pavlas, D., Fiore, S., Salas, E., 2010. A framework to develop task complexity. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 54, 2338–2342. doi:10.1177/154193121005402725, arXiv:https://doi.org/10.1177/154193121005402725.

Liu, P., Li, Z., 2011. Toward understanding the relationship between task complexity and task performance, in: Rau, P.L.P. (Ed.), Internationalization, Design and Global Development, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 192–200.

Liu, P., Li, Z., 2012. Task complexity: A review and conceptualization framework. International Journal of Industrial Ergonomics 42, 553–568. URL: https://www.sciencedirect.com/science/article/pii/S0169814112000868, doi:10.1016/j.ergon.2012.09.001.

Liu, Z., Crouser, R.J., Ottley, A., 2020. Survey on individual differences in visualization. Computer Graphics Forum 39, 693–712. doi:10.1111/cgf.14033, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14033.

Lu, F., 2021. [dc] glanceable ar: Towards an always-on augmented reality future, in: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces

Abstracts and Workshops (VRW), pp. 717–718. doi:10.1109/VRW52623.2021.00241.

Lu, F., Bowman, D.A., 2021. Evaluating the potential of glanceable ar interfaces for authentic everyday uses, in: 2021 IEEE Virtual Reality and 3D User Interfaces (VR), pp. 768–777. doi:10.1109/VR50410.2021.00104.

Lu, F., Davari, S., Bowman, D., 2021. Exploration of techniques for rapid activation of glanceable information in head-worn augmented reality, in: Proceedings of the 2021 ACM Symposium on Spatial User Interaction, Association for Computing Machinery, New York, NY, USA. pp. 1–11. URL: https://doi.org/10.1145/3485279.3485286, doi:10.1145/3485279.3485286.

Matthews, T., 2006. Designing and evaluating glanceable peripheral displays, in: Proceedings of the 6th Conference on Designing Interactive Systems, Association for Computing Machinery, New York, NY, USA. pp. 343–345. doi:10.1145/1142405.1142457.

Maynard, D.C., Hakel, M.D., 1997. Effects of objective and subjective task complexity on performance. Human Performance 10, 303–330. doi:10.1207/s15327043hup1004_1, arXiv:$https://doi.org/10.1207/s15327043hup1004_1$.

Moreira, J., Mendes, D., Gonçalves, D., 2023a. Impact of incidental visualizations on primary tasks. Information Visualization 0, 14738716231180892. doi:10.1177/14738716231180892, arXiv:https://doi.org/10.1177/14738716231180892.

Moreira, J., Mendes, D., Gonçalves, D., 2023b. Incidental graphical perception: How marks and display time influence accuracy. Information Visualization , 14738716231189218.

Moreira, J.a., Mendes, D., Gonçalves, D., 2020. Incidental visualizations: Pre-attentive primitive visual tasks, in: Proceedings of the 2020 International Conference on Advanced Visual Interfaces, Association for Computing Machinery, New York, NY, USA. pp. 1–9. URL: https://doi.org/10.1145/3399715.3399841, doi:10.1145/3399715.3399841.

Neshati, A., Sakamoto, Y., Leboe-Mcgowan, L., Leboe-Mcgowan, J., Serrano, M., Irani, P., 2019. G-Sparks: Glanceable Sparklines on Smartwatches, in: 45th Conference on Graphics Interface (GI 2019), Kingston, Ontario, Canada. pp. 1–9. URL: https://hal.archives-ouvertes.fr/hal-02891743, doi:10.20380/GI2019.23.

Panavas, L., Worth, A.E., Crnovrsanin, T., Sathyamurthi, T., Cordes, S., Borkin, M.A., Dunne, C., 2022. Juvenile graphical perception: A comparison between children and adults, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. pp. 1–14. URL: https://doi.org/10.1145/3491102.3501893, doi:10.1145/3491102.3501893.

Prolific, . Prolific. https://www.prolific.co/. Accessed on November 14, 2023.

Quadri, G.J., Rosen, P., 2022. A survey of perception-based visualization studies by task. IEEE Transactions on Visualization and Computer Graphics 28, 5026–5048. doi:10.1109/TVCG.2021.3098240.

Schiewe, A., Krekhov, A., Kerber, F., Daiber, F., Krüger, J., 2020. A study on real-time visualizations during sports activities on smartwatches, in: 19th International Conference on Mobile and Ubiquitous Multimedia, Association for Computing Machinery, New York, NY, USA. pp. 18–31. doi:10.1145/3428361.3428409.

Skog, T., 2004. Activity wallpaper: Ambient visualization of activity information, in: Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, Association for Computing Machinery, New York, NY, USA. pp. 325–328. doi:10.1145/1013115.1013171.

Statistics, L., 2015. Three-way mixed anova (with one between-subjects and two within-subjects factors) using spss statistics. Retrieved from. URL: https://statistics.laerd.com/. statistical tutorials and software guides.

Turner, M.L., Engle, R.W., 1989. Is working memory capacity task dependent? Journal of Memory and Language 28, 127–154. URL: https://www.

sciencedirect.com/science/article/pii/0749596X89900405, doi:10.1016/0749-596X(89)90040-5.

Weiser, M., Brown, J.S., 1996. Designing calm technology, Citeseer. pp. 75–85.

Wilkinson, D., Bahirat, P., Namara, M., Lyu, J., Alsubhi, A., Qiu, J., Wisniewski, P., Knijnenburg, B.P., 2020. Privacy at a glance: The user-centric design of glanceable data exposure visualizations. Proceedings on Privacy Enhancing Technologies 2020, 416–435. doi:10.2478/popets-2020-0034.

Wood, R.E., 1986. Task complexity: Definition of the construct. Organizational Behavior and Human Decision Processes 37, 60–82. URL: https://www.sciencedirect.com/science/article/pii/0749597886900440, doi:10.1016/0749-5978(86)90044-0.

Wood, R.E., Mento, A.J., Locke, E.A., 1987. Task complexity as a moderator of goal effects: A meta-analysis. Journal of Applied Psychology 72, 416–425. doi:10.1037/0021-9010.72.3.416.

Zeng, Z., Battle, L., 2021. Are we there yet? a systematic review of visual perception knowledge for visualization recommendation systems. doi:10.48550/ARXIV.2109.01271, arXiv:2109.01271.

Zhang, L., Tu, N., Vronay, D., 2005. Info-lotus: A peripheral visualization for email notification, in: CHI '05 Extended Abstracts on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. pp. 1901–1904. doi:10.1145/1056808.1057051.

Zhang, Z., Pan, Z., Li, W., Su, Z., 2022. X-board: an egocentric adaptive AR assistant for perception in indoor environments. Virtual Reality doi:10.1007/s10055-022-00742-3.

Author Contributions Section

João Moreira: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization

Daniel Mendes: Conceptualization, Methodology, Writing - Review & Editing, Supervision

Daniel Gonçalves: Conceptualization, Methodology, Writing - Review & Editing, Supervision

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: