

# How do users interact with Virtual Geographic Environments? Users' behavior evaluation in urban participatory planning

Thibaud Chassin<sup>a,b,\*</sup>, Jens Ingensand<sup>b</sup>, Guillaume Touya<sup>c</sup>, Sidonie Christophe<sup>c</sup>

<sup>a</sup> Laboratory of Geographic Information Systems (LaSIG), EPFL, Lausanne, Switzerland - [thibaud.chassin@epfl.ch](mailto:thibaud.chassin@epfl.ch),

<sup>b</sup> InsIT, University of Applied Sciences and Arts Western Switzerland (HEIG-VD), Yverdon-Les-Bains, Switzerland - [\[firstname.lastname\]@heig-vd.ch](mailto:[firstname.lastname]@heig-vd.ch)

<sup>c</sup> LASTIG, Univ Gustave Eiffel, ENSG, IGN, F-94160 Saint-Mandé, France - [{firstname.lastname}@ign.fr](mailto:{firstname.lastname}@ign.fr)

\* Corresponding author

**Abstract:** For the past twenty years, the adoption of Virtual Geographic Environments is thriving. This democratization is due to numerous new opportunities offered by this medium. However, in participatory urban planning these interactive 3D geovisualizations are still labeled as very advanced means, and are only scarcely used. The involvement of citizens in urban decision-making is indeed carefully planned ahead to limit off-topic feedback. A better comprehension of Virtual Geographic Environments, and more specifically of users' strategic behaviors while interacting with this medium could enhance participants' contributions. The users' strategic behavior was assessed in this article through an experimental study. A total of 107 participants completed online tasks about the identification of 3D scenes' footprints, the comparison of buildings' heights, and the visibility of objects through the scenes. The interactions of the participants were recorded (i.e. pressed keys, pointing device interactions), as well as the camera positions adopted to complete specific tasks. The results show that: (1) users get more efficient throughout the study; (2) interruptions in 3D manipulation appear to highlight difficulties in interacting with the virtual environments; (3) users tend to centralize their positions within the scene, notably around their starting position; (4) the type of task strongly affects the behavior of users, limiting or broadening their explorations. The results of this experimental study are a valuable resource that can be used to improve the design of future urban planning projects involving Virtual Geographic Environments, e.g. with the creation of personalized 3D tools.

**Keywords:** Virtual Geographic Environments, 3D geovisualization, strategic behavior, user interactions, participatory urban planning

## 1. Introduction

The world-wide presence of Google Earth<sup>1</sup> and the development of highly realistic video games or Computer-Generated Imagery (CGI) place non-experts as regular consumers of 3D, also considered as Virtual Geographic Environments (VGEs) (Lin and Gong 2001). Additionally, the development of intuitive techniques to interact with these digital 3D objects lowers their entry cost (Jankowski and Hachet 2015). VGEs are, therefore, gradually perceived as a tool not limited to expert usage. This shift inspires the design of several applications that are based on the participation of a broad, non-expert population. This interest is experienced by a large range of sectors (Biljecki et al. 2015), including urban planning, where the benefits of 3D geo-visualizations have been acknowledged for twenty years (Al-Kodmany 2002). VGEs considered as 3D communication medium provide a valuable opportunity to fulfil the primary objectives of participatory approaches in urban planning, namely participants gaining knowledge and broadening their perspectives about the urban project (Joerin et al. 2009). However, the authorities currently disregard the adoption of VGEs in a real-world context because of concerns about the usability of these tools by citizens. The position

of the authorities is thus not in line with the previous considerations. In this study, we argue that a better understanding of how users interact with VGE could address the authorities' concerns and ultimately democratize the practice of VGEs in urban participatory practices.

Through an experimental study, we aim to investigate the aspects of VGEs affecting users' behavior developed when interacting with a 3D scene. Firstly, some aspects of VGEs in participatory planning other than users' behavior related studies are presented. Next, we describe the experimental setup designed to evaluate the behavioral aspects. Then, the results are presented regarding four elements encapsulating the mentioned behavioral aspects. Lastly, we extensively discuss the results and their significance for the practice of VGEs in urban participatory planning (and other sectors).

## 2. Related work

Several aspects of VGEs are favorable for participatory practices. They facilitate communication to a broad public and help the appropriation of a hypothetical future urban project by all actors of a participatory approach: experts, politicians or population (Jacquino and Bonaccorsi 2019). The visualization of these future projects in their surroundings also supports the

<sup>1</sup> [google.com/earth/](https://www.google.com/earth/)

convergence of users' mental representations in one shared understanding that is common to all users (van der Land et al. 2013). This direct overview of a project proposal tends to limit actors' misconception and heterogeneous understanding, which often comes from mental representations based on a conceptual description such as 2D maps or texts. Moreover, the interactivity provided by VGEs encourages users to explore the 3D scene by using their own perspectives and vantage points, which reduces biases introduced by predefined point of views (Downes and Lange 2015). This exploration also improves the perception and the understanding of the 3D scene's depths, heights, distances, etc (Dübel and Schumann 2017; Sheppard and Cizek 2009; Herman, Juřík, et al. 2018).

However, while interacting with VGEs experts still perform better than laypersons (Herman, Řezník, et al. 2018), who are often accustomed to virtual environments only as passive customers. Indeed, interactivity increases the complexity and the cognitive load conveyed by 3D scenes, which are already cognitively intensive (van der Land et al. 2013). The resulting users' fatigue leads to less effective interaction with the VGE, which may constrain the task(s) to perform. This fatigue varies between users according to specific socio-demographic characteristics, such as age or spatial cognition abilities (Stanney, Mourant, and Kennedy 1998).

The design of VGEs, as a 3D geovisualization tool, implies several challenges articulated around three categories: the *data* (co-visualization of information from various sources on the same medium); the *users* (communication of a consistent (neutral) message to heterogeneous users, having unique skills, cognition abilities, objectives, and interactive behavior); the *representation* (depiction of information with its set of style and interaction) (Christophe 2020; Çöltekin et al. 2017). The design of (non-expert) citizen involvement tools should carefully consider these categories, in order to limit the risk of enacting inappropriate VGE practices, which could lead to bias, misinterpretation, inequality in its application (Schroth et al. 2011), and ultimately poor decision-making. A better understanding of how these categories are connected could limit these hindrances and enhance participatory practices.

This research implements an experimental study to investigate one possible connection between two of these categories: representation and users. This connection is evaluated by assessing the development of users' strategic behavior when interacting with VGEs. The users' behavior has already been studied in the literature in an interactive 3D context (Wilkening and Fabrikant 2013; Ugwitz et al. 2019; Herman and Stachoň 2016), notably by the reconstruction of a *trajectory visualization* (or user's camera path) from raw users' inputs (Herman, Řezník, et al. 2018). An analogous method will be used in this study.

### 3. Methodology and Experimental study

#### 3.1 Research questions

Designing an urban participatory approach is challenging for authorities. They may experience a tense situation with citizens, which will decide the future of a project. Also, their reputation may be jeopardized, especially if the selected medium proves inappropriate. The authorities are, thus, concerned about the usability of VGEs in a participatory context. These concerns could be addressed by a better knowledge of the interaction between 3D scenes and the users' behavior (understanding, hesitation, performance, inputs, etc). Therefore, investigating how the 3D representation (in terms of style and detail) affects the users' behavior while interacting with the VGE could contribute to this knowledge. This knowledge can promote the design of VGE that is responsive according to users' idiosyncrasies or the tasks to perform by assisting users with customized assistance, for instance. In this study, the users' behavior will be analysed via a trajectory visualisation over four aspects:

(1) *Learning curve*: VGEs are known to have a rapid learning curve (Zhang and Moore 2014). Users indeed get accustomed to the 3D medium, i.e. its controls, rules, portrayals, and specificities. Therefore, we expect users to develop fast understanding of this medium that translates into the development of strategies to get more effective independently from the 3D scene representation.

(2) *Uncertainties*: Manipulating a VGE is challenging for users, even if the controls are kept simple. Times when users are not interacting with the VGE are observed in the literature (Herman, Řezník, et al. 2018). We presume this inactivity to be linked to uncertainties or time for mentally rebuilding an understanding of the 3D scene, which could be related to the user's socio-demographic characteristics (age, previous experiences with VGEs).

(3) *Anchor points*: Users create mental images of the scene that anchor their cognition (Couclelis et al. 1987). If the task requires the collection of supplementary information, we suppose users will aim to limit the development of these mental images to ease their cognitive burden.

(4) *Impact of the task*: Distinct tasks require users' specific information to gather in a particular representation. Therefore, we expect that the nature of the task to complete should highly channel users' interactions with the VGE.

#### 3.2 Experiment design

This research presents the second phase of an experimental study aiming at assessing participants' performances with VGEs<sup>2</sup>. The VGEs were built from reduced city models with 16-19 buildings of various representations. The participants were asked to perform tasks articulated around participatory session interests, such as heights, angles of view, parallels between 2D and 3D representation. In total, 18 questions were set up for the overall experiment with a combination of tasks,

<sup>2</sup> The first phase is presented in a different article under submission

interactive methods (or metaphors) and representations. The study required an approximate time of 20 minutes for completion. Hereafter, we will adopt the term *batch* to describe all the tasks in a specific representation and metaphor.

**VGE representations.** Fig. 1 shows all four representations adopted in this study. These representations are based on two Levels of Detail (LoD): LoD 1 and LoD 2 (Biljecki, Ledoux, and Stoter 2016). On these raw LoDs, there were three types of drapings enveloping the buildings: (1) *color*, where the dominant color of the building from the satellite images was assessed visually and applied to the model; (2) *texture*, i.e. a photo-realistic texture applied to the building; (3) *focus*, where the color method was applied to each building, but one, the highest, which was colored white with its roof bright red. We introduced the last representation to investigate the impact of visual cues on users' behavior. The four representations were named: *LoD 1 & color*, *LoD 2 & color* (control scene), *LoD 2 & texture*, *LoD 2 & focus* (see Fig. 1).

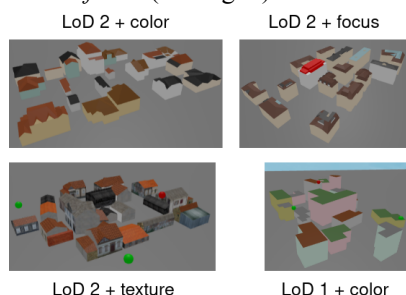


Figure 1. Illustration of the VGE representations used in this study for the first four batches.

**Metaphors.** The metaphors implemented for the study were aiming at simulating participatory session setups: (1) “*unrestricted*” interactions (batches 1 to 4, Fig. 2), where users could freely manipulate the VGEs without constraints of time or number of inputs; (2) *memorization* (batch 5, Fig. 2), where the 3D scene was not visible when participants were answering the questions related to a task – this setup aims to simulate the projection of a video, seen only once, before debating topics specific to a participatory session; (3) *static* (batch 6, Fig. 2), where the 3D scene was not interactive, simulating the use of photomontage in a participatory setup. The last metaphor will not be discussed in this study. Participants could manipulate the interactive VGEs via two degrees of freedom (rotation: pitch and yaw) with the arrow keys of the keyboard. The pitch was locked from 0° (horizontal view) to 45° (oblique view) to avoid viewing the 3D scene from under, or from a vertical view (being too similar to the 2D layout of maps). The distance to the VGEs (zoom) was also locked. These interactive rules were introduced to limit the entry cost, and ultimately open the study to anyone with or without technical skills.

**Tasks.** Participants were asked to complete three kinds of tasks related to urban participatory planning interests. The layout of the pages was similar: a question was asked on the top of the page, on the left was the VGE, with the

alternative answers on the right. The three kinds of tasks were: (1) *guess layout*, where the participants were asked to identify which one of 2D map layouts corresponded to the 3D scene; (2) *guess highest* or *guess lowest*, which was about building heights – participants were invited to select the highest or the lowest building on the 2D layout that was previously chosen; (3) *guess viewpoint*, where three spheres were located within the VGE (one was red, the two others were green) – participants were asked to determine which of the green spheres were visible from the red one (both, the closest, the farthest, none).

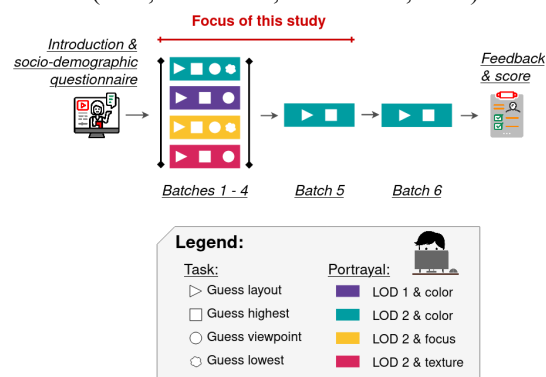


Figure 2. Detailed outline of the user study.

### 3.3 Collecting User interactions

While interacting with the VGE to complete their tasks, the inputs of the participants were recorded and transmitted to a database. For instance, a participant pressing the key ← triggers a record of the starting position, the ending position, the duration of the input and the orientation of the rotation (here left). An aggregated result of this raw data is depicted in Fig. 3: all camera positions for a specific task and representation are displayed.

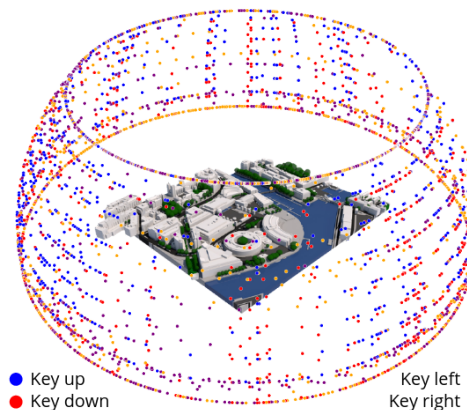


Figure 3. Raw camera positions of users to answer a specific task, in this case, *Guess highest* on the portrayal *LoD 2 + color*. The participants control the VGE via vertical and horizontal rotations, forming the spherical segment shape.

The shape of a spherical segment is constrained by the rules of interaction limiting the participants' position. In addition to the user's camera path, we recorded the user response and time. The VGE is virtually present at the center of this shape (as depicted in Fig. 3). Moreover, other information on the setup still needs to be outlined: (1) the starting position of the camera was random

(equally distributed between top or bottom of the VGE); (2) the position was always the same for one full batch to avoid user confusion during the task related to the same VGE; (3) none of the 2D layouts were rotated and therefore they were always in the same orientation.

#### 4. Participants of the study

This experimentation was available online between June and September 2020. The study was fully open to the public, with the participants using their own laptop to complete the tasks. Participants were reached over social networks, forums for professionals, and word-of-mouth. The online platform hosting the experiment was from an in-house development, and accessible only from a computer (smartphones were excluded). A total of 107 participants have completed the study. Out of this number, 35.5% were female and 64.5% male. The average age was 37.7 years old ( $sd = 12.3$ ), fluctuating from 17 to 67 years old. The participants were accustomed to 3D with 81.3% using this medium on a monthly basis.

### 5. Results

#### 5.1 VGE interaction learning

The learning of new tools (or technologies) by users begins at its first use. This assimilation is manifested by an increase of efficiency in using the tool. The estimation of this learning curve is assessed by two parameters in this study: (1) the number of inputs, (2) the time to complete the task. The second parameter was introduced to inquire about the participants' fatigue, with the assumption that this aspect is translated by fewer inputs and shorter answer time. For this analysis, each batch is defined by the combined inputs of *guess layout* and *guess highest*. *Guess viewpoint* and *guess lowest* were excluded from the calculation because these tasks were not asked in every batch.

Fig. 4 shows a decrease in the number of inputs over time. This negative correlation was validated statistically by a Somers' Delta, non-parametric test for ordinal dependent variables ( $\delta = -.123$ ,  $p\text{-value} < .001$ ). The strength of the relationship between the two parameters is *weak*, but notable. For estimating this relation, the number of inputs was aggregated by slices of 10. The same operation was conducted on the time to complete a batch per participant, and no correlation was found.

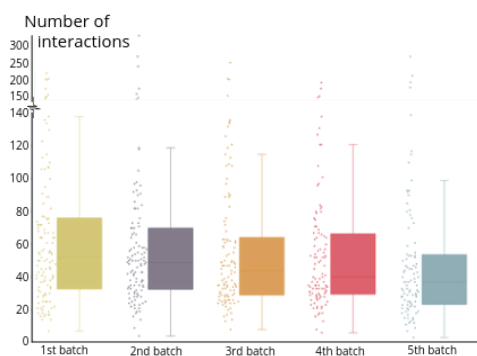


Figure 4. Number of inputs per batch. The batches have been classified by order of appearance for each participant.

#### 5.2 Participants' interaction uncertainties

This concept is estimated by the ratio of VGE interaction time divided by the total time needed to complete the task. Indeed, users experiencing hardship manipulating a VGE will pause for a significant amount of time between their interactions. This interruption may help the users to reassess their situation in the VGE, i.e. evaluate their new position, study the new orientation of the scene, consider their next input, etc. Fig. 5 depicts the average ratio of interaction time/total time according to the age and the frequency of 3D use for the task *guess layout* and *guess highest*.

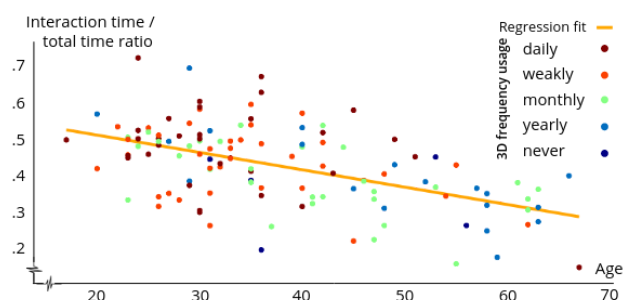


Figure 5. Interaction/total time ratio according to age; the color represents the discrete frequency of 3D use.

A few ratios (21) for the *LoD 2 + texture* representation were deleted from the analysis due to a computational error. For the remaining ratios, a compelling relation for the age has been identified with a Somers' Delta test ( $\delta = -.343$ ,  $p\text{-value} < .001$ ). A moderate opposite relation was demonstrated for the frequency of 3D use with  $\delta = .264$ ,  $p\text{-value} < .001$ .

#### 5.3 Anchor points

To analyze how the users anchor their mental images, two anchor points affecting participants' cognition are considered: (1) the participants' starting orientation that is defined as the (randomly distributed) yaw angle (i.e. horizontal position, between 0 and 360°) at which a user lands on the VGE; (2) the orientation of the 2D maps' layouts that is established as the (yaw) angle of the VGE at which the 2D maps and the 3D scene orientations are aligned (Fig. 6). These two anchor points have been selected because: (1) the first orientation in which a user sees a VGE is crucial – all next orientations that are generated by manipulating the 3D scene are indeed resulting from this first image; (2) users had to exploit specific 2D layouts' orientations that are non-rotatable, thus, users had to adapt their VGE images to match the 2D layouts' orientations. These aspects were analyzed through the camera path which a user adopted to complete a specific task. The two anchor points were characterized by a sphere section encapsulating their orientations (Fig. 6). This section was calculated with a 5% margin around the orientation angle, e.g. if the orientation angle was 100°, the bounds of the section were 82°-118° ( $100-360*0.05$ ;  $100+360*0.05$ ).

Fig. 6 displays an example of these anchor points for a specific user. This participant started at the center of the green surface, through interaction the VGE was shifted to



visualize the 3D scene from a bird's-eye view and then the scene was rotated to match the 2D layouts' orientation (represented by the red section).

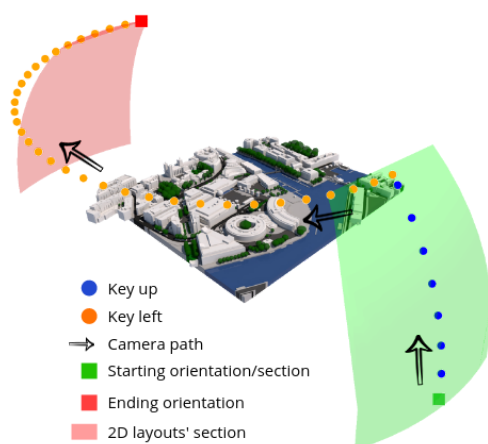


Figure 6. Camera path of a specific user for a specific task. The green area represents the landing position  $\pm 5\%$  and the red area represents the layout orientation  $\pm 5\%$ .

The importance of the anchor points was analyzed and shown in Fig. 7, which depicts the proportion of the users' camera positions that is enclosed in these anchor points' sections. This proportion is calculated for all the participants. For the task *guess layout*, these two sections represent 60.8% of all the participants' interactions (34.8 for starting orientation + 26.0 for 2D layouts' orientation). For the task *guess highest*, the users performed 48% of their interactions within these sections (31.5 starting orientation + 16.53 for 2D layouts' orientation). Therefore, these two sections channel a non-negligible proportion of the users' interactions, considering the fact that these sections represent from 10% to 20% of the spherical segment (i.e. all potential camera positions), depending on their overlay. These two sections were, thus, impactful orientations for the participants.

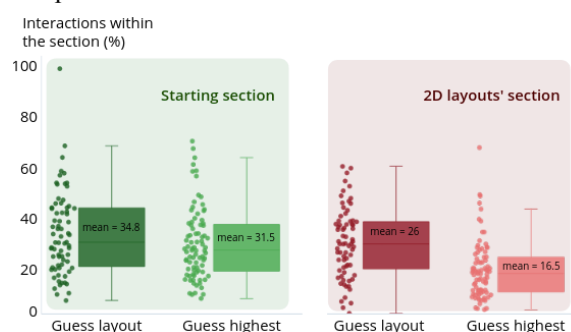


Figure 7. Average proportion of users' inputs enclosed in the starting and the 2D layouts' orientation area.

The relation between the starting and the 2D layouts' sections, as well as the camera path interpretation, are depicted in Fig. 8. For each user, all of the camera positions (for all pressed keys) were divided into five groups in chronological order (normalization of the number of inputs). The first group contains the first 20% of the positions, the second ranges from 20 to 40%, and so on. For example, if a user manipulated the camera 15

times (numbered by their chronological order), the distribution would be group 0-20%: 1, 2, 3; group 20-40%: 4, 5, 6; group 40-60%: 7, 8, 9; etc. An average distance was then estimated for each group between the averaged camera positions and the two sections. The distance is defined as the yaw angle difference between the averaged camera position and the central orientation of the section. Thus, this distance ranges from  $0^\circ$  to  $180^\circ$ . Fig. 8 depicts these distances for the task *guess layout* of all the five representations. A regression line (for the average user distance for each representation) has been calculated and drawn on the figure to highlight the global trend. The top graph on Fig. 8 depicts the distance compared to the 2D maps' layout orientation. We observe a diminution of the distance, i.e. participants tend to get closer to the 2D layouts' orientation processing towards the completion of their task. This relation was demonstrated statistically via a Somers' Delta estimation ( $\delta = -.289$ , p-value  $< .001$ ). This estimation indicates a *moderate-strong* negative correlation between the two ordinal values (groups: independent value; distance: dependent value, aggregated by  $10^\circ$  distance). The bottom graph on Fig. 8 illustrates an opposite relation. The distance from the origin increases over the users' inputs. A Somers' Delta value was calculated:  $\delta = .326$ , p-value  $< .001$ . This calculation shows a *strong-moderate* positive correlation. Therefore, a relation can be highlighted between the two anchor points (starting position and 2D layouts' orientations). Users begin their task by exploring the 3D scene around their starting orientation and complete their assignment when attaining an orientation aligned with the 2D layouts.

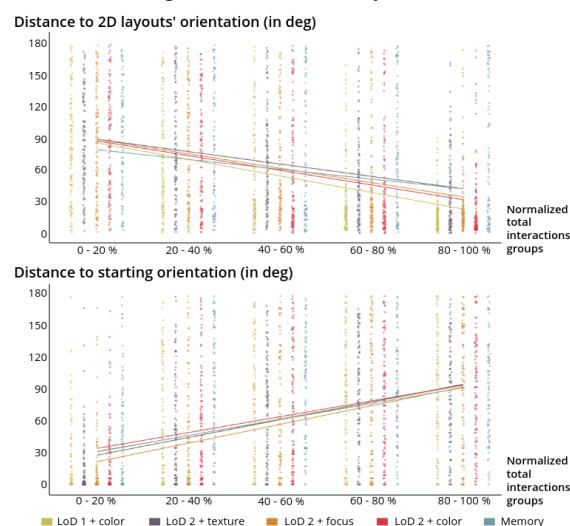


Figure 8. The evolution of distance from the users' camera positions to the 2D layouts' orientation and the starting orientation at five temporal snapshots.

#### 5.4 Channeling users behavior

The density of users' camera positions depicted in Fig. 9 provides an opportunity to evaluate how the nature of the task channels the users' behavior to a reduced set of actions. This figure presents the inputs of the participants for one full batch (*guess layout*, *guess highest*, and *guess viewpoint*) of a specific scene (*LoD 1 + color* in this

particular case).

Results show that the task *guess layout* attracts camera positions on the upper part, revealing that the participants looked at the VGE from a bird's-eye view (closer to the 2D layouts). Camera's positions for the task *guess highest* are more distributed, but the bottom part is denser, indicating a view closer to the horizontal plane (which helps to distinguish the heights of the buildings). Last, the task *guess viewport* attracts the users' interactions on the two lines between the red sphere and the two different green spheres (which is needed to assess their relative visibility). These three tasks demonstrate three heterogeneous distributions of the participants' camera positions.

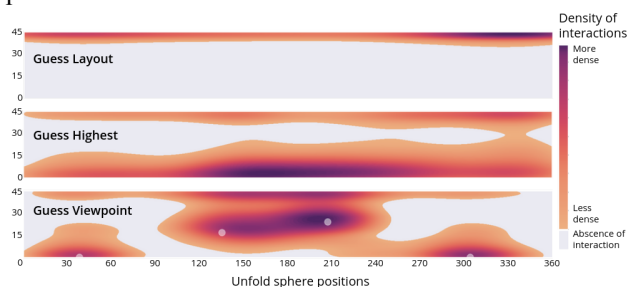


Figure 9. Density of users' camera positions on the unfolded spherical segment.

A second aspect has been considered for analyzing the impact of the task on users' behavior. Participants were interacting with the VGE via a keyboard; more specifically, with their arrow keys. When starting a new step of the study, the participants discovered a task and a VGE setup; from their understanding and starting position, they instinctively decided their first inputs. This input translates what their first intention is, which is crucial for analyzing the users' behavior.

Fig. 10 portrays this first intention. For the task *guess layout*, the principal input is *up* (48.2% of all first inputs), allowing visualizing the VGE from above. The second most pressed input is *down*, which represents 23.7%. For the second task *guess highest*, the number of inputs *down* becomes greater (44%) than the number of inputs *up* (36.8%).

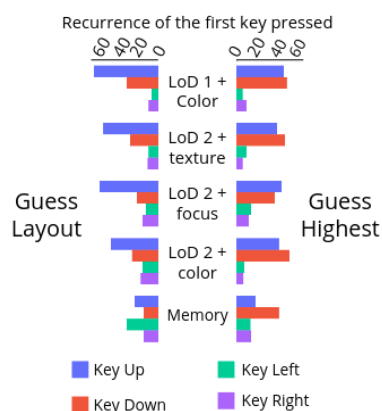


Figure 10. The number of first inputs (up, down, left, right) per 3D scene.

## 6. Discussions

### 6.1 Participants learn to interact with the VGE

Participants executed fewer interactions to complete specific tasks over the course of the study. However, the time needed to complete the task stayed consistent. Hence, the drop of the interaction number is linked to an improvement in VGE manipulation and not to fatigue, boredom, randomness, etc. Furthermore, no significant difference in the accuracy of the answers was established<sup>3</sup>.

Participants were discovering the layout of the various scenes through their assignments. The different scenes were off-context to avoid any recognition (that could lead to bias). Therefore, participants learnt to manipulate the VGE (with its rules, tasks, representations), in only a few interactions, notably by developing strategies to get more efficient (as soon as they start interacting with the VGE).

The learning process highlighted in this study is crucial for participatory approaches. Indeed, challenges faced during the first utilization of an unfamiliar VGE could be quickly overcome. Therefore, participants do not have to be trained in advance. A brief targeted tutorial explaining the VGE specificities could enhance its use by the participants, as well as improve their feedback.

### 6.2 Participant's idiosyncrasy contributes to interaction uncertainties

The ratio of the VGE's manipulation time in relation to the total time required to complete a task is strongly correlated to socio-demographic characteristics such as age or frequency of 3D use. These two characteristics are acknowledged as limiting users' performance with 3D (Stanney, Mourant, and Kennedy 1998). Therefore, the ratio could predict hardship in the usage of VGE.

A ratio tending toward zero indicates a lot of pausing in the VGE interaction, possibly due to an understanding reassessment of the 3D scene and its following positions. On the opposite, a ratio close to one depicts an intensive interaction introduced by a certain proficiency in 3D. Identifying struggle or ease in interacting with VGEs should, however, not be only limited by the evaluation of this ratio. Indeed, several other parameters (spatial cognition, computing performances, screen size, disturbing factors, etc.) could have affected the users' performances and weighed in this relation between the users and the VGE. Nevertheless, the interdependence between age and frequency of 3D use did not skew the results: (1) 11 participants reported using 3D tools once a year or less frequently (5 < 32 years old and 5 > 55 years old); (2) a monthly consumption of 3D tools (considered scarce in this study) cannot reflect unfamiliarity.

Identifying struggle in manipulating VGE is central to the design of such tools for urban participatory planning. The ratio introduced in this study suggests an accurate estimation of the users' skills and is easy to evaluate in

<sup>3</sup> The result on accuracy is under submission in a parallel study.

real time. Therefore, a customized assistance could be proposed dynamically to the participants with a low ratio (defined by a threshold), to whom the VGE could offer extensive guidance. In contrast, participants with a high ratio could be offered fewer indications.

### 6.3 Participants' cognitions are articulated around the initial position and the task to complete

The analysis of the anchor points highlighted two essential positions: users' initial viewpoints and 2D map layouts'. Participants remain in a section of about 10% around their initial position for about 30% of their overall positions around the scene. The same importance has also been shown for the 2D map layouts, yet the task *guess highest* reached up to about 15% of the positions.

When starting a task, the participants tend to remain in their initial position's area to frame an understanding of the scene, whereupon the next images of the scene (introduced by interaction) will be constructed. Once an overall understanding is crystallized, users explore the 3D scene to build the knowledge needed to perform the task they have to complete. In this study, the task is characterized by driving the users to align the VGE orientation with the 2D maps' layouts or by identifying the highest building. In the second task, the 2D maps' layout orientation affects the users positions less: (1) users had to identify the building that seems to be the highest, and therefore rotate the scene before locating the footprint of this building; (2) the task *guess highest* is always following the task *guess layout*, thus, the participants already have a mental image of the 2D layouts orientation and the VGE.

These two anchor positions are crucial for the design of participatory 3D tools. The initial position supports most of the users' understandings, therefore it should be selected thoughtfully, for instance, to emphasize an element within the VGE, or to have an overview of the system. The task to complete can deeply focus users' attention and either drive them to explore the environment (for example, by introducing a task on height comparison or by a sequence of tasks aiming for different localization) or to limit their displacement (via a task located at the same area of the initial position).

### 6.4 The assigned tasks channel participants' behavior

The tasks that users have to accomplish channel the resulting interaction. The results demonstrate that the task *guess layout* focuses the interaction towards the top of the VGE, i.e. a bird's-eye-view on the scene. In contrast, the task *guess highest* drives the participant to adopt a horizontal view that is useful to identify the height of a building compared to the others. Last, the task *guess viewpoint* highly spatializes the position of the participants around the lines connecting the red sphere to the green spheres.

These interaction behaviors demonstrate the priority given by the participant to visualize the scene in a certain way to perform a task. A higher point of view close to the 2D layout, supports a global representation of the VGE,

and also eases the translation between 2D and 3D representations. Other points of view encourage the distinction of heights or specific perspectives. However, the bird's-eye-view perspective appears to be central in the development of interaction behavior, where participants in addition to performing their current task anchor their VGE's mental image from this top view. Furthermore, users seem to be more comfortable with this top view, similar to a 2D map, with which they are more familiar.

Following the anchor points images for the design of VGE in urban participatory planning, the task affects the interaction behavior developed by the users. The findings of these users' camera positions support the fact that the task's configuration could drive users to broaden or centralize their cognitions within VGEs.

## 7. Conclusion and perspectives

In this paper, we have found evidence suggesting that a better knowledge of user strategic behavior in interacting with VGEs can be beneficial for improving the design of interactive 3D tools for participatory approaches. Users' behavior was tested in an online user study that simulates specific tasks for the practice of VGE in urban participatory planning. From this experimental study, recommendations for adopting VGE in these practices were established:

1 - Users improve their interaction efficiency as soon as they start their tasks. Therefore, the users have a strong appropriation of the VGE. As this learning process is immediate, a specific tutorial could highly improve the handling of the interactive VGE by the users.

2 - The manipulation time with the 3D scene compared to the inaction time (or uncertainties) is an accurate indicator for interaction ease with VGE. This computationally inexpensive indicator could automatically be estimated during the interaction to provide users with customized guidance.

3 - Users fix their attention to anchor points. These positions monopolize a significant amount of angles of views adopted by the users. Therefore, the cognitive load conveyed by the interaction with the VGE could be reduced by a mindful selection of the starting position within the system and a consistency with the task to complete.

4 - The required task drives the user's interactions. Thus, the angles of view that a user adopts while interacting with a VGE can be stimulated, offering an opportunity to passively guide users. However, in these circumstances, special attention should be addressed on ethics in the system's design.

The method developed in this study for evaluating strategic behaviors of the users demonstrates its relevance. Indeed, despite the heaviness of the effort asked from users, the dropout was limited to 47% of all attempts. Following this study, an experimental comparison between a digital and a physical setup, notably with interlocking blocks, is projected to assess the impact of the digital medium on users' behavior

(notably on elderly). Furthermore, the guidelines outlined in this study will be implemented in a parallel development of a 3D platform for urban participatory planning.

This platform aims to apply these findings in a real-world context, broadening the use of VGEs in urban participatory planning, in the hope of enhancing the practices and ultimately leading to better urban decisions.

## 8. Acknowledgment

This study was partly funded by the Computers & Geosciences Research Scholarships co-sponsored by Elsevier and the International Association for Mathematical Geosciences (IAMG). The in-house code used in this study is under MIT licence available on github: <https://github.com/thibaud-c/3DperceptionUX>.

The VGEs are published on Zenodo, doi: [10.5281/zenodo.5137307](https://doi.org/10.5281/zenodo.5137307).

## 9. References

- Al-Kodmany, Kheir. 2002. "Visualization Tools and Methods in Community Planning: From Freehand Sketches to Virtual Reality." *Journal of Planning Literature* 17 (2): 189–211.
- Biljecki, Filip, Hugo Ledoux, and Jantien Stoter. 2016. "An Improved LOD Specification for 3D Building Models." *Computers, Environment and Urban Systems* 59 (September): 25–37.
- Biljecki, Filip, Jantien Stoter, Hugo Ledoux, Sisi Zlatanova, and Arzu Çöltekin. 2015. "Applications of 3D City Models: State of the Art Review." *ISPRS International Journal of Geo-Information* 4 (4): 2842–89.
- Christophe, Sidonie. 2020. "Geovisualization: Multidimensional Exploration of the Territory." In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*.
- Çöltekin, Arzu, Susanne Bleisch, Gennady Andrienko, and Jason Dykes. 2017. "Persistent Challenges in Geovisualization – a Community Perspective." *International Journal of Cartography* 3 (sup1): 115–39.
- Couclelis, H., R. G. Golledge, N. Gale, and W. Tobler. 1987. "Exploring the Anchor-Point Hypothesis of Spatial Cognition." *Journal of Environmental Psychology* 7 (2): 99–122.
- Downes, Melanie, and Eckart Lange. 2015. "What You See Is Not Always What You Get: A Qualitative, Comparative Analysis of Ex Ante Visualizations with Ex Post Photography of Landscape and Architectural Projects." *Landscape and Urban Planning* 142 (October): 136–46.
- Dübel, Steve, and Heidrun Schumann. 2017. "Visualization of Features in 3D Terrain." *ISPRS International Journal of Geo-Information* 6 (11): 357.
- Herman, L., and Z. Stachoň. 2016. "Comparison of User Performance with Interactive and Static 3d Visualization – Pilot Study." *ISPRS - International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences XLI-B2* (June): 655–61.
- Herman, Lukáš, Vojtěch Juřík, Zdeněk Stachoň, Daniel Vrbík, Jan Russnák, and Tomáš Řezník. 2018. "Evaluation of User Performance in Interactive and Static 3D Maps." *ISPRS International Journal of Geo-Information* 7 (11): 415.
- Herman, Lukáš, Tomáš Řezník, Zdeněk Stachoň, and Jan Russnák. 2018. "The Design and Testing of 3DmoveR: An Experimental Tool for Usability Studies of Interactive 3D Maps." *Cartographic Perspectives*, (90): 31–63.
- Jacquino, Florence, and Julia Bonaccorsi. 2019. "Studying Social Uses of 3D Geovisualizations: Lessons Learned from Action-Research Projects in the Field of Flood Mitigation Planning." *ISPRS International Journal of Geo-Information* 8 (2): 84.
- Jankowski, J., and M. Hachet. 2015. "Advances in Interaction with 3D Environments." *Computer Graphics Forum: Journal of the European Association for Computer Graphics* 34 (1): 152–90.
- Joerin, Florent, Gilles Desthieux, Sandrine Billeau Beuze, and Aurore Nembrini. 2009. "Participatory Diagnosis in Urban Planning: Proposal for a Learning Process Based on Geographical Information." *Journal of Environmental Management* 90 (6): 2002–11.
- Land, Sarah van der, Alexander P. Schouten, Frans Feldberg, Bart van den Hooff, and Marleen Huysman. 2013. "Lost in Space? Cognitive Fit and Cognitive Load in 3D Virtual Environments." *Computers in Human Behavior* 29 (3): 1054–64.
- Lin, Hui, and Jianhua Gong. 2001. "Exploring Virtual Geographic Environments." *Annals of GIS* 7 (1): 1–7.
- Schroth, Olaf, Ellen Pond, Cam Campbell, Petr Cizek, Stephen Bohus, and Stephen R. J. Sheppard. 2011. "Tool or Toy? Virtual Globes in Landscape Planning." *Future Internet* 3 (4): 204–27.
- Sheppard, Stephen R. J., and Petr Cizek. 2009. "The Ethics of Google Earth: Crossing Thresholds from Spatial Data to Landscape Visualisation." *Journal of Environmental Management* 90 (6): 2102–17.
- Stanney, Kay M., Ronald R. Mourant, and Robert S. Kennedy. 1998. "Human Factors Issues in Virtual Environments: A Review of the Literature." *Presence* 7 (4): 327–51.
- Ugwitz, Pavel, Vojtěch Juřík, Lukáš Herman, Zdeněk Stachoň, Petr Kubiček, and Čeněk Šašinka. 2019. "Spatial Analysis of Navigation in Virtual Geographic Environments." *APPS. Applied Sciences* 9 (9): 1873.
- Wilkening, Jan, and Sara Irina Fabrikant. 2013. "How Users Interact with a 3D Geo-Browser under Time Pressure." *Cartography and Geographic Information Science*.
- Zhang, Sisi, and Antoni B. Moore. 2014. "The Usability of Online Geographic Virtual Reality for Urban Planning." *Lecture Notes in Geoinformation and Cartography*.