Pip and Pop 2.0

Andrea Keiser, Ann-Marie H. B. Bech, Frederik S. V. Bærentsen and Nicky M. Kjærgaard

Abstract—Searching for an object in a set of distractors has been researched thoroughly, however all the current articles use simple lines or simple three-dimensional graphics for both target and distractors. This paper will research how advanced 3D graphics created in Autodesk Maya will affect the visual search task in a pip and pop similar experiment.

Two different experiments were made; one where participants were asked to search for a target object in a visual search setup using three different types of graphics, accompanied by sound, and one where participants used visual search to look for a target object using advanced graphics, both with and without sound. The first experiment shows that using different types of graphics does not change the pip and pop effect when 3D modelled objects are used. The second experiment confirms the assumption that the sound has a positive effect on the search time when applying auditory beep to the advanced 3D graphics compared to the lack of the auditory beep.

Keywords—*pip and pop, visual search, attention*

I. INTRODUCTION

This paper contains information about visual search and attention with a focus on theory about the pip and pop effect together with further research made regarding this effect. The focus in this research paper was on graphics and how applying more advanced graphics to the original pip and pop experiment could change anything. Previous studies have proven that audio can guide a person into finding a target object whether the audio is spatially [2] or non spatially [5] supplied. These studies have given a focus on the importance combining the auditory senses with visual events when these are synchronized. Therefore, they discovered the improvement in performance resulting from it.

Two experiments have been conducted: One to see if the pip and pop effect when changing the graphics still apply and works, and if there is a big difference in the performance time between each of three different images of different graphics level used; 2D lines, 3D rectangles, and 3D objects with high details. The second experiment is focused only on the most advanced of the three graphics, testing if the pip and pop actually makes a difference or not when using advanced graphics.

The motivation for researching the pip and pop effect was regarding a focus on the graphics elements. To see whether or not using different graphics would change the functionalities of the pip and pop effect and to see if the auditory beep can be helpful even when working with advanced 3D graphics.

II. THEORY

A. Attention

This part of the theory section will include some basic information about attention before moving on to a more in depth research on how visual search works and how it depends on attention.

"Attention is the name we give to a family of mechanisms that restrict processing in various ways" [6, p. 187]. Attention can be given to all senses and not only visual, which will be the main focus in this paper. How visual attention works cannot be explained completely but there is a theory that it works like a spotlight. The idea is that visual selective attention will most of the time be directed on one thing and in one location. [6, p. 189-191].

Attention can be described using multiple different categories and terms; it can be internal, which concerns selection, more specifically how to select one response opposing another or how to focus on one line of thought and not the other. Attention can also be external, which is the attention given to the stimuli in the world. Two other terms could be overt and covert. Overt is when using a sense organ like the eye and directing that sense organ on a specific stimuli. The idea of covert attention was developed by Hermann von Helmholtz, a German psychologist and physicist, who demonstrated that even though the vision is focused on a specific spot, the visual attention can be directed in another area without moving the eyes [3, p. 117-118]. Covert attention can also refer to two studies, which define this type of attention as endogenous attention and exogenous attention. Endogenous refers to the information gathered from within, own knowledge while exogenous refers to the outside information gathered by the attentional process, for example the sound of a breaking glass or salient colors [3, p. 117-118]. It is also possible to have divided attention. This occurs when the attention is split between different tasks, for example reading while listening to music. Studies performed on divided attention show that performance is slowed down when trying to focus on two visual sources at once. Performing two tasks at once also hinders the accuracy or reaction time of the person, this phenomenon is called dual-task interference [3, p. 112]. Sustained attention is fixating on a specific thing tracking the development of that current event, like water boiling. The last type that will be mentioned is selective attention, which is the ability to choose a few or maybe only one out of multiple stimuli. [6, p. 187-188].

B. Visual search

Visual search is a process where the goal is to find a target surrounded by other distracting objects, these are called distractors. When experimenting with visual search the task is often to find a target mixed together with distractors on a display [4]. It is however not only when conducting experiments that visual search is experienced but also in the real world. Every time it is necessary to locate a specific object between others it is visual search, for example if trying to find a certain movie in a movie store or a queen of spades mixed

between the additional 51 cards or distractors. The difficulty within these real world searches can vary from situation to situation. Some might even be so easy that they are not really experienced as searches. Working with this difficulty within visual search has been one of the main considerations the last past quarter century in the field of cognitive psychology. When setting up an experiment one of the things to consider is the set size. The set size is the amount of total items shown in the display. A rule that might be logical itself is that the difficulty of the task, the task being finding the target, increases when the amount of items in the display increases. If the goal is to measure the efficiency, time can be used as a main factor in visual search experiments. [6, p. 191].

The term efficiency is a word that can be used to describe the process of visual search very well. If the target is found quickly the search can be considered efficient. If each item is looked at it will most likely take more time and therefore turn out to be a slow search, which will be considered an inefficient search. The efficiency can differ depending on the task requiring a visual search. [6, p. 192-194].

There are different searches such as:

- Feature search, when searching for a target that has been defined by one attribute, this could be the orientation of the target or the salient color.
- Parallel search, when several stimuli are being handled all at once within a search.
- Serial self-terminating search, working your way through the items, item to item and stopping only when the target has been found.
- Guided search, when directing attention to a limited group of potential items based on information used to define the target, this could be the color.
- Conjunction search, when two or more attributes are used to define the target of the search, this could be a green horizontal target surrounded by red vertical distractors.

[6, p. 192-194].

The binding problem refers to connecting different information handled by different brain circuits such as connecting a color to blue, motion to moving up, and orientation to horizontal, ending up with the result: a blue horizontal object moving upwards [6, p. 196]. When looking at Figure 2, it does not take long to see that we are looking at blue and orange rectangles, however when asked to find the target, being the only one completely horizontal, it is needed to direct the attention to the items individually. Pre-given information about what features to look after can help guide to what the attention should be focused on, for example to find a round blue circle, the attention should be focused on the color blue and items with the shape of a circle and not the color yellow or a triangle shape. [6, p. 197].

Illusory conjunction is worth mentioning when explaining the binding problem. Illusory conjunction occurs when combining things together wrong. For example looking at different shapes in different colors for a short amount of time, followed by being asked to mention them using both attributes, it is possible to combine them together incorrectly, see Figure 1. So instead of ending up with e.g. a purple triangle, which is an
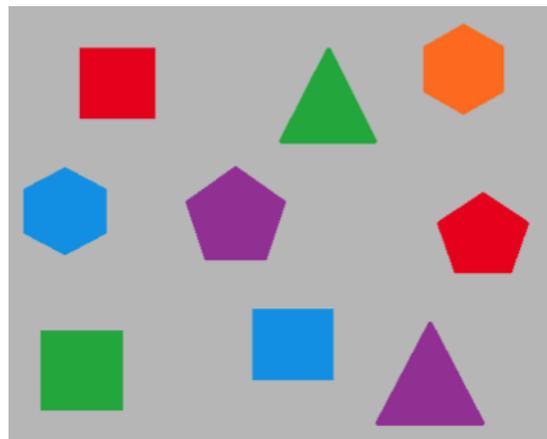

Fig. 1: Image used to demonstrate the binding problem and illusory conjunction.

actual item on the display, what is combined is a purple circle, even though there is none on the display. It should however be mentioned that it is not possible to combine features that are not represented on the display, Figure 1 demonstrates again it would not be possible to end up with a yellow circle, since there are no circles on the display or anything with the color yellow. [6, p. 197].

III. STATE OF THE ART

A. Pip and Pop experiment

The pip and pop experiment is a visual search experiment that focuses on nonspatial auditory signals, which is said to improve the search within the experiment. This project is based on the pip and pop experiment performed at Vrije University [5] and recreates the experiment performed within this research.

Before it is possible to understand the pip and pop experiment it is necessary to know the definition of a visual search task. As mentioned before, a visual search is based on a scan of the visual environment for a particular target object either with or without eye movements. In the case of the pip and pop experiment, the visual search is performed without eye movements and participants are not allowed to scan through the image one by one object but are asked to focus on a white fixation dot placed in the middle of the image.

The dot is surrounded by various amounts of objects, the used set sizes in Vrije University's experiment are 24, 36 and 48 continuous changing red or green lines. The researchers of the study claim that the reason these set sizes were chosen was to avoid immediate detection of the target object. These lines were placed in an invisible 10x10 grid around the fixation dot and their orientation was slightly rotated by plus or minus 22.5 degrees horizontal or vertical, excluding the target object, which was either horizontal or vertical. Another constraint was that the target would never be placed near the central positions of the image. The lines changed colors in cycles which consisted of 9 intervals each, each of these intervals

took place randomly every 50, 100 or 150 ms, however the target change was always followed by a 100 ms interval and only occurs after a 150 ms interval. The amount of objects that changed colors within the set depended on the set size, for example when the set size was 24 an amount of 1, 2 or 3 objects changed color and when the set size was 36; 1, 3 or 5 items changed.

Another characteristic of the pip and pop experiment is the continuous changing of the colors of a random number of objects within the display, however the target change was always alone and always accompanied by a tone, in the tone present condition.

The researchers focused on the reaction time of the participant from searching the display and finding the target object. When the participant found the target object, they were requested to either press the z or m key respectively to state whether its horizontal or vertical. Participants were asked to perform two practice tests, followed by four without sound and four with sound. For the results, practice setups results and incorrect results were excluded.

The researchers performed five experiments where they tested the tone and tone absence importance. In the first experiment, the sound acted as a cue to when to expect the target change, which always changed alone within the interval, the tone may have increased general alertness. Reaction Time (RT) were faster with tone, in the tone absent condition the search time was dependent on the set size, however in the case of the sound the set size was not found to be significant.

Experiment 2 uses a different approach on the visual search task, and the sound was replaced by visual cues to when the target will change color. One of these cues was to briefly remove the fixation dot, another was to narrow the participants window of attention while they focused on the fixation dot. However, it was proven that these cues were ineffective.

Experiment 3 went in more depth with the results of experiment 1, where the researchers focused on manipulating the intervals when the sound would take place and tested whether providing it before the target change, after or simultaneous to it, would be the most optimal. The experiment concluded that the simultaneous tone with target change proved to be the most optimal when it comes to performing the visual search task.

In experiment 4 the researchers tested the tone change coinciding with the color change of one of the distractors, respectively the tone change with the target, which was present 80% for one trial and 20% of the time for the other, whether this played an importance, considering that the short reaction time is due to a stimulus driven process. The results show that participants had an easier time finding the target when the auditory signal coincided with the target change.

Their last experiment consisted of removing the synchronization of the tone with any event within the visual search setup or with one of the distractor changes. The results show that reaction time was faster when the sound was present and synchronized rather than when it was not synchronized with no event at all.

B. Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli

Another study that takes a different approach on the pip and pop effect is a study conducted at Oxford University that focuses on spatial informative cues rather than non-spatial cues such as the pip and pop experiment performed at Vrije University. They also observe that spatial informative cues help improve performance in visual search tasks.

The Oxford University experiment also focuses on cues importance on a visual search task. The study performed four experiments each with different conditions, the first experiment being based on the pip and pop experiment, which tests the non-spatial auditory signals and its connection to a visual search task. It is to be noted that they presented the auditory cues differently from Vrije University, using loudspeakers rather than headphones. The results from the first experiment lead to the same conclusion, meaning that the visual search tasks presented with the auditory cue led to faster search latency.

During the second experiment, four different cue types were tested. Two of them being the same as experiment one, with an auditory cue and without, however, they also tested the vibrotactile influence placed on the waist and used bimodal audiotactile cues, meaning that the target change was accompanied by both sound and vibrations. The results also showed that no matter the type of the cue, the response time is significantly increased compared to the task without any type of cue. [2, p. 1668].

In the 3rd experiment, the researchers focused on spatial informative cues while testing both the sound and the vibrations influence on the search time. For example when the target object was placed in the left side of the setup, the participant would get auditory feedback in the left ear while the vibrotactile feedback was given by a tactor on the corresponding side. Experiment 3 focuses mainly on whether the combination between synchronous and spatially informative cues improve performance when completing visual search tasks in comparison to tasks that included temporally synchronous cues however spatially uninformative. The results prove that by adding both synchronous and spatially informative cues will lead to an improved search time and performance. [2, p. 1658]

The last experiment focuses on giving the cues through different medias, in this case they use headphones and compare them to loudspeakers. It was discovered that auditory cues via loudspeakers can improve performance in visual search rather when using headphones. [2, p. 1660].

C. Sensitivity to three-dimensional orientation in visual search

Visual search has been researched before, and early finding shows that two visual subsystems exist [1, p. 323]; the preattentive system and the attentive system. More recent findings argue that spatial information can be used at the preattentive stage, which goes against the early finding that the preattentive state cannot use spatial information. Enns and Rensink mention that scene-based properties like three-dimensional orientation are likely candidates for properties at the preattentive level.

Enns and Rensink conducted research into 3D orientation using different objects created by a computer and presented in a grid [1, p. 323]. They performed three experiments using different kinds of shading on very simple 3D objects to test the orientation. The first experiment is about testing different kinds of shading and seeing how fast participants can find them in a group of distractors with just one target. Experiment 2 refers to removing sides of the objects so the objects are not whole but parts are missing. Using these new objects, participants were again timed in their search for the target. The last experiment tests three-dimensional orientation as a preattentive feature, this is done using one to three different distractors for each target and using scene- or image-based orientation. Their findings suggest that the scene-based features are present at the preattentive level. Enns and Rensink have two unique findings: *"...rapid search can be based entirely on the spatial relations between linesshading is not required"* [1, p. 325] and *"...rapid search can be based on the three-dimensional orientation of simple blocks and not simply on the convexity/concavity of surfaces"* [1, p. 325].

Reserchers Enns and Rensink conclude in their experiment, that visual search that uses target synchronized non-spatial auditory signal, while the participant is performing the visual search task, decreases the search time required to find the target object.

IV. DESIGN AND PROCEDURE

This study contains two different experiments, this section will go in depth with the implementation and design of the first experiment, which uses three different graphical representations for three similar setups. These setups were made to follow a certain consistency, where each of the sets use a 36 set size, however each featuring different graphical representations. Figure 2, 3 and 4 show the images of each setup implemented for the test. In each of the setups, 35 of the objects were slightly deviated by either 22.5 degrees plus or minus horizontal or vertical and only one of the objects, which was the target object was placed horizontally (Figure 2) or vertically (Figure 4). All three setups were implemented and tested using the sound used in the original experiment [5], which occurred whenever the target object, horizontal or vertical changed color.

The design of all the experiments are based on the pip and pop effect experiment performed at Vrije University and each of the setups feature a white fixation dot in the middle of the image and follow the following guidelines:

- The tone will occur when the target object changes color
- The set size must be 36, with 35 distractors and 1 target object
- 1, 3, or 5 of the distractors will change color randomly every 50, 100, or 150ms
- The target object changes must occur every 900ms
- The target object must not be placed in the central position around the fixation dot
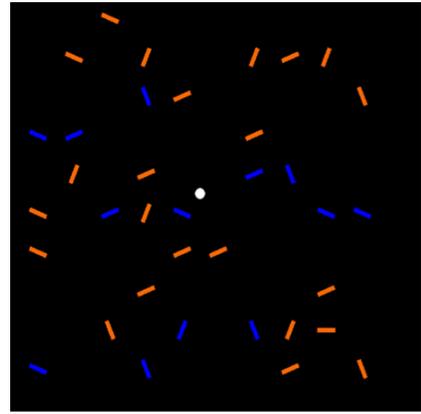- The target object must be vertical or horizontal

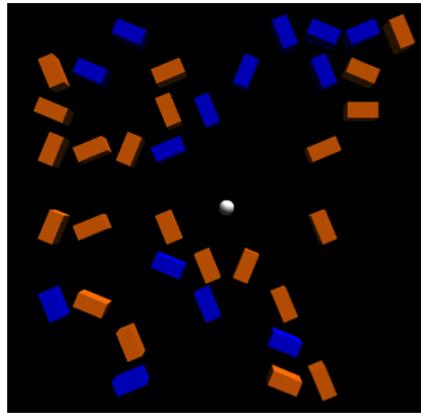Fig. 2: Recreation of the original pip and pop experiment.

Fig. 3: 3D rectangles pip and pop experiment.

A. Experiment 1 - Visual search based on graphical difference

As seen in the figures, there is a small difference in the graphical representations, Figure 2 was created based on the original experiment performing the pip and pop effect [5], therefore, only simple lines were used, which changed colors randomly. The target line changed colors every 900 ms and was the only one of the objects in the setup that changed the color during that interval. This setup was fully created in Unity engine and programmed to create lines within a 10x10 grid.

The second setup, Figure 3, is created very similarly to the original experiment, however the only difference was within the graphical representations, where simple lines were replaced by 3D rectangles. The objects are shown from a certain angle, creating shadows that indicate the objects are 3D. This setup also follows the 10x10 grid and was created in Unity Engine, following the same code as the former experiment.

The last setup (see Figure 4) follows the same idea as the former two setups, however in this case 36 random 3D objects were modelled in the program Autodesk Maya. The objects created have no connection to each other. The objects were imported in Unity and placed in a 10x10 grid, which was
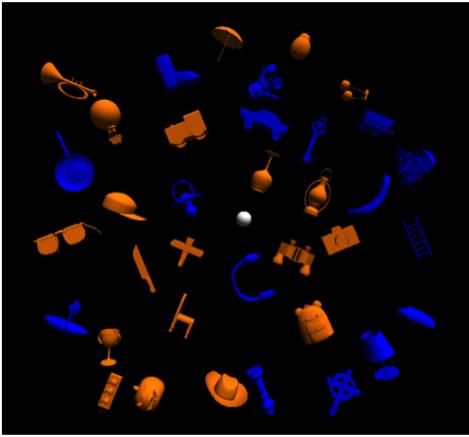
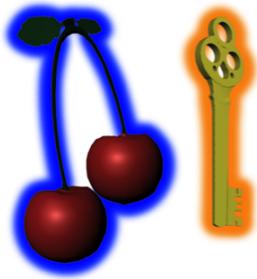Fig. 4: Advanced Graphics pip and pop experiment.

Fig. 5: Graphics using glow instead of full color.

used as a guideline, since the objects varied in size. More choices that were made during the design of this last setup will be described in the section below, *Experiment 2*, which is an experiment that thoroughly follows the experiment described in pip and pop experiment from Vrije University.

B. Experiment 2 - Pip and pop effect using improved graphical representations

For the second experiment the image with the advanced 3D objects was reused to test these kinds of graphics with and without applying the tone to the target change. The objects used for this experiment were randomly selected and were modelled in Autodesk Maya. Some examples of how the objects look before being added to the setup can be seen in Figure 6 to the right. It was decided to only use two colors for this setup and not assign each object their original colors such as a pumpkin being orange, or a cherry having two colors red and green, as seen in Figure 6 to the left.

It was thought to use a colored glow for the items or to make them change colors, such as it happens in the original experiment. An example of how the glow would be used can be seen in Figure 5. However, after further debate regarding the similarity to the original experiment and by trying to make sure the design of the current experiment was as similar as the
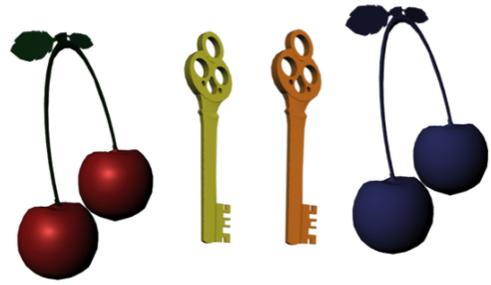

Fig. 6: Example of objects original colors to the left and fully colored objects to the right.

original pip and pop experiment as possible, the objects were colored completely and change colors at the same rate as the original experiment. Therefore the objects used in the final experiment can be seen in Figure 6 to the right, where both objects have one color each, blue and orange.

C. Implementation

This section will explain some of the implementation. For the full code with comments see the CD. The final programs can also be found on the attached CD under Programs.

The implementation was made in the Unity engine, with the imported 3D models made in Autodesk Maya. Each of the three images were made from the same code, but use different objects. The objects were placed in a 10x10 grid, with the white dot placed in the middle of it. This was done in the Unity Editor and not in the code (see Figure 7).
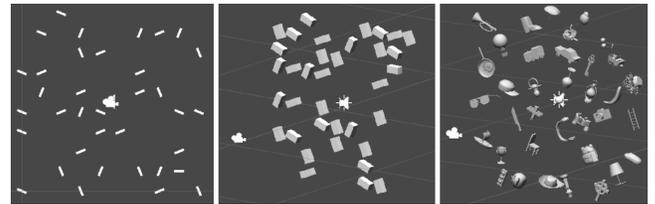
Fig. 7: The three images as they look in the Unity Editor.

To make the objects snap to the grid, a small script was added. This code basically checked the object's position on each axis and rounds the position value to the closest integer (see Code 1).

```
1  x = Mathf.Round(transform.position.x /
       ↪ cell_size) * cell_size;
2  y = Mathf.Round(transform.position.y /
       ↪ cell_size) * cell_size;
3  z = Mathf.Round(transform.position.z /
       ↪ cell_size) * cell_size;
4  transform.position = new Vector3(x, y, z);
```

Code 1: Snippet of code for making the objects snap to the grid in Unity.

Since the objects are placed in the Unity Editor they were visible as soon as the program was opened. To avoid this, the objects were all set to not being rendered as the program starts.

The color changing was done using two timers. The target object simply changed once every cycle, 900 ms (Code 2), while the distractor objects changed at different intervals within each cycle. As the target object should change color alone, the if statement `if(cycle > 0.2 && cycle < 0.8)` was added so the distractors only changed if there were more than 200 ms left of the cycle, but also less than 800 ms left (see Code 3 Line 1). That way there were 300 ms where nothing but the target was changing color.

```
1  if ( cycle < 0) {
2      if ( targetObject . transform . renderer . material .
           ↪ color == Color . blue ) {
3          targetObject . transform . renderer . material .
               ↪ color = new Color (1.0 f , 0.5 f , 0.0 f
               ↪ );
4      } else {
5          targetObject . transform . renderer . material .
               ↪ color = Color . blue ;
6      }
7      audio . Play ();
8      cycle = cycleTime ;
9  }
```
Code 2: Snippet of code for changing the color of the target object.

The distractors could change color at three different intervals: 50 ms, 100 ms, and 150 ms. At each interval the program also chose how many distractors should change color this time. This can be either 1, 3, or 5 distractors at a time. Line 3 in Code 3 shows code of this being chosen. In the `for-loop` it was chosen which distractors were to change color and then placed in a temporary array. If the same distractor was chosen a second time, a new distractor was chosen at random again, until 1, 3, or 5 different distractors are chosen, depending on the amount to change were chosen.

```
1   if ( cycle > 0.2 && cycle < 0.8) {
2       if ( timeLeft < 0) {
3           for ( int k = 0; k < numbersToChange [ Random .
               ↪ Range (0 ,3) ]; k++) {
4               tmp[ k ] = Random . Range (0 , distractorArr .
                   ↪ Length );
5               for ( int j = 1; j <= k; j++) {
6                   if (k > 0) {
7                       if (tmp[ k ] == tmp[k−j ]) {
8                           tmp[ k ] = Random . Range (0 ,
                               ↪ distractorArr . Length );
9                           j = 0;
10                      }
11                  }
12                  if ( distractorArr [tmp[ k ]]. transform .
                       ↪ renderer . material . color ==
                       ↪ Color . blue ) {
13                      distractorArr [tmp[ k ]]. transform .
                           ↪ renderer . material . color =
                           ↪ new Color (1.0 f , 0.5 f , 0.0 f );
14                  } else {
15                      distractorArr [tmp[ k ]]. transform .
                           ↪ renderer . material . color =
                           ↪ Color . blue ;
```

```
16                  }
17              }
18          }
19      }
20  }
```
Code 3: Snippet of code for how the distractors are chosen to change color.

In order to save the participants' answers and answer time, a text file was saved in the end. This file was saved under the name: 'Participant' plus their participant number (see Code 4), e.g. `Participant1`. The file contained text based on the variable `answer` and the amount of time it took them to answer.

```
1  System . IO . File . WriteAllText ( path +
       ↪ participantNumber + ". txt ", answer + "
       ↪ ," + timeTaken . ToString ());
```
Code 4: Snippet of code for saving a file with participant answers.

The variable `answer` was based on which key the participant pressed (see Code 5). Pressing `d` for horizontal would save `answer` as either `h,c` or `h,w`, depending on whether it was correct or wrong. Pressing `l` for vertical basically does the same, but with a `v` instead of `h`. If the participant pressed `q` for quit, `answer` would be saved as `f,f` for forfeit. An example of a saved file could therefore contain: `v,c,12.53`, if the participant pressed `l` after 12.53 seconds and it was the correct answer.

```
1   if ( Input . GetKeyDown ("d")) { //——— Horizontal
2       if ( targetObject . transform . eulerAngles . x >=
           ↪ 45) {
3           answer = "h,c";
4           EndGame ();
5       } else {
6           answer = "h,w";
7           EndGame ();
8       }
9   }
10  if ( Input . GetKeyDown ("l")) { //——— Vertical
11      if ( targetObject . transform . eulerAngles . x <
           ↪ 45) {
12          answer = "v,c";
13          EndGame ();
14      } else {
15          answer = "v,w";
16          EndGame ();
17      }
18  }
19  if ( Input . GetKeyDown ("q")) { //——— Quit
20      answer = "f,f";
21      EndGame ();
22  }
```
Code 5: Snippet of code for the user inputs.

V. EXPERIMENT PROCEDURE

Before any of the experiments were conducted, a pilot test was conducted on 5 participants to make sure that most issues

would be detected and fixed before the actual tests. Two issues were found, one was regarding the changes needed to be made on the written guide and the second was regarding the who participants forgot which button was refering to which orientation, this was fixed by placing two signs on the buttons, one with a horizontal line and one with a vertical one.

The first experiment explained in IV Design and Procedure focused on three different displays where the difference lay in the advancements of the graphics. The test was conducted on 30 participants, each testing all three displays with the pip and pop effect applied. Observation together with small talk after the test showed that in some cases, some had misunderstood the experiment and thought that they had to find a line in the image created by the objects. This was fixed by giving an verbal explanation to the participants before the test would start.

The second experiments point was to test the advanced 3D graphics with and without applying the auditory beep. For this experiment, 30 participants were tested on, however it was discovered after the test, that the difficulty between the two images was too different from another and it was decided that this was a big bias. The solution for this was to redo the experiment, now using between group experiments with 15 participants in each group so 30 in total. The first 15 participants were reused from the previous test, the original, to make sure it was the same image used for the first 15 participants which was now used for the new 15 participants. The final test included 15 participants testing the pip and pop with the auditory beep applied and 15 participants without, all testing on the same image to remove the bias of difference in difficulty.

In both experiments regarding the advanced 3D graphics it was for some participants difficult to locate or find the target object because of the orientation of the 3D objects making it hard to see if they were horizontal, vertical, or not. Another thing the two experiments had in common, was that the test could be conducted in various locations since only a computer and a pair of headphones was needed. This meant that there might have been people talking in the background or walking around, which could be a distraction, however, since headphones were used it was decided that it made no difference for the test whether or not the environmental conditions changed. It also made it a lot more convenient to gain participants since the test could be set up in their location. The participants were told to focus on the white dot in the middle while searching, however there was no way to be sure if they actually followed this requirement.

In total 75 participants were tested on, but due to the change in experiment 2 only 60 of the results were used for the result analysis.

VI. METHODS

The experiment conducted consists of two different group experiments; a between group experiment and a within group experiment. The within group experiment consists of three setups, which were randomized between each participant. The first setup was represented by an image consisting of lines, the second setup by an image consisting of 3D rectangles and the third setup by an image consisting of 3D objects. The second experiment, the between group experiment, consisted of each participant getting the same image, but it was randomly chosen whether the image was accompanied by sound or not.

Each experiment was afterwards evaluated using qualitative and quantitative methods. After each experiment the participants' times and answers were logged in a .txt file for later analysis and each participant was asked to fill out a short questionnaire. The questionnaire asked basic demographic questions together with questions about their previous experience with these type of experiments.

For the first experiment, the time it took for each participant to find the object was analyzed using a one way ANOVA test to determine if the mean of the three samples differed. The null hypothesis (H_0) states that there will be no significant difference between the three means, whereas the alternative hypothesis (H_1) states that there will be a significant difference between the three means of the groups.

For the second experiment a t-test was used to compare the two datasets of time recorded from the participants. The t test looked at the difference between the datasets for both correct answers, wrong and correct answers and all the combined answers (wrong, correct and failed to respond).

VII. RESULTS AND DISCUSSION

A. Within Group

The within group experiment has three conditions; Advanced graphics, Simple Graphics and Original Graphics. Figure 7 shows an example of these conditions.

For the within group experiment 30 participants were chosen randomly from various studies at Aalborg University Copenhagen. Out of the 30 participants 11 were female and 19 were male, with a median age of 23.

Each participant was given a random order in which they performed three different tests and the time it took for each participant was recorded, an overview can be seen in Table I.

TABLE I: Table showing the results.

	Mean time	Median time	Fastest time	Slowest time
Advanced Graphics	39.30s	29.30s	4.47s	181.12s
Simple Graphics	28.70s	17.27s	4.87s	70.73s
Original Graphics	28.02s	17.93s	2.88s	182.99s

The table shows that Advanced Graphics had a longer mean and median time, but the fastest time and slowest seemed to be random.

Before using an ANOVA test on the results, two hypotheses were defined:

- H_0: There is no significant difference in performance time of datasets when adding different graphical representations in a visual search task.
- H_1: There is a significant difference in performance time of datasets when adding different graphical representations in a visual search task.
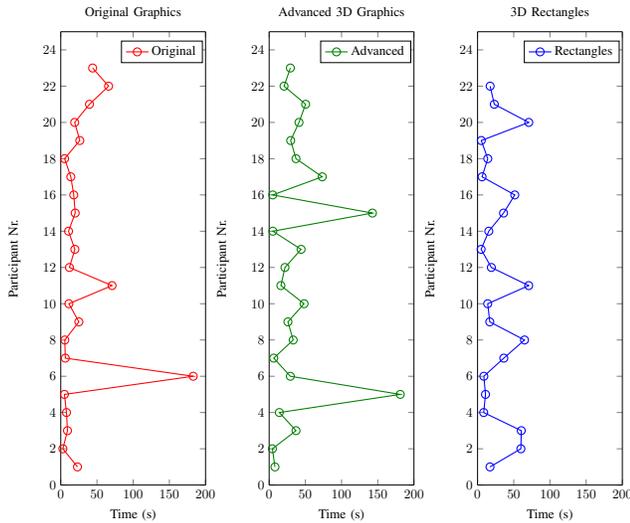
Fig. 8: Representation of participants response time.

The first t-test was done on the datasets of Advanced Graphics and Original Graphics, where only the participants that found the objects were considered. The hypothesis were the same and the result showed that it was not possible to reject the null hypothesis, so there is no significant difference between the datasets. For the second t-test the datasets were the Original Graphics and the Simple Graphics, the results showed that the null hypothesis cannot be rejected.

These tests show that there is no significant difference between the datasets.

B. Between Group

The second experiment was conducted as a between group experiment where two groups of 15 participants tried to find the target in an image only consisting of Advanced Graphics; one group had sound, the other did not.

TABLE III: Overview of the sound/no sound results.

	Mean time	Median time	Fastest time	Slowest time
Sound	43.36s	29.77s	12.73s	125.94s
No Sound	77.65s	64.99s	8.68s	303.05s

30 participants were randomly chosen from Aalborg University Copenhagen and 7 were female and 23 were male. They were randomly given a condition (with or without sound) by selecting a slip of paper from a hat.

Each participant had to complete the task once and when they found the target object, their time and answer were recorded (see Table III) and they were presented with a short questionnaire.

The results showed that the mean and median time for the group that had sound was clearly lower, however, to see if there was a significant difference between the two, a t-test was conducted. The hypotheses were the following:

- H_0: There is no significant difference in the completion time.
- H_1: Adding sound will be significantly faster than not having the pip sound.

In both groups seven participants found the correct target, while four thought they found the correct target and four gave up. Figure 9 shows two graphs of the time it took to find the target, for each particiant that found the correct target. The graph

Fig. 9: Graph showing the responce time for each participant that found the target.

for the 'Sound' (—⊖—) shows a much more stable line and the mean is just 23.76 whereas the mean for the 'No Sound' (—⊖—) is 78.78.
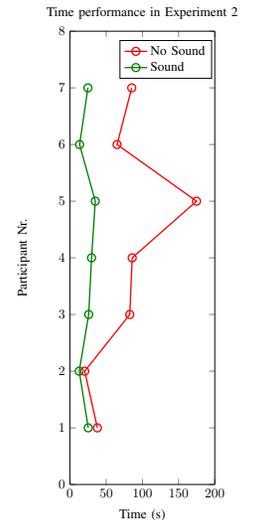
Using an ANOVA test (made in MATLAB using `anova1(x)`) on the three datasets it was possible to see that there was no significant difference between the datasets. This wass because the $F = 0.7383$ and to reject the null hypothesis F has to be more than 3.06 ($F > 3.06$). Since this was not the case it was not possible to reject the null hypothesis.

This means that there was no significant difference when using either of the three types of graphics while using the pip and pop effect.

This ANOVA test was made on all the answers and not just the ones that found the target. Table II shows a representation of how many participants found the object for each of the three setups.

In Figure 8 it's possible to see the difference in performance time between the participants that had correct answers in Experiment 1. The figure compares the time each participant required in order to find the target object amongst the distractors, which is shown on the x axis while the participant number is shown on the y axis. Since it was made sure that the order of the first experiments test was random for each participant, meaning there were 10 participants starting with advanced graphics (—⊖—), 10 with the original (—⊖—) and 10 with 3D rectangles setup (—⊖—), the bias that some participants performed better because it was their third trial, can be excluded.

TABLE II: Table showing distribution of Correct, Wrong and Forfeit.

	Correct	Wrong	Forfeit
Advanced Graphics	23	2	5
Simple Graphics	22	5	3
Original Graphics	23	2	5

It was not possible to make an ANOVA test on a dataset with different amount of data, so instead two t-tests were conducted.

This gives a small dataset to work with and by running a `ttest2(x,y,'tail','left')` on the dataset shows that it is possible to reject the null hypothesis (H_0) and accept the alternative hypothesis (H_1). This means that there is a significant difference between using sound or not. It also shows that using sound would have a significant lower mean time than not using sound.

C. Questionnaire

Based on the questionnaires it was discovered that 12 participants, out of the 30 that tested the first experiment, had tried the visual search tasks before and 18 had not. Based on their response in the questionnaire, on whether they had found the line during the experiment, it is possible to say that it did not have any influence on their performance. This conclusion was based on the question *which horizontal/vertical objects did you find?* Where 21 of the participants believe they had found the target line in the original simple line graphics, 19 believe they found the Rectangle in the 3D graphics and 16 believe they found the target object in the advanced 3D display. What can be concluded from these results is that previous knowledge of the visual search experiments did not help participants with the task, since, by looking at the participants who had tried such experiments before did not perform better when it came to finding the target.

TABLE IV: Categorization of the three setups according to difficulty.

	Easiest	Slightly more difficult	Most difficult
Lines	13	10	6
Rectangles	11	10	6
Objects	6	10	18

The question asking which target the participants found most difficult or easiest to find, had various results which can be seen in Table IV. Few believed that finding the horizontal object amongst the advanced graphics was easy, therefore 18 had the most troubles with this type of setup. The results in Table IV could be considered a bias since some participants gave the same answer to all three categories, resulting in having the same setup being considered as easiest, slightly more difficult and the most difficult. Therefore, not every setup is represented 30 times as the amount of participants in the experiment.

After being asked about which was the most difficult of the images, the participants were asked to explain why they thought their choice was the most difficult. The participants that chose the same image as the most difficult more or less answered the same. One of the participants who chose lines as the most difficult wrote: *"there were a lot of lines and they were slightly smaller"*, while another wrote: *"the object was smaller, so couldn't find it as fast as the others"*. Both agree that the smaller lines makes it difficult to find the target, even when the grid is the same for all three images.

A participant who chose rectangles wrote: *"the similaritie and size of the rectangles"*, as the reason for why it was the most difficult. Another participant agreed by writing: *"I am not

sure, but maybe because the rectangles were bigger, making it more difficult to somehow get an overview of all of them, and hereby difficult searching for the vertical/horizontal object"*. Both mention the bigger size of the rectangles as the reason.

The participants that chose the advanced graphics, mentioned in the comments as objects, mostly agreed the shapes were the reason. One wrote: *"more details had to identify the objects and the way they usually are"* and another wrote: *"objects were different shapes and sizes, also difficult to tell correct position of the object"*. The fact that the objects had different shapes could mean that the participants need to identify the objects before they can figure out the objects orientation. Another participant wrote: *"Objects orientation depends on shape. Differences"*. The shape of the objects influences the orientation of the objects, therefore after the object is identified, the normal orientation of the object, as it would be in the real world, needs to be identified as well.

In the ending of the questionnaire, participants were asked if they had any other comments about the test. Different comments regarding their confusion were mentioned by multiple participants. One comment which was mentioned was that it was difficult to keep focus on the fixation dot. The participant wrote: *"somewhat difficult to keep focused on the white dot, eyes sometimes wander around scanning objects"*. Another participant blamed the color changing for his/her loss of focus: *"I hated the constant color changes. I made me lose my focus"*.

Another comment was regarding the sound, which they believed it was annoying. The participant wrote: *"The sound sort of confused me, but maybe i did not understand that from the intro"*. The participants were not told what the meaning of the sound was and were just informed that there will be and auditory beep. Two other participants mentioned that the sound was stressful: *"the sound is a bit stressin.. but maybe that was the point?"* and *"The sound kind of stressed me out. It was very hard to find the objects. The sound did not help me, but felt annoying"*. The participants did not figure out the meaning of the sound, however one of them believed it was supposed to help them, hence mentioning *"the sound did not help me"*.

The last comment which was mentioned multiple times by various participants, was the fact that it was difficult to distinguish the objects' orientation in the advanced graphics image. One participant simply wrote: *"The objects' orientations were too confusing to distinguish"*, while another wrote: *"I was very confused on what I was looking for - I was both looking for the colour changes, shapes, but none of them REALLY felt perfectly horizontal or vertical"*. These comments relate to the shapes of the objects in the advanced graphics, such as when is a banana actually horizontal or vertical.

VIII. DISCUSSION

During the experiment performed, some biases were encountered and these will be discussed in this section and some solutions will be suggested.

Since it was important to keep the focus on the white dot, it would have been a better approach to implement eye tracking to make sure that participants were not scanning the image

for the target object. This way it would have been possible to disregard the results which show that participants did not keep their focus on the white dot, however it would not have made sure that they do in fact keep attention on the fixation point. Therefore, by adding this feature it would not have improved the results completely, however it would have been possible to exclude the participants who were unable to keep their focus on the fixation dot.

Another issue encountered was regarding the participants attention during the experiment. It was observed that numerous participants interrupted their focus while performing the experiment, in order to ask questions and give comments regarding it. Even though participants were asked whether they have any questions regarding the test before they started, they might have not been able to form any questions before they actually encountered the task they were put with. This might have affected the time they spent on it which should be remembered when going through the results. A way to avoid this would have been to use the same approach which is used in the original pip and pop experiment from Vrije University, which was to have trial runs before performing the actual test.

The environment tested in was not prearranged, and therefore numerous participants tested in different locations which were set throughout the campus in either busy or empty rooms. Even though participants were all wearing headphones and it was made sure that the screen was placed next to a wall, reducing the possibilities of being distracted. By borrowing a room and setting the experiment up in a fixed location would have made sure that each participant performs the test in a quiet and still environment, however this would have limited the opportunities to gather as many participants.

When detecting the target, the participants should have pressed the buttons D or L depending on the orientation of the target. The issue with doing it this way was that participants had problems pressing the correct button for the given orientation. It was observed in the pilot test that people had troubles remembering what button was for which orientation, therefore some small signs were made and placed on the two buttons showing which was for horizontal and which was for the vertical orientation. However, even this solution had its problems since the participants had their fingers above the buttons during the test and therefore might have not really seen the signs. A solution could have been to ask them to remember that horizontal was the left button and vertical the right button, but even that could have caused some confusion. The fact that the participants had troubles with this could have resulted in them using reaction time on figuring out which button to press or in the end press the wrong button without even realizing it.

For the advanced graphics it was realized that the participants had troubles with the orientation recognition of the objects, a solution for this could have been to add a question in the questionnaire asking the participants to state what object they had found as horizontal or vertical. To see whether the participant had selected an object as vertical but it was not the target object then it would have been possible to exclude that result as correct.

Even though the images used for the first experiment were randomized it could have been possible that some of them were biased due to the fact that the target might be placed in slightly more obvious position, being surrounded by less distractors. This could affect the time which the participants take to find the target object. In general a problem that occurred was participants pointing out that the target object was not completely vertical in the advanced graphics setup. However, the reason to why participants might have seen the target object as not completely vertical was due to the perspective of the grid used for the advanced graphics image.

Overall the results pointed to a certain conclusion for both experiments, therefore it was possible to decide whether to reject or accept the null hypothesis. The first experiment showed that there is no difference between using advanced graphics, simple lines or 3D rectangles in a visual search experiment where a tone is synchronized with the target color change. By performing an ANOVA test on the overall results, correct, wrong and the ones forfeited, it was not possible to reject the null hypothesis stated in the sections before. Even when performing two t-tests, one for the data sets from the original graphics and the simple 3D graphics and one between original graphics and the advanced 3D graphics, it was still not possible to report any significant difference between the three experiments. The amount of results with correct responses from all three experiments also sustains the fact there is no significant difference in performance time in a visual search task where the pip sound is involved.

The second experiment performed was completely based on the original experiment where only the advanced graphics were tested and the connection to the perceptual effect of the pip and pop experiment. For this experiment T-tests were performed on all the results, correct, wrong and the ones skipped however this was not ideal since the wrong and forfeited answers cannot contribute to the rejection or approval of the null-hypothesis, since they cannot provide us with any useful data in time performance. Therefore it was decided to focus only on the correct answers and perform a t-test only using this data, which lead to the null hypothesis being rejected and alternative hypothesis being accepted therefore adding sound will be significantly faster than not having the pip sound.

Since the second experiment was changed after it was realized that the image for the setup without the sound was biased due to the grid size and placement of the target object, the amount of participants used was rather low for both groups. Only 15 participants were used for each of the experiment groups in the second experiment, however very few participants had correct answers, only 7 participants for both groups were useful results. Ideally it would have been best to have 30 participants for each group, which could create a clearer approach to the conclusion.

IX. CONCLUSION

This study focuses on the difference between performance time when using different graphical representations such as simple lines, 3D rectangles and advanced 3D graphics in a visual search task. It was based on the pip and pop experiment conducted by Vrije University which tests auditory signals and their connection to visual search experiments. The experiment

performed in this study uses a slight different approach of the graphical representations that build the sets used for the tasks. The experiments conducted also included an imitation of the original experiment from Vrije University which was compared to the two different approaches using 3D rectangles and advanced 3D graphics, such as 3D modelled familiar objects. From the results it was discovered that 3D graphical representations do not improve or decrease search time in a visual search task accompanied by a target specific tone, therefore it was not possible to reject the null hypothesis stating that there is no difference between the three setups. The second experiments results led to the assumption that a visual search task using advanced graphics with a target accompanied tone compared to a visual search task without the tone will make the target object more salient during the task. Overall, it is not possible to conclude based on the results but merely make assumptions of the conclusions due to the low participant numbers that took place in the experiments performed.

REFERENCES

[1] Enns J. T. and Rensink R. A., *Sensitivity to Three-Dimensional Orientation in Visual Search*. Psychological Science, September 1990, vol. 1, no. 5, page 323-326.

[2] Ngo M. K. and Spence C., *Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli*. Attention, Perception, & Psychophysics, August 2010, vol. 72, no. 6, page 1654-1665.

[3] Smith E. E. and Kosslyn S. M., *Cognitive Psychology: Mind & Brain*. Pearson, 2007, Chapter 3, page 103-120.

[4] Treisman A. M.and Gelade G., *A feature-integration theory of attention*. Cognitive Psychology, January 1980, vol. 12, no. 1, page 97-136.

[5] Van der Burg E. and Olivers C. N. L. and Bronkhorst A .W., *Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search*. Journal of Experimental Psychology: Human Perception and Performance, October 2008, vol. 34, no. 5, page 1053-1065.

[6] Wolfe J. M. and Kluender K. R. and Levi D. M., *Sensation & Perception*, 3rd ed. Sunderland, Massachusetts U.S.A., 2009.