Improving HSL Recognition Skills with a Color Game

ALEXANDRE ANDRIEUX

Förbättra färgigenkänning i HSL genom ett färgspel

Förbättring av färgmatchningsförmåga kräver ett väl utformat gränssnitt såväl som en väl beräknad återkoppling av den underliggande färgmodellen. Detta exjobb avser produktionen av ett spel för inlärning av HSL och utvecklar kunskapen kring inlärningsmönster med avseende på kulör och progression i spel. En kvantifiering av lärande baserad på variation i prestation föreslås. I ett försök att balansera datamängd och relevans, definieras flera variabler av rent analysintresse. Resultat kring inhomogeniteter i inlärning presenteras genom s.k. Hue learning curves med streamgrafer och detaljerade stapeldiagram. Konsekvenserna av designval och spelande på prestanda och lärande diskuteras.

Improving HSL Recognition Skills with a Color Game

Alexandre Andrieux

KTH Royal Institute of Technology, School of Computer Science and Communication
andrieux@kth.se

ABSTRACT

Improving color matching skills requires a specifically designed interface as much as well-calculated feedback on the underlying color model. This thesis relates to the production of a game for learning HSL and elaborates on learning patterns with regard to hue and game progression. A quantification of learning based on performance variation is proposed. In an effort to balance data quantity and relevance, several variables of pure analysis interest are defined. Results on learning inhomogeneities are presented through Hue Learning Curves with streamgraphs and detailed bar charts. The consequences of design choices and gameplay on performance and learning are discussed.

Keywords
HSL, Hue, E-Learning, Color Distance, Streamgraph, Hue Learning Curve, Color Matching

1. INTRODUCTION

The HSL color model stands for Hue, Saturation and Lightness. Introduced in the 1970s, its usage in industry is presently worldwide as color selection has become a routine task. The World Wide Web Consortium incorporated the HSL and HSLA palettes in W3C Recommendation CSS Color Module Level 3 in 2011, stating that the advantage of HSL over RGB is its intuitiveness since one can guess at the wanted colors, and then tweak. "It is also easier to create sets of matching colors by keeping the hue the same and varying the lightness/darkness, and saturation" [1]. Since then, the HSL color space has indeed empowered web designers to create more coherent corporate identities and style guides. In 2016 HSL is supported by all modern desktop and mobile browsers [2].

1.1 The HSL color space
1.1.1 Description
Smith [3], Joblove and Greenberg [4] introduced the HSL color space in 1978. HSL, also called Double Hexcone Model due to its roots in the RGB color space, is a cylindrical color space with two linear dimensions – *saturation* (S) and *lightness* (L) – and the angular *hue* (H) dimension. Hue is an angle on the color circle ranging from 0° to 360°. Saturation can be thought of as the perceived intensity, or "vividness" of the color [5]. A 100% saturation results in a vibrant color while a 0% saturation results in a shade of grey. Lightness is the relative brightness of a color: any color with 100% lightness is white and any color with 0% lightness is black.

1.1.2 Usage
Image editing software have widely used the HSL color palette, in particular for color selection tools. The most common form of color picker is a two-dimensional slice through the hexcone with saturation and lightness as axes, coming together with a slider for hues.

In the specific case of web applications and web design, HSL is the main color space for color picking. All major Content Management Systems (CMS) and front-end frameworks for the web such as *Drupal* or *Wordpress* display HSL color pickers in their core release [6][7]. The *Adobe Creative Cloud* software collection remarkably exploits the power of hue-based color spaces with the *Adobe Color* tool [8], one of the most popular color scheme generators among designers.

Recent studies reinforce the theory that a website's credibility is judged with soft criteria such as background colors [9]. One common technique used by digital artists and designers for style guide creation is to allow high values of saturation for only one hue - thus creating one predominant vibrant color - and limiting other hues to low levels of saturation. This method is confirmed by Papachristos et al. [10], as they found that the dominant color's brightness had a stronger emotional effect than its hue and conclude that the optimal combination is a "highly saturated dominant color with a low saturated secondary color". Camgoz et al. reached similar conclusions, stating that brightness and saturation levels were more significant than the hue component in color preferences [11].

Designing coherent corporate style guides has become of a paramount importance for communication. Fortunately, front-end developers and web designers have gained access to the manipulation of the HSL space with the incorporation of the HSL and HSLA palettes in *W3C Recommendation CSS Color Module Level 3* by the World Wide Web Consortium in 2011 [1]. In spite of being a more intuitive model, learning how the hues are positioned on the color wheel is a necessary condition for the mastery of designing with HSL.

1.2 Learning HSL
1.2.1 Color pickers, software and practice
Due to the complexity of color models such as HSL, specifying the three exact values needed to generate a given color requires tools for direct manipulation. Color selection is only possible through interfaces that allow incremental exploration and refinement. Scarcely are raw HSL characteristics presented; the use of such color pickers most often drives users to being unaware of the underlying color model. Yet, understanding the HSL color scale is key to a thorough investigation of the color theory as applied to design [12].

As the HSL color model requires learning to use effectively [5] but no significant e-learning environment that addresses color recognition skills is available online, designers have to rely on sole experience to associate a color with its HSL components.

Contrary to within a controlled lab environment, establishing a game environment helps subjects behave more naturally [13]. Besides, gamification in psychophysical studies improve user experience and the quality of the study [13]. One might then consider implementing a color game designed for the special needs of studying learning in HSL.

1.2.2 Existing games
A few color recognition games have tried to bring back the users to direct manipulation of color model dimensions. The game *Hex Invaders* is a color recognition skill test where the player success relies on their ability to quickly associate a color with its description in the hexadecimal web format [14]. *RBG Challenge* is a guessing game of a moderate difficulty where one has to choose between three random RGB color values [15]. Neither of those games train users nor use HSL.

Whathecolor [16], however, can be played with HSL. The player has to type the three values to match a target color in an input controller. The comparison of the target color and current chosen color is live and the aim is to find the right set of values as fast as possible. The precision required for target reach is very high, thus the end game is fine tuning and the player relies on the displayed information of the distance to find local maxima rather than looking at the colors. Driving the user away from paying close attention to the HSL values inevitably infringes on learning.

Even though all described games involve color in a specific context, the player's recognition abilities are evaluated rather than trained.

1.3 Aim of the project
This thesis attempts a new approach of learning color models with an online game. HSL serves as a scale to measure color recognition abilities in the context of a color-based game specifically designed to enhance learning, named *Hue Hunt*.

As HSL is a tri-dimensional color space, a three-step

approach was seen as most relevant in order to progressively enhance user ability to associate colors with their HSL values. A first level of difficulty focuses on the hue component of a color, involving colors with fixed saturation and lightness. A second level adds lightness as a variable, with fixed saturation. A last level allows colors with almost any HSL set of values.

Throughout the study, the focus remains on the measurement of learning across the color space. While the visual environment of the user interface is designed to enhance learning, collected data are analyzed only to quantify performance and learning, as well as to reveal potential patterns. The analysis aims at answering two research questions:

- *Which hue areas show pessimal and optimal learning, if any?*

- *Does the game reveal trends on learning HSL with regard to game progression?*

It is noteworthy that the relevant information about in-game variables are restricted to what, in color finding, is called the *convergence phase* – i.e. the first few actions of the player – as learning might not occur in the later refinement phase [5]. Since short-term involvement has to be maintained to complete the game, some gamification mechanics like leaderboards are integrated to the prototype for extrinsic user motivation [17].

2. RELATED WORK
This section introduces the research elements that drove the study and led to strict requirements for the interface and user interaction elements, as well as the definition of criteria for quantifying performance and learning.

2.1 Choosing HSL
2.1.1 HSL versus RGB
The RGB color space arose as a paramount color model as early as the nineteenth century, and remained the first color model until the late twentieth century, principally owing to its close relation to the technology behind the production of light and colors. RGB corresponded directly to graphic systems' underlying hardware, i.e. the three red, blue and green electron beams on a color Cathode Ray Tube [4][18]. As a consequence, programmers have repeatedly used RGB for the implementation of colors in user interfaces because there was no easier and more widely-used model [5].

In the late 1970's though, Smith [3] introduced the HSV color model, naming *hue*, *saturation* and *value* as three axes for a more intuitive color description. The same year, Joblove and Greenberg [4] defined HSL, where L stands for lightness and is closely related to *value* in HSV.

Subsequent research amply acknowledged the superiority of HSL over RGB for color description. Tastl and Purgathofer [19] assess that specifying a color in terms of hue, lightness and saturation is easier than using red,

green and blue. Berk et al. [18] compare HSL with the *Color Naming System*, or CNS, which is a three-dimensional adjective-based description system used in textile and paint with colors such as "dark vivid purple". It turns out that the very definition of the CNS is closely related to HSL, as the discrete choices of adjectives can be mapped to the hue, saturation and lightness scales. They conclude that the HSL system is a successful attempt to exploit the psychology of color recognition with dimensions that are perceptual rather than physical. Smith and Lyons later took a step further in seeking the most intuitive color space by proposing the *Hue Whiteness Blackness* (HWB) model [20], but it has as of yet not reached an equivalent popularity.

The study performed by Schwarz et al. [5] is a reference work for empirical data on several color models. They led an experiment with the purpose of comparing RGB, HSL and other models – *LAB*, *YIQ*, and *Opponent* – for color matching. A main outcome was that the participants matched the target colors quicker but with less accuracy with the RGB color space while HSL brought more accuracy but a slower progression towards the color target. They also differentiated between two phases in color finding and used the notion of *Color Distance Unit*. Both are further discussed in section 2.2.

Their analyses also provide elements about the learning differences between the models. While a significant learning occurred for all studied color models, there was no evidence that any learning happened for the RGB color model. This last criterion is a strong edge for HSL as opposed to using RGB in color training.

2.1.2 Using the Macbeth Color Checker
Quite a number of color-based user studies employ a set of colors called the *Macbeth Color Checker* [18][21][22], which is an array of 24 natural, chromatic, primary and grayscale colored squares in a wide range of colors that represent natural objects. Using such colors for a learning game is a prima facie good basis to explore the HSL space.

Douglas et al. advocate the use of the Macbeth Color Checker and state that it is essential for experimental reliability to use colors that are rationally selected independently of a particular model [21]. However, if we parse the color components to HSL, we notice that lightness and saturation for most colors fit within restricted sets of values – e.g. a saturation between 25% and 50% for 13 out of 18 colors without counting the shades of grey. This is not optimal when aiming at evaluating a set of colors evenly distributed along the three HSL axes.

2.1.3 HSL limits
Due to impossible recognition of hues at certain locations of the HSL space, the colors displayed in an HSL-based color study interface should be restricted to only a portion of the space. Indeed, with extremely low saturation, the color is a shade of grey and all hues produce the same shade of grey. Extremal lightness – both low and high – also results in non-recognizable hues [20].

2.2 The process of color finding
2.2.1 Measuring color distance
In the process of evaluating color finding and learning, there is a need for a scale to measure distances between colors. A linear definition of a color distance in the HSL space could be the usual Euclidean squared root of the added squared differences in H, S and L. However, experience showed that the variable sensitivities of the eye photopigments across the color spectrum engendered differences in the perception of color distances. Bachy et al. [23] defined the *Hue Discriminaion Curve* (HDC) and showed that it consistently displayed minima and maxima – color discrimination is optimal in the yellow-oranges and blue-mangetas while being pessimal around green. Murch [24] also described the changes in the green area as being among the less readily perceived due to photopigment sensitivity.

Schwarz et al. [5], Douglas et al. [21] and Berk et al. [18] use a *Color Distance Unit* (CDU) that corresponds to the Euclidean distance in the CIELAB color space, an international standard defined for perceptual uniformity [25]. The definition of CDU in the CIELAB color space has evolved during the twentieth century with growing complexity and accuracy [26], together with the evaluation of the *Just Noticeable Difference* (JND) between similar colors.

2.2.2 The two phases of color finding
The process of color finding involves two phases: the *convergence phase* and the *refinement phase* [5][21]. During the former, subjects approach the target color with leaps of substantially high distances, and a significant likelihood of overstepping the target. In the latter however, the subject is already in the vicinity of the target and attempts to fine-tune the match.

Behavioral problems arise in the refinement phase, where users often base their choices not on the understanding of the underlying model but rather on incremental techniques such as hill-climbing [10] with tiny steps along one dimension at a time. This technique can be problematic as the search space is very large and small local changes might not lead directly to the goal. It happens in the HSL space that making a local choice of each separable dimension does not result in getting any closer to the target color.

The convergence phase is more of a one-shot attempt carefully thought through. This is precisely where learning is required [5], and happens [5]. For those reasons, the game experience has been oriented towards the first phase of color finding: the convergence phase.

2.2.3 Color feedback and importance of the UI
Douglas et al. studied the effects of low and high visual feedback – about the user's current location within the color space into play – on speed and accuracy of color matching for RGB and HSV [21]. They mostly emphasize the fact that the color model into play might have a lesser importance than its presentation on the interface, and the way the user can manipulate the model. Visual feedback and design of the interface turn out to be more important factors in improving the usability of a

color selection UI than the model itself. Douglas criticized the Schwarz study [5] in that the feedback was poor – no information about where the current color is in the space, nor about the effects of each parameter – and that users had difficulty creating a mental model of the color spaces [21]. A color study interface ought to be designed with a balance in this regard. With too much feedback (clues, precision rate or distance to the target) comes a higher risk of distracting users from understanding the HSL model by enabling hill-climbing techniques, whereas one can hardly expect any progress if H, S and L mean nothing to the user.

3. METHOD
As stated above, the core element of the thesis is an experimental game prototype. The game, *Hue Hunt*, assigns target colors with random HSL values. The goal is to find these values using color recognition skills. The game proposes a progressive approach where the player is able to compare the current color – corresponding to the chosen HSL values – to the target color at each step.

The online prototype of the game follows a set of good-practice rules for e-learning and displays a UI with neutral colors that do not interfere with target colors. A recording module gathers players data and enables investigation about the learning process, i.e. how users get better at associating the colors with their corresponding HSL values.

This section firstly presents the final version of the prototype with illustrative in-game images. Then, details are provided on how research-based and feedback-based choices shaped the development of the game, the user study and the subsequent analysis methodology.

3.1 Game flow, aesthetics and colors
3.1.1 Game flow and elements
Gulliksen et al. [27] propose key principles for *User-Centered Systems Design* (UCSD) together with a process to design high-quality products with a focus on usability. The study aimed at following most UCSD guidelines, with an early prototyping, some iterations on the design choices, and regular user involvement. Game design is an extension of those principles, as all interfaces must be mapped to a specific stage of the user experience.

Hue Hunt has a three-step game flow. On the first screen (Fig. 1), players:
- get an overview of their current progression in terms of experience and unlocked levels
- select the level to play
- preview the target color for the selected level
- are given animated explanations on the meaning of each color dimension
- enter a player name (first round only)

Clicking on the *Play* button leads to the second screen (Fig. 2), where players:
- view the color target
- view the last color guess (after first shot)

- can enter HSL coordinates to match the color target
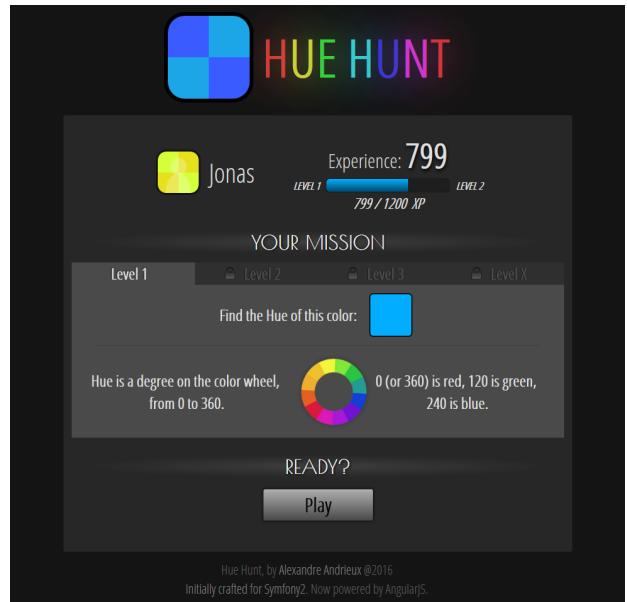- get information about the current round, i.e. number of remaining shots and current score

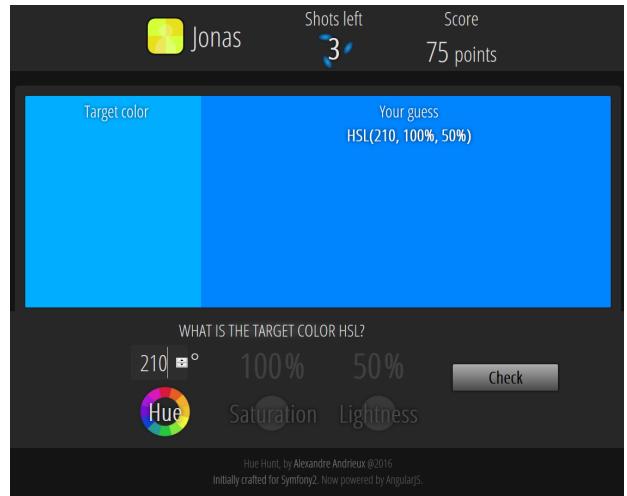**Figure 1. Hue Hunt first game screen**

Figure 2. Hue Hunt second game screen, for a target color of 203° hue and at level 1

Displayed after each round, the last screen (Fig. 3) shows:
- the target color and the according HSL values
- the performance for the round
- a visual history of the color guesses, together with the entered HSL values on hover
- achievements unlocked throughout the game
- a leaderboard showing the ten highscores for the current level with the associated colors and player names

From the last screen, the *Choose level* button brings back to the start screen. Such a short gameplay loop with immediate feedback on progress was key to stimulating and maintaining user involvement.
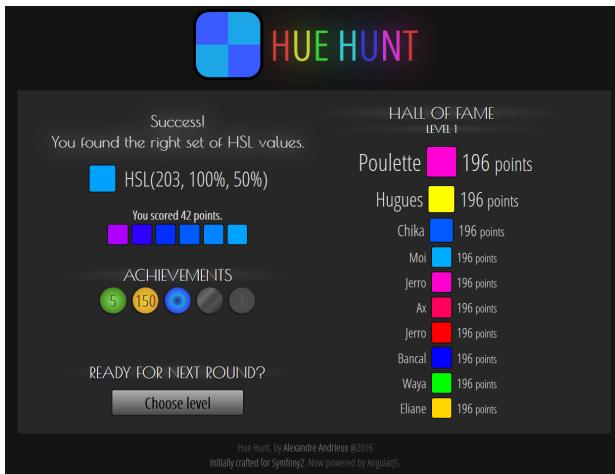
Figure 3. Hue Hunt third game screen

3.1.2 Levels and experience

The game brings into play a progressive approach to the manipulation of HSL. Three main levels and a bonus level are proposed. The first level, and purposely the easiest, is about getting familiar with the hue coordinate. It has one degree of freedom, i.e. one variable of investigation. Saturation is set to 100% and luminosity to 50%.

The second level requires guessing both hue and saturation – it has two degrees of freedom (hue and saturation) and a luminosity set to 50%. The third level introduces lightness, thus having all three degrees of freedom with both hue, saturation and lightness as variables. A last level was added for competitive players, similar to the third but with only three attempts to reach the target's vicinity.

In order to assess progress and unlock the next levels, player experience is measured. After each round, the player receives an amount of experience equal to the performance of the round (see Appendix 2 for the definition of performance).

Experience thresholds trigger the unlocking of a new level, which requires an average of 15 rounds, depending on the player's abilities. Players who perform well – i.e. who demonstrate a good mastery of the current coordinates set – are thus quickly brought to higher levels.

3.1.3 Colors and design

Standards for good designs exist, and most users of the Web are familiar with them. Designers train users to expect certain items to be located in certain places, like buttons, menus or keys. Not fulfilling those expectations can cause frustration and annoyance. Cooper [28] strongly advises designers to obey standards unless there is a truly superior alternative.

Overall design

Creating a suitable environment for learning online is a complex task. The user interface has to follow the most general guidelines about ergonomy and usability. Three key criteria for the evaluation of e-learning systems'

visual design are usability, readability and attractiveness [29].

Defining a graphic identity and style guides for Hue Hunt was critical. Several visual elements were created to strengthen the game identity, such as a colorful logo, player avatars reminding the logo design and icons for achievements. Despite displaying grey backgrounds, the overall dark appearance is livened up by vivid colors, most often carrying meaningful color information. An *almost flat design* approach shaped the elements' visual properties (shadows, borders, buttons).

Neutral background colors

Whether or not they actually are easier to use, aesthetic designs look easier to use and have a higher probability of being used [30]. Together with the need for good aesthetics and alignment to the web design standards at play at the time of the study, the colors of the UI had to be carefully chosen. Indeed, the game should enhance the learning for a wide range of colors. The surrounding colors of the interface should therefore be neutral. In this regard, the various game screens only display shades of grey in the background and on the elements, except for a colorful logo which is not visible during a round of color finding and the game controls which provide an indication on the effects of each dimension.

Minimal saturation, min and max lightness

As explained in section 2.1.3, grey, black or white must be avoided as evaluation colors in the study because an infinite number of hues can lead to those colors. That is why colors with a saturation lower than 20% and a lightness out of the range [20%, 80%] are discarded from the game's proposed color possibilities. Within those ranges, the colors are randomly chosen. Thus, a balanced distribution of hue, saturation and lightness is ensured throughout the game.

3.2 The game development process
3.2.1 User testing

Following UCSD principles, Hue Hunt received iterative feedback, with varying amount of users testing the prototype at each step. Over a period of three months, several game features were implemented, or underwent an important change. Along with users without prior knowledge of web design nor color models, the game was tested by experienced UX designers of a web agency. Their feedback on the game experience ranged from the visualization of color to the gamification features.

Hue Hunt final user study lasted about three weeks after an online release and communication within personal and professional networks, mostly via e-mail. Nothing else than playing the game was expected from participants. It gathered 218 players, among which 97 reached the second level and 24 completed the third level. A total of 2972 rounds were played with 16713 color guesses.

3.2.2 Research-based and feedback-based elements

This section inventories the game features that either were implemented after deepened research analysis, or underwent a significant change in the process of iterative feedback.

Inputting HSL data
Initially, entering HSL data was done through typing the three color components in the page's web address – or URL – separated by a slash: *http://[game.path]/h/s/l*. This mode of interaction drove the user's attention outside of the UI scope, and the game playability was low. Replacing the URL-based navigation mechanics by dynamic inputs allowed a balance between ease of use and attention to the color model itself. As explained in section 2.2.3, not distracting the user from the understanding of the color model itself and preventing blind hill-climbing techniques are essential. In Hue Hunt, players do need to reflect on their previous attempt to gauge how much is needed to hit the color target, and in which direction.

Levels
The initial game had only one level, with all three degrees of freedom. The player had to guess both hue, saturation and lightness at the same time, and the goal was to approach the target color using as few shots as possible. The threshold for winning was set to a value corresponding to approximately 5 CDU expressed in the Euclidean definition of distance, depending on the number of dimensions.

The first sessions of user feedback showed that having to guess the three components of a formerly unknown model was an oversized challenge – with an average of 20 shots to reach the color target vicinity – and did not engender learning. This is why the final game proposes several levels of increasing difficulty.

A limited number of shots
As explained in section 2.2.2, the focus of this study is the convergence phase of color matching. Excluding the refinement phase was at first solely ensured by ending the game as soon as the color guess was in the vicinity of the target, with distance thresholds of increasing tolerance across levels, as the number of degrees of freedom grew. With such a setup, the calibration of minimal distance to target was critical to getting a measurement of player performance. The quality of the definition of performance – and thus learning – could not rely on approximate estimations of what a satisfactory distance threshold should be.

An alternative means of limiting the study to the convergence phase while ensuring data quality is to allow a low maximal number of shots. With only six shots to reach the target, and the automatic win condition when reaching the target's vicinity, the collected data is pertinent as players have to give their best guesses. Furthermore, it is made clear with an in-game message (Fig. 4) that the first shots are worth more points than the last, as the points value of a shot is halved at each step.

Another element was removed to prevent hill-climbing techniques: clues on which coordinate to change and in which direction in the color space. Such an indication totally disinterested players in the actual HSL values.

Last but not least, the game used to display the matching precision. Informing players about the precision – i.e. the

closeness of the color guess – with a numerical quantity led them to not even look at the colors but to rely uniquely on figures. As an incentive to engage in progression, precision was replaced by a *score* indication. Players cannot rely on their current score to guess the color target. The leaderboard's ranking criterion was accordingly moved from *number of shots* taken to match the color, to *final precision*, to *score*.
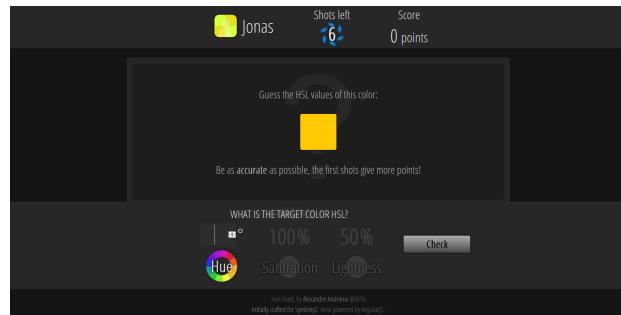

Figure 4. Game screen for a first color guess

Layout and target color
At first, header elements such as *precision* (later *score*) were placed far from the vertical central axis of the UI. It turned out most players did not even notice the element's existence, as the key elements are the game colors (guess and target) and the inputs right below. Thus, all header elements were horizontally centered.

Initially, the color target was a small square circled by a black border and positioned at the corner of a wider area with the current guess as background color. Other than hampering the perception of the target color with risks of an optical illusion called the contrast effect, such a display was redundant with a similar target color square placed in the UI header. Both were removed and remodeled into a single larger color area directly adjacent to and in direct contact with the current guess color area (see Fig. 2).

Selected feedback on the color model
Most of the user interaction happens on the second in-game screen. The information given on the start screen about the effects of hue (see Fig. 2), saturation and lightness are not sufficient for players to remember them all the way through the round. In order to leave this information available at any time without overloading the UI, the animated explanatory icons of the start screen were also placed below the inputs – a rotating segmented color wheel for *Hue*, a disk with a color changing from vivid blue to a shade of grey for *Saturation*, and from black to white through blue for *Lightness* (Fig. 5). In fact, some users reportedly used the hue color wheel icon to locate and orient themselves along the hue dimension. The impact of the color wheel icon is discussed in section 5.1.2.
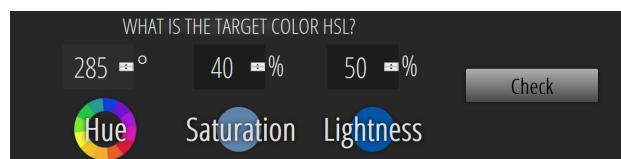

Figure 5. HSL inputs and animated visual clues

On the win screen, players could view the target color and its HSL components, but had no idea how far they were from the target along the way. A visualization of past color guesses was added to the final screen, with aligned squares containing the colors. It successfully helped players compare their color guesses to targets, while enriching the overall appearance of the game.

The improvements on the history of guessed colors on the win screen went further. More than seeing the color squares, users expressed the need to reflect back on the HSL components of their guesses. This last piece of information is now printed under the color squares on cursor hover (Fig. 6).
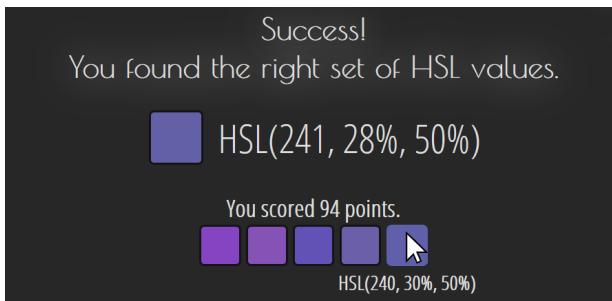
Figure 6. Win screen color guess history

Achievements
Due to the repetitiveness of the game, the sole leveling mechanics were not sufficient to maintain proper user involvement. Achievements with gradual difficulties were added as extrinsic motivation features. Two achievements require playing the game for a certain duration while three achievements reward specific performance, like finding the right HSL values with only one guess. A panel on the win screen inventories the obtained achievements (Fig. 7).

Figure 7. Win screen achievements

3.3 Result analysis
3.3.1 Hue Hunt Results
Retrieved data from Hue Hunt amounted to 2972 sets of rounds with the pattern *(username, roundLevel, performance, targetH, targetS, targetL, timestamp)*. In order to visualize and analyze the color-specific dataset, a data visualization interface – *Hue Hunt Results* – has been specially created (Fig. 8). The interface allowed to make sense out of the dataset and highlight colored patterns with the use of a Streamgraph [31], bar charts and data processing functions.

Parsing performance data into learning data requires several computational steps detailed hereafter.

Figure 8. Hue Hunt Results, a dedicated tool for game data analysis

3.3.2 Technical details
Hue Hunt[1] was initiated under the *Symfony 2* PHP framework, but was later transposed to the *AngularJS* framework (version 1.4.7). The project with its results interface amounts to 5150 code lines altogether (overall Git minus data export), mostly JavaScript and CSS. Other than *AngularJS*, libraries and tools such as *jQuery* and *Underscore.js* were used. Layouting is coded with CSS3 Flexbox.

Hue Hunt Results involved *D3.js*, *FontAwesome* and the *Less* CSS preprocessor. *Firebase* was used to store and retrieve data asynchronously. The final exported JSON dataset weighs 734Ko and contains 2972 *round* objects.

3.3.3 Quantification of performance and learning
Color distance
Even though the CIELAB color space evens out the perceptual differences across the hue dimension, the focus of the present study remains the learning. In this regard, the quantification of learning remains based on the Euclidean distance in the HSL color space. Similarly to the notion of HDC (section 2.2.1), the *Hue Performance Curve* (HPC) and *Hue Learning Curve* (HLC) are defined and used. Deriving them from color distance measurements requires mathematical definitions to quantify performance at each step of the game. Details on the experimental models are given in Appendix 2 with the associated formulas.

Performance
A round of color finding has a fixed color target, and a maximal number of shots. As the aim is to measure the immediate perception of the color as expressed in the HSL color model, the performance has to emphasize the first guesses accuracy (convergence phase) and diminish the weighting of the last shots, which tend more to border on the refinement phase threshold. Thus, the first shot has twice the weight of the second shot, which has twice the weight of the third, and so on.

Schwarz et al. [5] and Douglas et al. [21] use a color distance threshold of 25 CIELAB CDU to separate the convergence phase from the refinement phase. This study uses a win threshold of 2 Euclidean CDU for the first game level – which corresponds to an average of 5 CEILAB CDU across the HSL color space – in order to

1   The game in its final version is playable online at *http://icosacid.com/huehunt* at the time of the study

stop the game only once the colors are so close that the eye might not perceive a difference. Levels 2 and 3 have more permissive winning conditions: respectively 5 and 10 Euclidean CDU. Those values allow to reasonably assess that we measure performance within the convergence phase without too much restriction – i.e. including potential guesses situated in the refinement phase. A maximal threshold in the impact of color distance on performance has been arbitrarily set to 75 Euclidean CDU (see appendix 2 for details on the filtering function). It prevents overly erroneous guesses from adding to the performance.

Excellent players might get into the refinement phase before they run out of shots, which is by no means an issue as the performance earned in the last shots will in all likelihood remain high. The total performance at the end of the round will mirror the good guess. That is why regardless of the player's expertise, the model measures performance on the sole convergence phase.

In Hue Hunt Results, a *Hue Performance Curve* (HPC) is built to compare performance across the hue dimension. Each hue (h_c) is mapped to the addition of all players' performance in the vicinity of h_c, averaged by the number of contributions.

Learning

As performance is peculiar to the player and strongly dependent on the color target's hue, the dataset has to be subdivided into several smaller datasets of hue width SHR (*Subset Hue Range*, in degrees) containing rounds with a target hue comprised between a minimum and a maximum value. Filtering operations based on username and hue subsets are the first steps towards highlighting differences in learning across the hue dimension.

Learning occurs when the performance gets better along the game. Since performance is quantified for each round, the *Hue Learning Curve* (HLC) can be obtained by comparing successive performance. As performance within a hue subset might present a high variability, averaging several shots' performance provides a more stable measurement. Thus, learning can be considered as the derivative of the performance, smoothed with an averaging *Learning Round Scope* (LRS) of 3 rounds. It is noteworthy that the learning is forced to zero in case of a negative computed value.

The HLC is represented by a Streamgraph [31] with layers of a width corresponding to learning for a player within a given hue data subset, stacked by color. The HLC is a function of the position of the round in the player game sequence.

Overall learning

Out of the definition of learning and the necessity for splitting the dataset into smaller subsets based on hue ranges, one can obtain an *Overall Hue Learning Curve* (OHLC) by summing the color-specific total learning of all players. While the HLC shows learning across rounds, the OHLC shows learning for all players as a function of hue. Appendix 2 provides details on the calculation of performance, learning and the construction of OHLCs.

4. RESULTS

Results on the HPC, HLC and OHLC are presented in this section. All figures originate from Hue Hunt Results interfaces. HPC and OHLC both present local minima and maxima in performance and learning respectively; the focus remains on which hues produce those extrema rather than on the total computed values, which derive from the specific chosen definitions of performance and learning. Variables used in this section are further detailed in Appendix 2.

4.1 Hue Performance Curve

The HPC is obtained by adding all players' performance regardless of level and player experience. In order to extract enough data for each target hue h_c on the x-axis, a dataset subdivision technique was used to cluster rounds with a target hue close to h_c. Averaging all rounds' performance produces a stable performance measurement for each target hue h_c from 0° to 360°. Figure 9 is the HPC obtained with all 2972 rounds from all players at all levels.
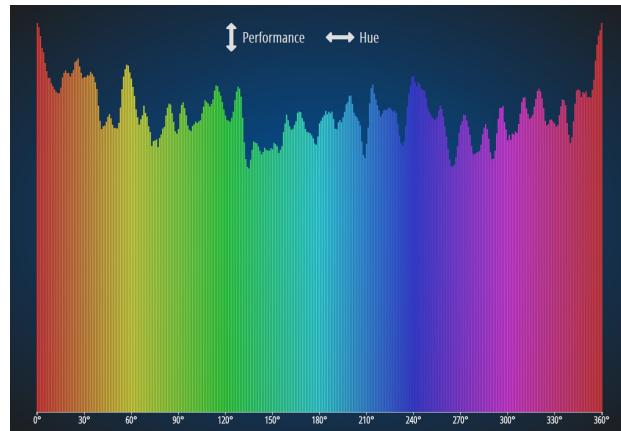

Figure 9. Hue Performance Curve

The HPC shows a few minima and maxima of note which are presented on Table 1.

Table 1. Hues with local maxima in performance on the Hue Performance Curve

HPC local maxima	**0°, 30°, 60°, 120°, 240°**
HPC local minima	**45°, 75°, 145°, 205°, 260°**

Assuming some learning progressively occurred, this HPC includes rounds of experienced players as well as beginners. The HPC involving exclusively rounds played on level 1 confirms the previous HPC and shows deeper performance gaps, with a few more notable local minima in greens and purples (130°, 290°, 325°).

At this point, it is only possible to assume that learning somehow contributed to filling the initial performance gaps on the HPC.

4.2 Hue Learning Curves

The HLC streamgraph displays several colored layers, each corresponding to the learning for rounds of a given player within a given hue range. Hue Hunt Results

allows to choose different hue data subsets based on color, but also to select all hues split into slices of equal hue width. For instance, selecting *40° slices* generates learning layers for hue subsets [0°, 40°], [40°, 80°], […], and [320°, 360°]. All stacked up layers form the HLC. A layer is colored with the medial hue of its hue range. Figure 10 shows the 40° slices HLC for player *Aegeon* (anonymized).
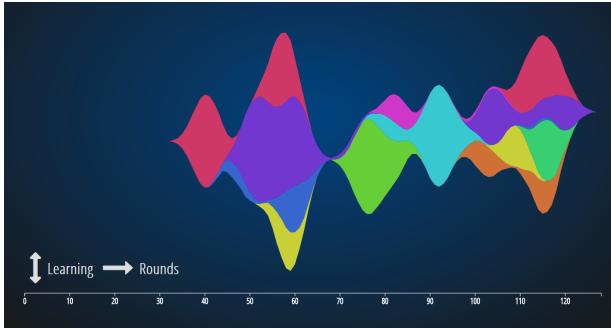

Figure 10. Hue Learning Curve with 40°-wide hue data subsets for one player

This player seems to have gone through two learning phases. The first phase (rounds 35 to 70) with a learning mostly in magenta, purple and dark blue – i.e colors of the second half of the hue circle – and a second phase (rounds 70 to 120) with a slower but more heterogeneous learning. The streamgraph representation was chosen to better highlight variations in learning with information on color. Figure 11 shows the 40° slices HLC for all players.
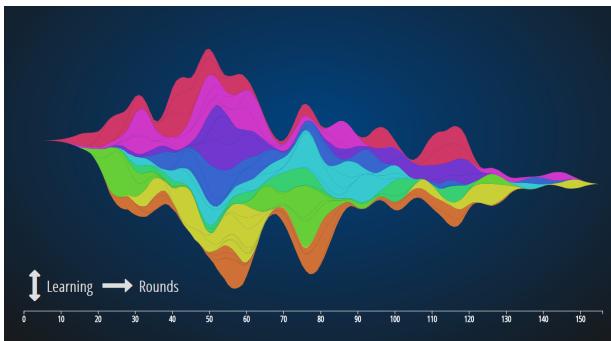
Figure 11. Hue Learning Curve with 40°-wide hue data subsets for all players

This streamgraph gathers in one single view all the data collected during the study. More than providing evidence that Hue Hunt has successfully engendered learning on the HSL color model, it reveals trends on the learning with respect to user color matching experience and color.

First, the repartition of colors within the streamgraph shows that *some colors take more time to learn than others*. In particular, an early learning happened for colors within the hue ranges [320°, 360°] (red) and [80°, 120°] (neon green), with a late learning for colors within [200°, 240°] (blue) and [240°, 280°] (violet).

Then, as a thick layer means high learning and a thin layer means low learning, the streamgraph clearly shows that more learning happens in the second half [180°,

360°] of the hue circle, until approximately round 70 when the greens and the central hue range [160°, 200°] start to rise in learning. Overall, learning for the hue range [120°, 160°] (green-blue) remains low.

Hue data subsets of 40° width (SHR) were chosen to highlight learning patterns, as the 40° SHR value ensures a good balance between data quantity and relevance with an average of 7 contributors per hue data subset, i.e. 7 players adding to the HLC per color layer. The results on OHLCs confirm that 40° is an acceptable SHR for HLCs.

4.3 Overall Hue Learning Curves

The SHR variable has been extensively used in the process of interpolating lacking data by averaging performance and learning values within hue data subsets. Its impact on a learning curve can be viusualized with the use of Hue Hunt Results' OHLC interface, which draws OHLCs based on a configurable SHR. An animated sequence of OHLC with SHRs growing from 10° to 120° highlights the emergence of patterns in the curve as early as 20°, which persist until around 75°. OHLC variability is high, as a difference of 5° in SHR can result in unequal local extrema.

One must look at the average number of contributors per hue data subset for hues on the x-axis in order to settle for a satisfactory SHR. The OHLC with an SHR of 58° has 15 contributors per hue data subset on average, and the OHLC with an SHR of 30° has 5 contributors. Given that SHR should ideally be minimized for learning precision in hues, but also that a minimum of contributors is necessary for acceptable data relevance, an *SHR of 48°* (Fig. 12) has been considered optimal, with an average of 10 contributors per hue data subset.
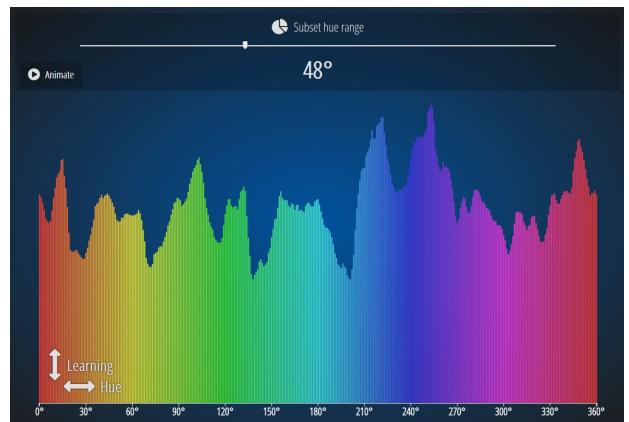

Figure 12. Overall Hue Learning Curve with an SHR of 48°

The chosen OHLC presents a few noteworthy extrema (Table 2).

Table 2. Hues with local maxima in learning on the Overall Hue Learning Curve (SHR = 48°)

OHLC local maxima	**15°, 105°, 130°, 220°, 255°, 350°**
OHLC local minima	**30°, 70°, 140°, 200°, 265°, 305°, 320°**

Most of the OHLC local minima correspond to the HPC minima. As a consequence, it appears that *hues with a low initial performance mostly resulted in low learning*.

However, in terms of local maxima, OHLC may have another relation to HPC. *0°, 120° and 240° HPC peaks seem to produce two OHLC peaks each, with hues placed approximately 15° before and after*. 0° on HPC might be the origin of OHLC 350° and 15°, 120° might result in OHLC 105° and 130°, while 240° might correspond to OHLC's 220° and 255° peaks.

This phenomenon could be explained as such: *players learned most in the close vicinity of the hues they were initially comfortable with*. The presence of OHLC narrow gaps on the very HPC maximal hues may be caused by high base performance resulting in lower learning.

As SHRs between 30° (5 contributors) and 58° (15 contributors) are acceptable, it is possible to validate the previous results against an average OHLC, which sums all OHLCs with SHRs from 30° to 58° (Fig. 13).
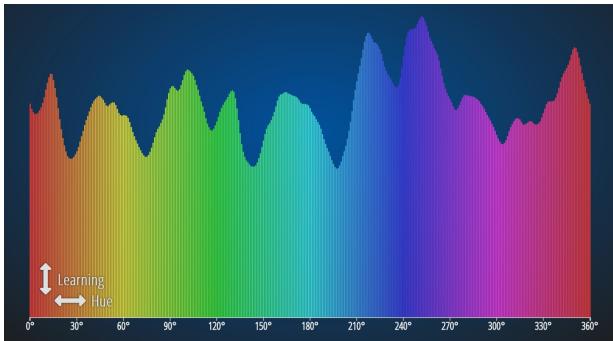

Figure 13. Averaged Overall Hue Learning Curve with SHRs comprised between 30° and 58°

The peaks pairs around the hues 0°, 120° and 240° appear distinctly, which corroborates the previous result. Also, *learning in the second half of the hue circle ([180°, 360°]) is higher*. Finally, the figure confirms a neat *learning in the hue ranges [40°, 50°] (yellow) and [160°, 180°] (turquoise)*.

5. DISCUSSION
5.1 On the results
5.1.1 LRS and learning inertia
The definition of learning, with a LRS of 3, requires that users encountered a substantial number of rounds within the selected hue range before showing any trace of learning. With 40°-wide hue data subsets, all HLCs clearly show that little or no learning is recorded before round 20, even though they may have already improved at HSL. Players might learn HSL at an earlier stage than the HLCs reflect.

The whole analysis was carried out anew with an LRS of 2. While the OHLC was evener with shallow peaks and gaps, the HLCs showed notable learning as early as round 10. Furthermore, it became possible to observe patterns on HLCs built out of 20°-wide hue data subsets:

colors with early, late, low and high learning remained unchanged overall. Despite resulting in less reliable data about learning, the results were consistent.

5.1.2 The color wheel's impact on performance
The central eye-catching color wheel displayed on the start screen was initially included as a mere indication of the meaning of hue. It turned out to be the most helpful clue for finding a target color's hue. As it was duplicated on the in-game UI below the hue input, players relied on the colored segmented scale to position their guesses. Some complained about its rotation because it made it more difficult to exploit.

The wheel is a visual clue of the nature of HSL. It thus distracts the user from building a pure mental model of the color space. Players presumably exited the convergence phase and reached the refinement phase quicker with the help of the color wheel. On the one hand, the purpose of the game might get bypassed with a shortcut to the correct answer. On the other hand, without any visual indication of what *hue* is, the player could only rely on textual information which concerns 0°, 120° and 240°. In this case, base performance would in all likelihood become higher around 0°, 120° and 240° with gaps in-between ; the effects of learning would be observable in those gaps. The continuous color clue evens out the peaks on the initial HPC that are caused by textual indications. It also prevents Hue Hunt from early player deterrence due to an excessively difficult first level.

In a nutshell, the use of the color wheel indicator introduced bias – as did the textual message – in the HPC, which should a priori match Bachy's HDC [23]; nonetheless, it served other key purposes with high efficacy: learning and user involvement.

5.1.3 Freedom to play back on low levels
Even though the last unlocked level is automatically selected for a new round, Hue Hunt's start screen allows switching to a previous level. Most players complete levels in a row without getting back to former levels but 19% of players who reached level 2 have done it at least once. Since a player's rounds are sorted by timestamp (i.e. date) and then sorted by level, some players might show late learning (after 50th played round) that still corresponds to level 1. Then learning for higher levels might appear further right, while other players started showing learning for higher levels as soon as their 20th played round on the same graph.

Due to players reaching upper levels within varying round counts and the definition of learning (delays the appearance of learning by 6 rounds), the streamgraph could not show the effect of leveling up on learning – and thus does not provide elements to conclude on the specific effects of introducing saturation and lightness.

5.2 On the method
5.2.1 The chosen color distance unit definition
One of the decisions taken after reviewing literature was not to use the CIELAB color distance unit for measurement but rather the standard Euclidean definition

of distance. CIELAB has the advantage of establishing a balanced hue distribution based on the human eye perception of color. It was expected and witnessed that not using it would result in a fluctuation of performance across the hue dimension. However this does not affect learning, to the extent that the variation of performance is extracted from subsets of restricted hue ranges, in which Bachy's HDC can be considered as constant.

5.2.2 The decreasing weighting of shots
The first shot screen design was successful inasmuch as players did try their best guesses for first shots. Less than 1% of all played rounds have a score of 0. The leaderboard on the win screen ended up displaying only the maximum achievable score for the first level, with very close and high scores on the following levels. New players might be deterred by the complexity of entering the hall of fame and becoming visible to all players. Instead of displaying only the 10 best scores, the hall of fame could contain all scores with a focus on the best ones and a possibility to scroll down to any score. An alternative could be to display the best scores along with a line for the current performance and scores coming directly before and after in the raking. This way, players could get a better feeling of impacting the game with good performance, even though not perfect.

5.2.3 Differences between subjects
Hue Hunt was released online and tested by players with diverse backgrounds. The initial level of the player, i.e. its base performance – determined by the average of the three first played rounds performance in the whole 360°-wide hue range – is evenly distributed from *90* to *155*, while the score upper limit for level one is *196*.

Some might perform very good from the beginning with the help of explanations. As a result, such players have less latitude for quantified improvement, although they might improve at mastering HSL. Some players were indeed highly capable and durably populated the highscores despite their HLC being empty.

An extreme version of Hue Hunt with only one guess to a target color would be sufficient to evaluate players' understanding of HSL and likely in the convergence phase. However, a one-shot approach – akin to the already existing online color games – would likely be of poor interest for learning as players are not able to iteratively reflect on their attempts.

5.2.4 The 0° 360° critical threshold
Hue is a periodic angle-based value. Level 1 indications explain that "*0° (or 360°) is red, [...]*". While typing negative hues or hues higher than 360°, the game will automatically parse it to the right interval, i.e. [0°, 360°]. However, some players might not 'dare' to enter negative values or to jump to the highest hues in the event that the last guess is a red above 0° hue while the target color has a slight rosy tone – with a hue of around 350°. Performance around 0° might be flawed in this regard, most likely for hues between 340° and 360°. This does not seem to be true on the HPC, but it is not possible to reach a conclusion as red is a local maximum of Bachy's HDC.

5.3 Future work
5.3.1 Game improvements
Feedback and global results showed that there were potential improvements to the game. Adding more levels and more gamification elements could retain more users. A global menu with detailed information on achievements, the possibility to consult all the leaderboards, to choose an avatar, to log in and out would improve user experience. Players could get an indication of their average scores in various hue ranges on the start or win screen. Bringing in sound effects is another step towards immersion.

Recognition challenges of a different nature could be proposed, e.g. famous artworks with a predominant hue to be guessed. Time could be introduced as well, with a gameplay similar to *Whathecolor* [16] and an uninterrupted sequence of quick rounds with lowered target precision.

Learning HSL, in all those options, remains the core of the game. However it could as well become a side-effect of a game of another genre somehow requiring angle input. For instance, games like *Crush the Castle* or *Angry Birds* use a swiveling cannon to project birds on a castle. The cannon orientation could be mapped to hue, with a necessity to guess a target color for optimal destruction.

5.3.2 Other color spaces
A similar approach to learning color could be applied to all other tri-dimensional models. *Hue Whiteness Blackness* (HWB) is suggested as a new standard for CSS Color Module Level 4 [32]. Whether this color model is indeed more intuitive for humans to use could be determined with an HWB version of Hue Hunt.

6. CONCLUSION
This study explored the process of improving HSL color matching skills with *Hue Hunt*, an online e-learning game. The creation of a game environment for research purposes was carried with a concern to respecting HCI and web standards, as well as color-specific UI requirements. Hue Hunt prototypes underwent several feedback sessions which highly contributed to improving its user interaction flow and gamification mechanics. The final game was played by 200 users and resulted in a dataset of approximately 3000 elements.

Through a definition of learning that corresponds to variations in performance, the occurrence of learning was successfully witnessed. With the help of *Hue Hunt Results*, an interface specially created for data exploration, the retrieved dataset allowed to carry in-depth analyses on learning trends, with regard to user progression in the game and hue.

Subdividing the dataset based on the color target's hue in order to extract meaningful information was a critical task, thus the adjustment of analysis parameters were thoroughly discussed. Definitions of *Hue Performance Curve* (HPC), *Hue Learning Curve* (HLC) and *Overall*

Hue Learning Curve (OHLC) were proposed with a view to maintaining a balance between data scarcity and quality.

HLCs showed that early learning happened for red [320°, 360°] and neon green [80°, 120°] while colors like blue [200°, 240°] and violet [240°, 280°] required more time to learn. Due to differences in user performance and fixed thresholds for unlocking higher game levels, no conclusions on the effects of introducing saturation and lightness could be reached.

OHLCs revealed that hues with a low initial performance mostly resulted in low learning, while players learned most in the close vicinity of the hues they were comfortable with at an early stage of the game: *0° red, 120° green and 240° blue*. Overall learning was highest in the second half of the hue circle, i.e. [180°, 360°].

7. ACKNOWLEDGEMENTS

8. REFERENCES

1. Çelik, Tantek, Chris Lilley, and L. David Baron. "CSS Color Module Level 3." (2012).

2. A. Deveria, "CanIuse". Web service for features browser compatibility. Retrieved from http://caniuse.com/#feat=css3-colors

3. Smith, Alvy Ray. "Color gamut transform pairs." *ACM Siggraph Computer Graphics* 12.3 (1978): 12-19.

4. Joblove, George H., and Donald Greenberg. "Color spaces for computer graphics." *ACM siggraph computer graphics*. Vol. 12. No. 3. ACM, 1978.

5. Schwarz, Michael W., William B. Cowan, and John C. Beatty. "An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models." *ACM Transactions on Graphics (TOG)* 6.2 (1987): 123-158.

6. "Color: changing the color scheme of certain themes" (2014). Drupal Color module, included as core module in versions 6.x/7.x/8.x. Retrieved from https://www.drupal.org/documentation/modules/color

7. Matt Wiebe, "New color picker in WP 3.5" (2012). Retrieved from https://make.wordpress.org/core/2012/11/30/new-color-picker-in-wp-3-5/

8. Adobe Systems, "Adobe Kuler: Color wheel | Color schemes", Retrieved from https://color.adobe.com/

9. Danielson, David R. "Transitional volatility in web navigation." *It & Society* 1.3 (2003): 131-158.

10. Papachristos, Eleftherios, Nikolaos Tselios, and Nikolaos Avouris. "Modeling perceived value of color in web sites." *Hellenic Conference on Artificial Intelligence*. Springer Berlin Heidelberg, 2006.

11. Camgoz, Nilgun, and Cengiz Yener. "Effects of hue, saturation, and brightness on preference: a study on Goethe's color circle with RGB color space." *9th Congress of the International Color Association*. International Society for Optics and Photonics, 2002.

12. Steven Bradley, "The Fundamentals of Color: Hue, Saturation, And Lightness" (2013). Retrieved from http://vanseodesign.com/web-design/color-theory/

13. Beigpour S., PEDERSEN M., "Color Play: Gamification for Color Vision Study"

14. 352 Inc, "Hex Invaders". Retrieved from http://www.hexinvaders.com/

15. Joey Germeraad & Peter Klaas Vrieswijk, Lastige Gasten, "RGB Challenge". Retrieved from http://www.rgbchallenge.com/

16. Lea Verou, "Introducing Whathecolor: A color game for web developers" (2014). Retrieved from http://lea.verou.me/2014/02/whathecolor-a-color-game-for-web-developers/

17. Nicholson, Scott. "A recipe for meaningful gamification." *Gamification in education and business*. Springer International Publishing, 2015. 1-20.

18. Berk, Toby, Arie Kaufman, and Lee Brownston. "A human factors study of color notation systems for computer graphics." *Communications of the ACM* 25.8 (1982): 547-550.

19. Tastl, Ingeborg, and Werner Purgathofer. "Color Spaces and Human Color Perception." *Photorealistic Rendering in Computer Graphics*. Springer Berlin Heidelberg, 1994. 219-226.

20. Smith, Alvy Ray, and Eric Ray Lyons. "HWB—A more intuitive hue-based color model." *Journal of graphics tools* 1.1 (1996): 3-17.

21. Douglas, Sarah A., and Arthur E. Kirkpatrick. "Model and representation: the effect of visual feedback on human performance in a color picker interface." *ACM Transactions on Graphics (TOG)* 18.2 (1999): 96-127.

22. Giesel, Martin, and Karl R. Gegenfurtner. "Color appearance of real objects varying in material, hue, and shape." *Journal of vision* 10.9 (2010): 10-10.

23. Bachy, Romain, et al. "Hue discrimination, unique hues and naming." *JOSA A* 29.2 (2012): A60-A68.

24. Murch, Gerald M. "Physiological principles for the effective use of color." *IEEE Computer Graphics and Applications* 4.11 (1984): 48-55.

25. CIE. Recommendations on Uniform Color Spaces, Color-Difference Equations, Psychrometric Color Terms. Bureau Central de la CIE (Supplement 2 of CIE Publication 15 (E-1.3.1) (1978)

26. Sharma, Gaurav, Wencheng Wu, and Edul N. Dalal. "The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and

mathematical observations." *Color Research & Application* 30.1 (2005): 21-30.

27. Gulliksen, Jan, et al. "Key principles for user-centred systems design." *Behaviour and Information Technology* 22.6 (2003): 397-409.

28. Cooper, Alan, Robert Reimann, and Dave Cronin. "About face: Interface-und Interaction-Design.*"* MITP-Verlags GmbH & Co. KG, 2010.

29. Tomko, Vladimir, and Larissa Zaitseva. "Visual Design of E-learning systems." *2009 Ninth IEEE International Conference on Advanced Learning Technologies*. 2009.

30. Lidwell, William, Kritina Holden, and Jill Butler. "Universal principles of design, revised and updated: 125 ways to enhance usability, influence perception, increase appeal, make better design decisions, and teach through design". Rockport Pub, 2010.

31. Byron, Lee, and Martin Wattenberg. "Stacked graphs–geometry & aesthetics." *IEEE transactions on visualization and computer graphics* 14.6 (2008): 1245-1252.

32. W3C. CSS Color Module Level 4 (2014) "HWB Colors: hwb() function". Retrieved from https://drafts.csswg.org/css-color/#the-hwb-notation

APPENDIX 1
Example game scenario and achievements

The pictures above illustrate a game sequence with a focus on key steps.
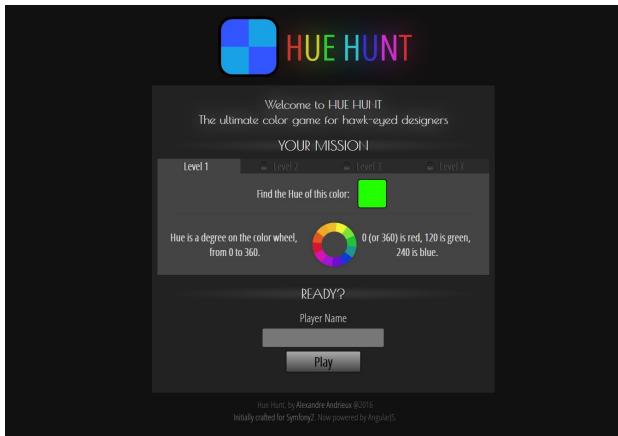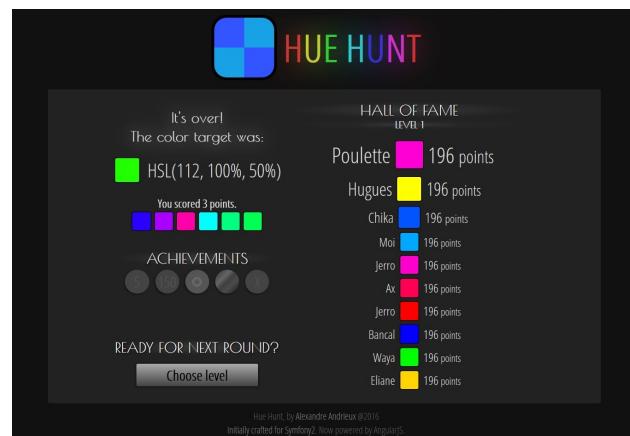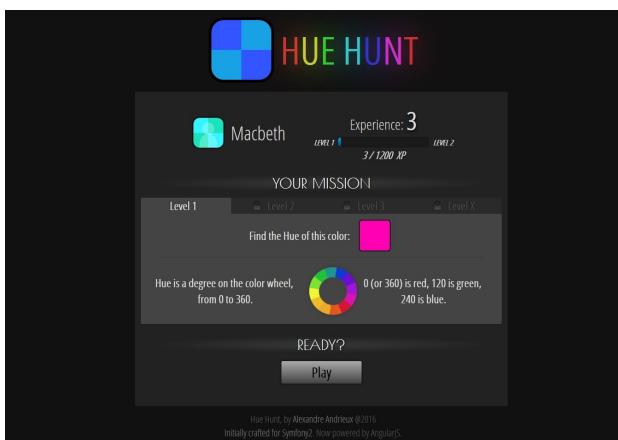

Start screen - Enter player name
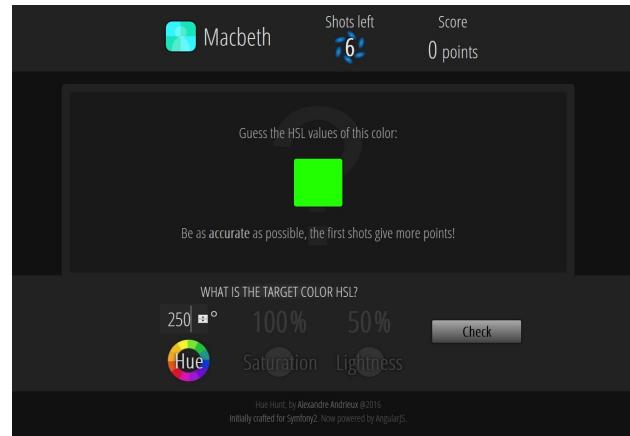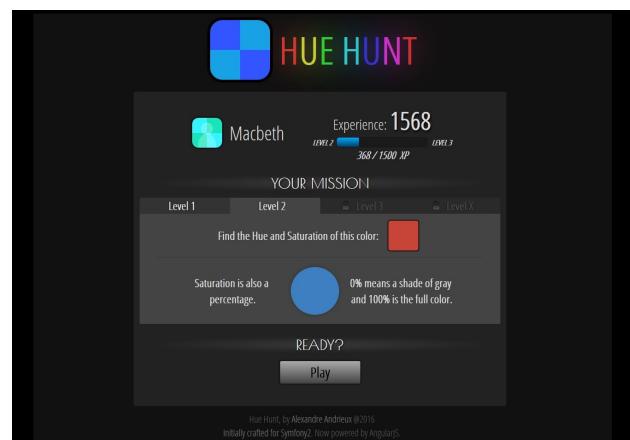

Game screen - Level 1 - Enter first guess
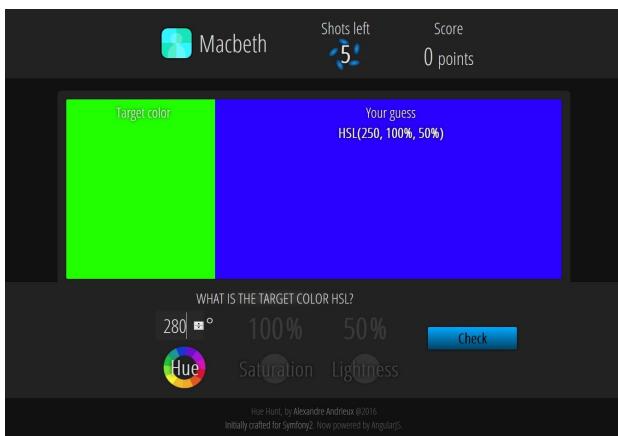
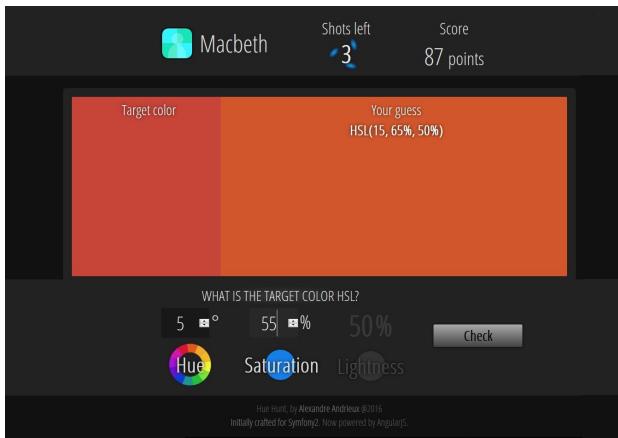**Game screen - Level 1 - Enter second guess**

Win screen - Failure - Round overview and highscores
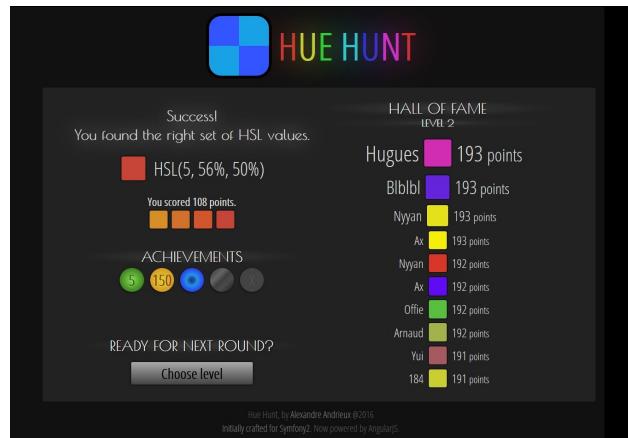

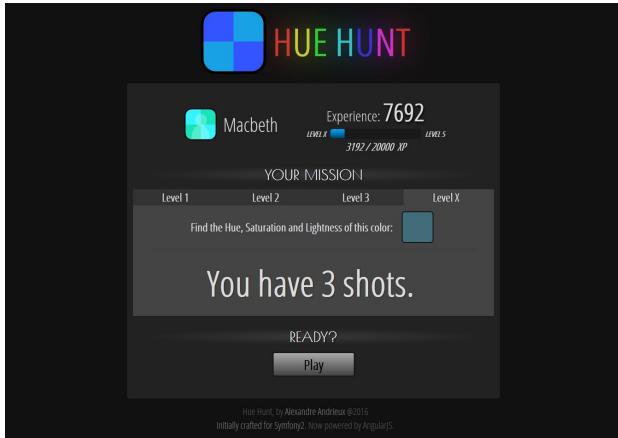**Start screen - Level 1 - Gained experience and new round to play**
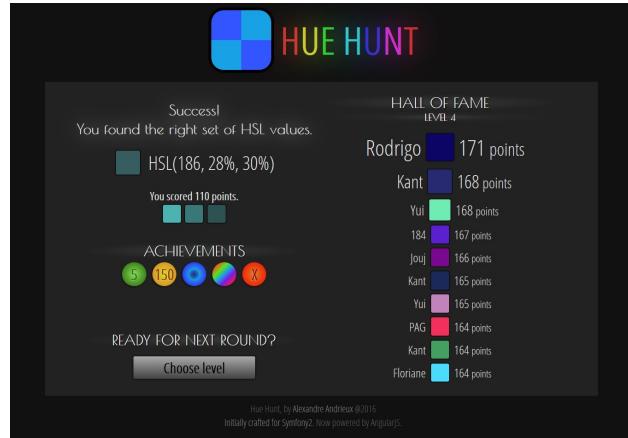

Start screen - Level 2 now available

Game screen - Level 2

Win screen - Successful guess before the end of round

Start screen - Level 4 unlocked

Win screen - All achievements obtained

Some achievements are easier to obtain than others. Here below is a table of all five achievements with their descriptions, which players can read when hovering on the according icons.

Achievement name	Description
On a roll	Play 5 rounds
Challenger	Get more than 150 points
Awwww yeah!	Perform a one-shot
Color master	Get 130 points in all color ranges
Highway to hell	Unlock level X

The achievement *Color master* is obtained by scoring more than 130 points in each of the following hue ranges: oranges [20°, 60°], yellows [60°, 100°], greens [100°, 180°], blues [180°, 260°], purples [260°, 340°] and reds [340°, 20°].

APPENDIX 2
Color distance, performance and learning

Quantifying learning out of HSL guesses required several layers of calculations, as well as defining mathematical models for various variables. Those calculations occur in Hue Hunt background JavaScript functions. This section translates them into formulas with explanations.

Color distance
HSL is a cylindrical three-dimensional color space. An Euclidean measurement of HSL distance can be obtained by adding the squared differences in each coordinate between the current color guess and the target color.

$$D_{target}(h,s,l)=\sqrt{d_H^2+|s-s_{target}|+|l-l_{target}|}$$

where the hue component is an angular distance with a period of 360°

$$d_H=min(|h-h_{target}|,|h_{target}-h|,|h-h_{target}-360°|)$$

As the game flow progressively introduces players to new variables, the primal definitions of distance are simplified, with $l=l_{target}$ for levels below 3 and $s=s_{target}$ for level 1. Scores in harder levels are thus lower on average.

Performance
Players get a score representing how good they performed once a round is over. The score – or performance – is the sum of several contributions, each corresponding to a color guess.

The first versions of the game used a raw HSL distance to stack up at each shot. However, in order to discriminate better between very good guesses and medium guesses as well as to nullify the contribution of a poor guess – i.e. high distance – to performance, a filtering function was used (Fig. 14):
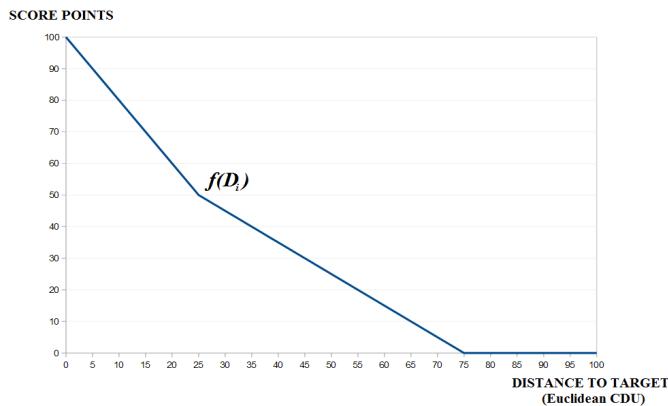

SCORE POINTS

f(D_i)

DISTANCE TO TARGET
(Euclidean CDU)

Figure 14. Distance filtering function

$$f(D_i)=\begin{pmatrix} D_i\in[0,25]\rightarrow 100-2D_i \\ D_i\in[25,75]\rightarrow 75-D_i \\ D_i\in[75,180]\rightarrow 0 \end{pmatrix}$$

where D_i is the Euclidean color distance between the ith color guess and the color target. 75 is a maximal distance above which the shot adds nohow to performance.

In order to reward initial guesses, the first shot has twice the weight of the second shot, which has twice the weight of the third, and so on. Hence the following definition of performance.

$$P_{round}=\sum_{i=1}^{i=m}\left(\frac{1}{2^{i-1}}f(D_i)\right)$$

where i is the current shot and m the maximal number of rounds for the current level.

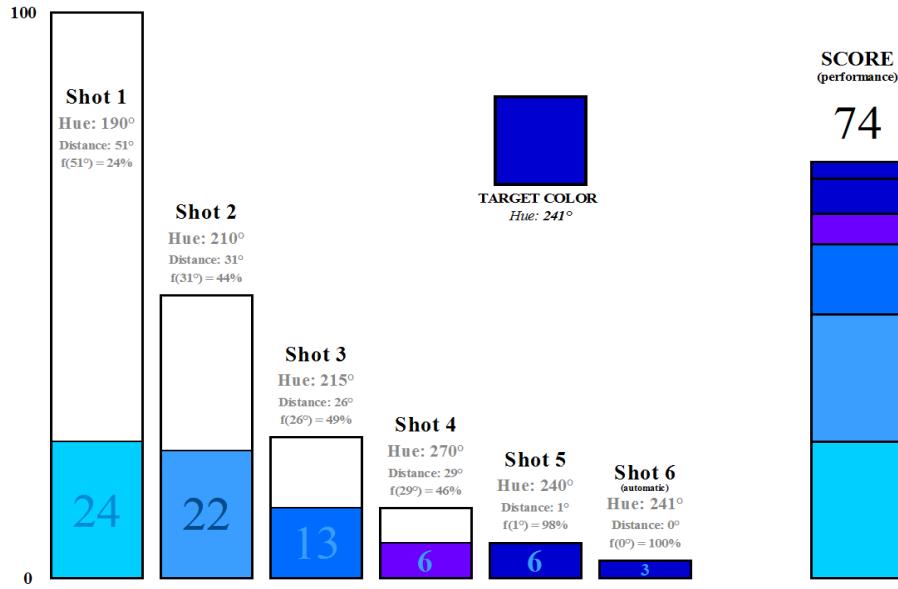Figure 15 illustrates the calculation of performance out of players guesses.

Figure 15. Contribution of color guesses to performance with 1 degree of freedom (hue)

Players might reach the target's vicinity (*2*, *5* and *10 Euclidean CDU* for level 1, 2 and 3 respectively) before the last shot, in which case the remaining shots are automatically considered perfect and the player is directly driven to the win screen with maximal value added to performance at each remaining step.

Learning and HLC

As a round is related to a specific color, the performance is a function of the color as much as the player previous experience with the game. It is only relevant to answer queries such as: "Let's look at learning within the hue range [0°, 30°] for player Jonas". In the process of filtering Hue Hunt's dataset, only rounds with username "Jonas" and a hue comprised between 0° and 30° will be used for the learning measurement. Definitions involving learning (*L*) in this section will be simplified:

$$L = L\left(player, h_{min}, h_{max}\right)$$

One main assumption of the study is that learning is the variation of performance. In this regard, one could define performance as such, where *i* is the *i*th played round:

$$L_i = P_i - P_{(i-1)}$$

With a high variability of performance for some players, one can add relevance to the definition of learning by averaging a sequence of performance, thus involving more rounds.

$$L_i = max\left(\frac{1}{LRS}\left(\sum_{r=i-LRS}^{r=i} P_r - \sum_{r=i-2LRS}^{r=i-(LRS+1)} P_r\right), 0\right)$$

where *LRS* is the *Learning Round Scope* and can be adjusted to average more rounds and get more inertia in learning. With this definition, there is a need for a minimal number of rounds played to get some learning: twice the *LRS*. For this reason, the analysis *LRS* was set to 3, which means that learning can be calculated after the 6th round, still within the selected hue range and for the selected player. That is why a substantial amount of data is necessary to get only a few learning values.

With twelve 30°-wide hue data subsets and the homogeneous repartition of target color hues, 6 (minimal round number) times twelve subsets make an average of 72 rounds played needed to get a single measurement of learning, and to get a few more, say six, this quantity has to be doubled. As the average number of rounds played per player amounts to 13 with only five users playing more than 50 rounds, it appears clearly that we are limited in sharpening the hue-based analysis.

With a *LRS* of 3, the definition of learning becomes:

$$L_i = max\left(\frac{1}{3}\left(P_i + P_{(i-1)} + P_{(i-2)} - \left(P_{(i-3)} + P_{(i-4)} + P_{(i-5)}\right)\right), 0\right)$$

It is assumed that playing Hue Hunt will not make users worse at HSL, thus learning has a minimum of zero. This last definition, together with the choice of players and hue intervals, is the base for plotting HLCs (Hue Learning Curves) on a streamgraph.

Overall Hue Learning Curve
The chosen definition of learning is ambitious and requires a lot of data. Finding a balance between data relevance and data quantity is key to reaching any conclusion on an HLC. In terms of relevance, does it make sense to calculate the learning happening with rounds between 0° and 180° hue? Even though such a hue data subset would provide enough data to get learning measurements, there is a need to narrow the analysis down in hues. A hue data subset such as *[0°, 10°]* would be highly meaningful but would not provide enough data, as explained earlier. We could settle for 60°-wide data subsets, but should it start with *[0°, 60°]* or *[5°, 65°]*? There is no apparent reason to favor one or the other.

On the x-axis of the OHLC, hues range from 0° to 360°. On the y-axis, each hue (h_c) is mapped to a learning quantity corresponding to the sum of learnings from all players at all rounds within a hue data subset centered in h_c with a given SHR.

The *SHR* (Subset Hue Range) is a variable of pure analysis interest and corresponds to the hue width of the datasubset from which learning is derived. A hue data subset centered in $h_c=50°$ with an SHR of 80° is made of rounds with a target hue comprised between 10° and 90°.

While the Streamgraph proposes a dozen pre-split hue data subsets, the OHLC approach is similar to selecting 360 data subsets, then extracting a single overall learning from each one – rather than a detailed round-based graph – and building a bar chart out of the extracted values.

The following formula defines the OHLC:

$$L(h_c) = \sum_p \sum_i L_i\left(p, h_c - \frac{SHR}{2}, h_c + \frac{SHR}{2}\right)$$

where p is the player, i is the player's i^{th} played round and h_c ranges from 0° to 360°. In Hue Hunt Results, only players having reached level 2 are taken into account for OHLCs.
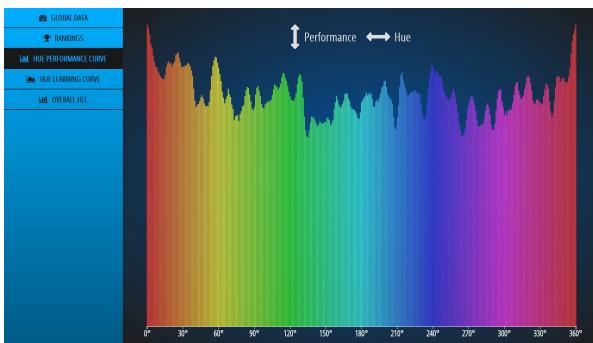
3. Hue Hunt Results UIs

Having built the game and shaped its data outcome, a special interface was needed to visualize and interpret rounds data. Hue Hunt Results comprise five UIs with different purposes.
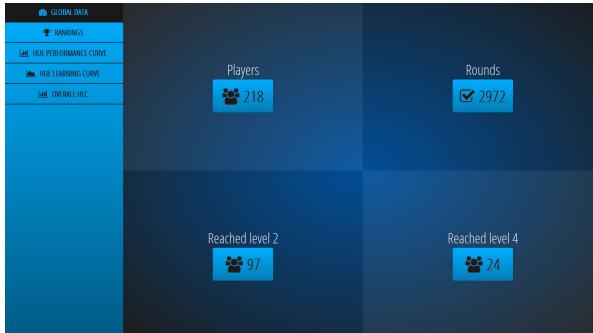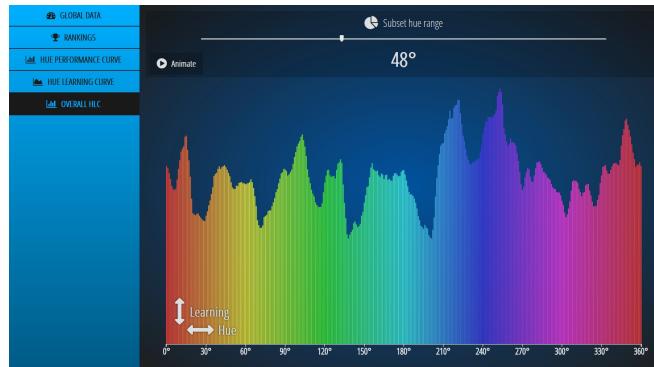

Hue Hunt Results - Global Data

Hue Hunt Results - Rankings

Hue Hunt Results - Hue Performance Curve

Hue Hunt Results - Hue Learning Curve

Hue Hunt Results - Overall Hue Learning Curve

Global Data
Overall statistics are displayed on the *Global Data* interface. One can consult at a glance the total number of players, the total number of rounds played, the number of payers having reached level 2 and the number of players having reached level 4.

Rankings
The win screens of the game include halls of fame, which are leaderboards of the current level where the player can compare their performance to others.
Rakings shows all four halls of fame at once.
It is a way to control what players view and detect any potential technical issue with leaderboards or score upload.

Hue Performance Curve
The project's HPC is displayed on the *Hue Performance Curve* interface. With a simple bar chart representation, visualizing patterns is made easy. Due to randomized target hues, a technique similar to using SHR was employed to cluster rounds around each hue and then average performance; which results in a curve with enough regularity to detect neat gaps and peaks.

Hue Learning Curve
Hue Learning Curve is a data manipulation interface. One can select a player and a hue range. The plotted streamgraph is the result of data filtering and calculation based on the selected parameters.
It is possible to select several parameters per section, in which case the output streamgraph will be a stack of layers, each corresponding to a learning data subset.
The *All players* parameter automatically selects all players and stacks up all resulting layers. The *20° slices* , *40° slices* and *60° slices* buttons respectively generate 18, 9 and 6 layers corresponding to their 20°-wide, 40°-wide and 60°-wide hue data subsets.

There is also a possibility to select *All levels*, or a single game level, resulting in the representation of specific data subsets. An HLC plotting with 40° slices for all players requires approximately 2.500.000 calculations to extract learning from the global dataset.

Overall Hue Learning Curve
Overall HLC displays Hue Hunt's OHLC (see Appendix 2 for explanations on its construction).
As the *SHR* is an analysis variable used to reveal persistent trends on learning across hues, a range slider enables one to input several SHR values.

Moving the SHR by a few degrees can sometimes result in significant changes in the plotted graph. In order to explore the appearance and disappearance of patterns, an *Animation* button starts a sequence that goes through all SHR values one by one, with a minimum of 10° and a maximum of 120°, moving to the next value after 1s.

Due to the definition of the OHLC, calculation times rise with growing SHR. Yet on all tested computers, the animation runs with success. Each OHLC plotting requires an average of 15.000.000 calculations (310.000 per SHR).