How to Arrange Icons on Smartwatches? Investigating the Effect of Hexagonal Layouts on Menu Search

Tassilo Bouwman
HCI MSc Final Project Report 2018
UCL Interaction Centre, University College London
Supervisor: Duncan Brumby

ABSTRACT
With small devices such as smartwatches emerging icons need to fit on smaller displays while users still need to be able to locate icons quickly. A way of arranging circular icons more space-efficiently is the hexagonal layout. This study investigated visual search performance of hexagonal icon arrangements at different density levels. Twenty-four participants performed known-item searches on icon menus in an online experiment using their own smartphone. These menus varied along four dimensions: (1) whether icons were arranged in hexagonal or in a grid layout, (2) the amount of space between icons, (3) the number of colour distractors, and (4) whether the menu stayed the same or changed with every trial. Results show that the effect of layout depends on the icon density and number of colour distractors. Hexagonal layouts perform better than grid layouts when the space between icons is zero and five colour distractors are present. When icons were further apart and one colour distractor was present no difference between layouts was found. Besides, the study confirmed previous findings on the effects of colour distractors and location learning. Participants were faster when fewer icons in the menu had the same colour as the target. With more practice with the same menu participants became increasingly faster finding icons. The findings imply that icons on small devices such as smartwatches should be arranged in a hexagonal layout.

Author Keywords
Visual search; icons; smartwatches; menus; practice

ACM Classification Keywords
H.5.2 User Interfaces: evaluation/methodology, screen design, theory and methods

MSc Contribution Type
Empirical.

INTRODUCTION
With small devices such as smartwatches and alternative display shapes emerging circular icons have become increasingly popular. Greater numbers of square icons are typically arranged in a grid e.g. on smartphones. An alternative way of arranging circular icons is the hexagonal layout (Figure 1). Compared to a grid layout, hexagonal layouts can occupy space more efficiently when using circular icons. This makes them ideal for smaller displays where space is limited. Locating an icon in a hexagonal layout is a classical visual search task. To date, however, there has been no empirical research on visual search in hexagonal icon arrangements. Knowledge on how people search hexagonal icon layouts could help to inform design decisions on arranging circular icons.

Icons have become a pervasive feature of Graphical User Interfaces. Icons are commonly used for representing files, commands or applications. They are faster to recognise and occupy less space than text labels [47]. Icons are used in a variety of places within operating systems. The most prominent application area is the app launcher, a part of the operating system that shows all installed applications. Locating an application within the app launcher is a typical visual search task.

A prominent area where circular icons are employed are smartwatches. In the current software release for the Apple Watch, the app launcher is presented in a hexagonal layout.
Icons on mobile devices such as smartphones are primarily distinguished by the visual features of colour, size and shape. On smartphones and smartwatches many designers employ 2D flat icons. Flat icons consist of simple pictograms or symbols and strong colours [47]. Flat icons are easier to distinguish than detailed icons because they are better at emphasising the primary visual features of colour, size and shape. These features guide the eyes in visual search because they can be more easily perceived when they are away from the point-of-gaze than detailed features [27]. Colour has been shown to have the biggest impact on visual search performance [55]. Although colour has been subject to many visual search experiments [57], very few studies have examined the role of colour in app icons [2] and its interaction with other visual features such as layout.

While visual search features within icons, such as colour, orientation and size, has been studied extensively [27,57], little research has been conducted on icon arrangements. Existing research on icon arrangements focused on icon distance [13,31] and visual grouping [19]. Studies on panel arrangements mainly compared horizontal and vertical icon arrangements [35,41,49]. One study by Chen and Chiang [6] looked at round panel arrangements that are similar to hexagonal layouts. The study indicates that round panel arrangements are less efficient than square arrangements. However, a menu of only eight items was investigated and the round panel differs from a hexagonal arrangement. A research area where hexagonal layouts have been used most is keyboards [46]. Hexagonal layouts have proven to improve pointing efficiency of keyboards because they reduce the travel distance between keys [30,58,59]. However, hexagonal layouts have been only investigated in terms of pointing efficiency in the research area of keyboards. Consequently, there is a notable paucity of empirical research focusing specifically on hexagonal layouts in the domain of visual search.

The aim of this study was to determine whether hexagonal icon layouts affect visual search performance. The study compared search times of hexagonal and grid layouts in a quantitative approach. Data was collected in an online experiment using a smartphone app. To investigate application areas for different screen sizes space between icons was systematically manipulated. Colour distractors and practice were incorporated to determine the effects for different real world search tasks. The key research question was how search performance of hexagonal icon layouts compares to grid layouts. The findings demonstrate when hexagonal layout are superior and confirm existing findings on colour distractors and the location learning effect. The main contributions of this work are: (1) provide new findings on icon arrangements; (2) introduce a novel approach to conducting visual search experiments online.

LITERATURE REVIEW

This section discusses literature related to this study of visual search in icon menus. It introduces the fundamentals of eye movement and visual acuity. The most relevant visual search features are considered before literature on icon arrangements are reviewed in detail. Current applications of hexagonal icon arrangements in keyboards and smartwatches are discussed. The section finishes with an outline of research on online studies.

Eye Movement and Perceptual Span

The eye has a limited acuity that restricts the amount of information that can be assessed in one gaze and forces people to sequentially search a display. Saccade is the continuous movement of the eyes. Saccades are broken into phases where the eye remains relatively still called fixation. Information is only read during a fixation which usually takes 200-300ms. The eye needs to perform frequent saccades to grasp a display because of its limited acuity. The area with a high acuity (fovea) is only 2° of our visual field. The acuity considerably decreases outside the area of the fovea (peripheral vision) because the density of neural cells in the retina is lower at the edges. The retina is the light-sensitive inner layer of the eye that consists of neural cells that are responsible for creating a focused two-dimensional image of the visual world [43]. Although the peripheral vision is less precise than the fovea, some information can still be obtained from this area. The region from which people obtain useful information is called perceptual span. The size of the perceptual span depends on the type of stimulus [45].

Designers can minimise the number of fixations and therefore improve visual search time by employing stimuli with a bigger perceptual span. The size of the perceptual span for a stimulus can be determined by using eye tracking. For example, a study by Näsänen found that four faces arranged in a 2×2 array can be grasped during one fixation [36]. According to [31] a 5×5 array of square icons can be assessed in one fixation regardless of their distance. However, the study did not take visual features of the icons into account. Especially colour is known to be visible further away from the fovea [18], making icons of different colours notably easier to distinguish [27] and therefore require fewer fixations. Fewer fixations are reflected by quicker visual searches. Thus, eye tracking data is not necessary to determine which stimuli is search more efficiently by the eye. The current study aims to investigate the visual search efficiency of different icon layouts. Therefore, no eye tracking data will be collected.

Visual Search Features

Visual search features such as colour, size or shape have been shown to impact search performance. Visual search is required to use many interactive computer systems, e.g., finding an application icon on a smartphone or smartwatch. Visual search has been the subject of many experiments in Human Computer Interaction (HCI). Prominent reviews can be found in [14]. In general, search time increases in a linear fashion with the number of items shown on a display. What research has demonstrated is that visual search is...
guided by certain visual features that make the search more efficient [27]. The features that are most relevant to searching icons will be reviewed in the following sections.

**Colour**

Colour is known to be the visual feature that can be distinguished the furthest away from the fovea [18, 57]. Thus, an icon with a unique colour can be quickly located within an icon menu. However, most icon menus contain multiple icons that share the same colour. Research has investigated the effect: If the target is red and all other items are green, search time will always be the same, regardless of the number of items (parallel search). As the target becomes less distinguishable (share more visual features), people’s attention needs to be deployed over multiple items and the search becomes less efficient (serial search) [57]. Items that share the same visual feature as the target are called distractors. With an increasing number of colour distractors, other visual features such as size and shape become more important [26]. Other search factors such as distance, organisation or the amount of practice have been mainly investigated for text based menus.

**Practice**

Existing research on text based menus suggest that search time decreases with experience with the same menu [1, 8, 12, 51]. Other studies have shown that an alphabetic and semantic organisation of menus is faster than an unordered layout [5, 20, 33, 34]. However, with practice well organised menus have no advantage over unorganised menus [1, 5, 33, 34, 51]. Cockburn et al.’s predictive model explains these findings: Expert users form and rely on spatial memory of menu item location, reducing the task to a Hick-Hyman Law decision task plus a Fitts’ Law pointing task [7]. Both Laws are logarithmic formulas that can be applied to predict people’s performance for an icon selection task. Fitts’ Law predicts the time it takes to rapidly move to a target area based on the distance to the target and the width of the target [15]. Hick-Hyman Law predicts the time required to make a decision based on the number of choices [22, 24].

The formulas of Fitts’ Law and Hick-Hyman Law can only be applied to predict a search task when no visual search takes place. Novice users need to visually search a menu because they have not formed a spatial model yet [7]. During the transition from novice to expert user, both visual search and spacial memory affect search performance. Knowledge on this transition process can help to design interfaces that can be searched more easily. While the interaction between organisation and location learning has been studied extensively [1], it is still unclear how other visual search features such as colour are affected by location learning. Thus, the interaction between location learning and visual search features needs further study.

**Distance**

There are contradictory findings on density as a visual search factor in menus. While Halverson and Hornof’s [21] findings suggest that sparse menus are searched faster, other studies showed opposing results [39, 52]. Tarling and Brumby [52] demonstrated that sparser layouts are searched first, even though they provide worse search performance. Ojanpää et al. [39] reported similar results with sparse layouts being searched more slowly. They also found longer fixations for more dense layouts, suggesting that people adjust their gaze to grasp multiple items at once. Studies on icon density show different results as demonstrated below.

**Icon Arrangements**

Existing research suggests that the way icons are arranged impact search performance. Previous studies on icon arrangements focused on orientation, aesthetics and grouping. According to [48] search performance increases with increasing aesthetics. Visual grouping of icons also improves search times [38]. An elliptical layout used in [50] to search images provided better visual comfort and search performance compared to an unordered, grid and diagonal layout. An eye tracking study on tag clouds found that fixation is strongly concentrated on the central part in circular tag cloud layouts. In contrast, the upper left area in linear tag clouds layouts was more fixated [32]. Overall, the studies from different application areas provide converging evidence that the arrangement of icons affect how people visually search icon menus.

It is unclear how hexagonal icon arrangements affect search performance. A hexagonal layout can be arranged both horizontally and vertically. An example of a horizontal hexagonal layout is the app launcher of the Apple Watch (Figure 1). Studies on panel orientation showed that horizontal arrangements are faster to search than vertical arrangements [35, 41], even with Chinese characters [49] that are read vertically. In contrast, one study suggests that the most efficient layout orientation depends on the language direction. According to this study search performance is better for the Chinese language in vertical panel arrangements [11]. Another study showed that square icon arrangements are faster to search than long vertical or horizontal arrangements [35]. Chen and Chiang [6] investigated round panel arrangements (Figure 2). They used panel arrangements (square, horizontal, vertical, round), icon size (20, 24, 27 mm) and movement distance (33, 47, 67, 74 mm) as independent variables. Two search times were recorded. The search time for finding an icon from the start position and the search time from the first target icon to a second target icon. The results showed that the round arrangement was significantly slower than the horizontal, vertical and square arrangement for both search times. The study indicates that hexagonal layouts may perform worse than grid layouts, since the round layout is similar to a hexagonal layout. A clear conclusion cannot be drawn because the study did not use a full hexagonal layout.
Most researchers agree that too dense icon arrangements impact search performance negatively [9,31,53]. Everett and Byrne’s results suggest that search strategy changes when icon spacing changes [13]. A model for visual search in mixed icon/text displays by Fleetwood and Byrne [16] predicts that smaller spacing improves search performance. However, experimental results show that search time increases when the space between icons is smaller [9,31,53]. Lindberg and Näsiäinen [31] conducted an experiment on graphical icons using real world icons (Figure 3). They employed icon spacing (0, ¼, ½, 1, 2 icons) and icon matrix configuration (2×2, 3×3, 4×4, 5×5, 7×7, 10×10) as independent variables. Participants had to search an icon on the screen while search time and eye movement was recorded. Their results showed that the most efficient space between icons was ½ icon for most set sizes. Even though 0 icon space showed the worst performance across most set sizes, the difference in search time was mostly negligible. In addition, their findings indicate that participants did not like the 0 and ¼ icon space. The 1 icon space was clearly preferred. However, these results were based upon real world icons from over 20 years ago and it is unclear whether their findings still apply to current icons. Compared to the flat icons of today, the icons are harder to distinguish due to their high visual complexity. Besides, Hornof [23] found that people are more careful when other objects are near the target. This indicates that different pointing strategies might impact search results. Overall, these studies highlight the need for further investigations on icon density with current generation designs.

**Keyboards**

Hexagonal key arrangements has been used to improve key proximity in keyboards, for example in [30,58,59]. An extensive list of keyboards can be found in [46]. However, search performance in hexagonal layouts have not been formally studied in the domain of keyboards. The main goal of keyboard researchers is to minimise travel distance between keys. Most keyboard layouts are optimised for everyday use. Only novice users need to perform a visual search on a keyboard. Thus, the limited interest of keyboard researchers in visual search performance may be due to the fact that the effect of visual features seems to diminish with practice [1,5,33,34]. While these findings can likely be applied to icons, there is still need to gather evidence in a study.

**Smartwatches**

A prominent area where circular icons are employed are smartwatches. In the current software release for the Apple Watch, the app launcher (an overview of all installed applications) is presented in a horizontal hexagonal layout (Figure 1). An in vivo study by Pizza et al. [42] investigated the everyday use of the Apple Watch using cameras. While most interactions only lasted a few seconds, people use the app launcher for 40s on average indicating long search times. Another study emphasised that interactions with smartwatches are more brief and occur more frequently [54]. Thus, finding a layout that supports quick interactions by minimising search time is important for smartwatches.
Despite icon size varies within and between operating systems, icon size was not investigated in this study due to limitations of the design.

We hypothesise that search performance of hexagonal layouts is worse than grid layouts. Even though hexagonal icon menus can be found in software for prominent devices such as the Apple Watch to date visual search in hexagonal layouts has not been systematically evaluated. Both an in vivo study on smartwatches and lab experiment on panel arrangements indicate slower search times in hexagonal icon layouts [6,42]. However, both studies do not provide direct evidence for this claim (see Literature Review). Knowledge on how users search hexagonal layouts could validate the design choice to arrange icons in a hexagon.

Distance Between Icons
We compare icon menus with two different distances between icons. For the close distance condition icons were right next to each other, replicating a previous study on icon density [31]. To increase ecological validity icons had the same distance as on current iPhones for the wide distance condition. Based on the literature on icon search we predict that search performance will be better with icons further apart. Literature on text-based menus showed opposing results, suggesting that more densely arranged menus are searched faster. However, we base our predictions on literature on icons because it is more relevant to the current study. Compared to a grid layout, hexagonal layouts increase screen density (the proportion of the screen displaying information) while maintaining the same minimum distance between circular icons as a grid layout. Studies have shown that increased screen density negatively impacts search performance of grid layouts [9,31,53]. It is unknown how these findings apply to hexagonal layouts. Knowledge on how the two icon layouts perform at different density levels could help designers choose the optimal screen density for each icon layout.

Colour Distractors
The study compares two different numbers of colour distractors: one and five. This was done for two reasons: First, we wanted to replicate a real world scenario. Most app launchers of operating systems contain both icons with many colour distractors and icons with a small number of distractors. For example, the default app launchers of Android and iOS contain many icons that are mostly white but only a few icons that are mostly green (Figure 4). Hence, depending on which target icon the user is looking for, there will be either more or fewer distractors with the same colour as the target.

Second, we wanted to investigate if the effects of layout, distance and practice are still present with a small number of distractors. Colour has been shown to overpower all other visual features such as size and shape [26,27,55]. However, the effect of colour distractors on distance and layout has only been explained in theory. Vision psychology demonstrates that people directly gaze at the icon regardless of their position when no colour distractors are present [57]. Thus, we predict that layout and distance will only have a significant effect when five colour distractor are present. In addition, it still remains unclear how the number of distractors affects learning (see below).

Practice
We compare icon menus that change with every new trial (random condition) with icon menus that remain the same (stable condition). These conditions were chosen for two reasons. First, the stable condition was chosen to investigate the effects of layout, distance and colour in a real world scenario. Böhmer and Krüger showed that the position of icons on smartphones and smartwatches does not change frequently [2]. Thus, a stable icon menu represents how user search icons on smartphones every day. Additionally, the search task was repeated every day over a period of three days to investigate long term learning effects.

Second, research on text based menus has shown that users form a spatial model that reduces the search to a pointing and decision task [7] reducing the impact of factors like organisation on search time [1,5,33,34]. However, it remains unclear how other factors are affected by practice. Colour and practice has been researched in isolation [27] making it hard for designers to determine whether practice mitigates the effect of distractors, or vice versa. Thus, comparing stable and random icon menus can explain the interaction between practice and colour distractors. Overall, we predict that the stable menu is searched significantly faster than the random menu, as numerous studies showed [1,8,12,51]. Based on Somberg [51] and Kaptelinin [25] we hypothesise that the learning effect overpowers the effect of colour distractor, distance and layout. Thus, we expect no difference between one and five distractors in the stable condition, whereas in the random condition search time is expected to be faster with one distractor than five distractors with present.

![Figure 4. Default app launcher of iOS 11 (left) and Android Oreo (right).](image)
Mobile App Based Experiment
The experiment was conducted online using a mobile application. The method of using a mobile app was chosen for four reasons. First, empirical evidence can be collected that validate if existing findings in the area of visual search are generalisable to smartphones. To date, visual search experiments has not been conducted on smartphones. Smartphones are currently the primary application area for multiple icons arranged in a grid. Thus, the results have a higher ecological validity (representing a real world use case). Second, by using an app the experiment can be made available to all smartphone users worldwide. As a result, participants can be recruited easier, from more diverse backgrounds and in greater numbers making the results more generalisable. Third, the experiment can be conducted outside a laboratory environment. This increases the ecological validity of the results. Fourth, participants could be prompted to complete another task using native push-notifications. Reminding participants directly on their smartphone increases the chance of the participant noticing the reminder and thus completing the experiment in time. However, this approach has limitations due to lack of control over how participants perform the assigned task. To minimise variability in task completion we focused on providing clear instructions when designing the app as recommended in [19].

METHOD
Participants
Twenty-four participants (7 male) aged between 18 and 46 (M = 26, SD = 6.97) were recruited through the website callforparticipants.com and through social media. Participants signed up to the study by downloading an app from the Apple App Store on their smartphone or tablet and filling in the consent form within the app. Age and gender were self-reported. Participants were incentivised through entry into a raffle with Amazon vouchers worth £20 as prizes. To ensure participants were independent a unique identifier was generated based on the participants iCloud account and associated with the study data. An analysis of the data revealed no duplicate participants. The display size of each participant’s device was automatically logged to analyse if display size affects search time.

Materials
Participants searched icon menus. Each menu consisted of 24 icons that were arranged in a 6×4 layout. All icons had white unique foreground shapes and a solid background colour (Table 1). Six icons were pink, six orange, and six blue. Two icons were green, two purple, and two turquoise. The distribution of colours was the same for all participants across all conditions to avoid confounding effects. In both the random and stable condition the blue and red icons were always in the three top rows or three bottom rows. There was always one orange icon in each row. To maintain the distribution of icon colours in the random condition the icons were shuffled in a controlled way. With every new trial the three top and bottom rows were interchanged and the icons were shuffled within each row.

For each participant the foreground shapes were randomly assigned to a colour to minimise confounding effects of shapes on search time. The position of each icon was the same for all participants in the stable condition. In the random condition the position of each icon changed with every new trial.

Stimuli were presented on a white background with two status bars at the top (Figure 5). The upper status bar showed the current time, cellular connection and battery status. The lower status bar showed the number of trials left to complete and a cancel button that allowed participants to stop the experiment at any point. The icon menu was centred horizontally with space of at least 50 points between the top bar and the top row of icons. The target icon was centred horizontally on the screen with a ‘Next’ button at the bottom. The vertical position of the target icon varied between three positions to vary where participants look when they start searching the menu. The vertical positions were based on the layout that was shown after the target icon. The top position was at the top row of the layout, the bottom position at the bottom row and the centre position between the third and the fourth row. The vertical position of the target icon was randomly chosen for each trial to avoid effects on search time in the stable condition.

The ratio of icon size and distance was picked based on the iPhone's home screen icons for ecological validity. The smallest iPhone's icon size was 60 points with a horizontal distance between each icon of 16 points (15 : 4 ratio). An

---

1 iCloud allows users to synchronise data between devices. Usually all iOS devices of a user are associated with the same iCloud account and is not change frequently.
The icon size of 55 was chosen to fit the hexagonal layout on the smallest iPhone screen. Thus, a distance of 13.33 points was used for the wide distance condition to maintain the 15 : 4 ratio. The distance for the close distance condition was one point. The icon size and distances of each condition were the same for all iPhone display sizes. For all iPads the icon size was increased by quarter of the iPhone's icon size because iPads are usually viewed from a greater distance. Thus, the perceived icon size remained similar to iPhones. The icon size to distance ratio was kept the same for all devices. The icon size and distance for all devices can be seen in Table 2.

A mobile application was developed by the author to conduct the experiment. The application consisted of three sections: dashboard, activities, and profile (Figure 6). The dashboard showed the average search time for each completed activity. The purpose of the dashboard was to motivate participants completing the tasks by allowing them to reflect on themselves. In the activities section, three search tasks, a survey task, and a task to sign up for a reward were shown in a list. A task could be launched by tapping the entry in a list. Only one task was available at a time. The purpose of presenting all tasks was to give participants an overview of what was expected from them. In the profile section, the participant's id, his assigned group, and the name of the study were displayed. In addition, participants could export their consent form and result files and withdraw from the study. The purpose of this section was to provide participants control over the data they share. Due to time constraints, the app was made for iOS devices only.

Apple’s ResearchKit framework was used to implement the experiment, sign-up process, and the graphs visualising the participant’s results. ResearchKit provided a consistent layout for each step of the experiment and allowed collecting search data in the background. The sign-up process was optimised for smartphone touch screens. It included reading and filling in the consent form, answering questions and drawing the signature using fingers. The signed consent form was generated as a PDF file at the end of the sign-up process to be accessed by both the participant and the researcher. ResearchKit allowed the visualisation of participant's results in a design consistent with iOS. The code is publicly available on GitHub\(^2\). A description of the design process can be found on the author’s portfolio\(^3\).

![Image](https://github.com/bouwman/LayoutResearch)

<table>
<thead>
<tr>
<th>Device</th>
<th>Icon Size</th>
<th>Icon Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>all iPhones (4&quot;, 4.8&quot;, 5.5&quot;, 5.8&quot;)</td>
<td>55 points</td>
<td>13.33 points</td>
</tr>
<tr>
<td>all iPads (9.7&quot;, 10.5&quot;, 12.9&quot;)</td>
<td>68.75 points</td>
<td>18.33 points</td>
</tr>
</tbody>
</table>

Table 2. Icon size and distance for each device in the wide distance condition.

\(^2\) [github.com/bouwman/LayoutResearch](https://github.com/bouwman/LayoutResearch)

\(^3\) [www.bouwman.me/layoutresearch.html](https://www.bouwman.me/layoutresearch.html)
Design
The experiment used a 2×2×2×2 (practice × distance × layout × colour distractors) mixed design. Practice and distance were manipulated as a between-subjects factors. For one half of the participants the order of the icons remained stable within each condition whereas for the other half the order of the menu was different with every new trial. The distance between icons was either zero for one half of the participants or at least 13.33 points apart for the other half (see Materials). Layout and colour distractors were manipulated as within-subjects factors. Participants searched icons arranged in a hexagonal layout and a grid layout. The order in which the layouts were searched was counterbalanced across conditions. The colour distractor factor was manipulated by using targets that shared the same colour with either one or five items within the layout.

Procedure
Participants downloaded the app on the Apple App Store, launched the app and filled in the consent form. Participants started the experiment by tapping on the highlighted row in the ‘Activities’ section of the app (Figure 6). An introduction screen informed participants about the task: Participants had to locate a given target icon in a menu as quickly as possible.

At the start of each trial participants were shown the icon they were looking for (Figure 5). After seeing the icon participants started the trial by clicking the ‘Next’ button. This made the target icon and the button disappear by sliding out to the left while the icon menu appeared by sliding in from the right. Participants selected an icon by tapping it using a finger. If an incorrect selection was made, the icon menu disappeared and the icon of the same target reappeared so participants could repeat the trial. If participants selected the correct icon the trial ended and they proceeded to the next trial.

Participants completed 63 trials in total. The first three trials were practice trials. The remaining 60 trials were divided into two blocks. One block of 30 trials contained hexagonal layouts and one block contained grid layouts. The sequence in which the two blocks appeared was counter-balanced between participants. Half of the participants performed all hexagonal trials before the grid trials, and vice versa. An introductory screen of the new layout appear before the first trial of the second block started (Figure 6). Trials in which the participants did not select the correct target on the first attempt were recorded as errors.

Participants searched 13 unique targets. Three targets were only used for practice trials. The main targets were searched six times, three times in each layout. Participants searched the same target after every 10th trial. Targets occurred in the same order and had the same colour for all participants. Targets were ordered in a way so that consecutive targets not appeared next to each other in the menu. This was done to ensure that participants did not learn the location from previous searches which could effect within-subjects conditions. The experiment took approximately five minutes to complete.

Participants were asked to repeated the experiment every day over a period of three days. After participants completed the first experiment a timer indicated when the next experiment was available. The current time and date was saved after the completion of each experiment to calculate the time interval between each attempt of the experiment. The app allowed participants to perform an experiment only once within 18 hours. Participants were reminded via an automated push-notification to perform the next experiment 18 hours after the last completed experiment.

The search time was recorded by the device by calculating the time difference between the appearance of the icon menu and the selection of the icon. The result was saved as one entity together with the participant’s anonymous identifier and the conditions. In addition, target description position, display size and hours since the completion of the last experiment were saved. At the end of the experiment all entities were written into a comma-separated values (CSV) file, encrypted and uploaded to a secure server which could only be accessed by the researchers.

RESULTS
Mixed-design ANOVAs with a significance level of 0.05 were used to determine the significance of effects and to detect potential interactions. The error rates were calculated by dividing the number of errors by the total number of trials. Fifty-one errors (from a total of 1,512) were excluded from the reaction time analysis. Overall, the mean accuracy was 96.6%. Eight exceedingly slow responses greater than 5 seconds were removed. The outliers were three times the interquartile range above the third quartile.

Figure 7 and 8 show mean search time data for each condition. While practice and colour distractors showed a significant main effect, no difference in search time was detected for distance and layout. In the left graph of Figure

![Figure 7. Mean search time of between-subjects factors practice (left) and distance (right). Error bars depict standard errors. *p < .05 **p < .001 ns = not significant.](image)
it can be seen that participants who practiced with the same menu were significantly faster than participants that searched a different menu with every new trial. In contrast, the right graph of Figure 7 shows similar search times for both levels, indicating no significant effect of icon distance.

Similar results can be seen in Figure 8. No significant difference in search time between hexagonal and grid layouts can be seen in the left graph. In contrast, the right graph indicates a notable difference between levels. Participants selected icons significantly faster with one icon sharing the same colour as the target than with five similar targets.

A mixed-design ANOVA revealed a main effect of practice and colour distractors. Participants located icons significantly faster with one colour distractor (M = 1.38, SD = 0.21) than with five colour distractors present (M = 1.54, SD = 0.31), F(1, 20) = 12.51, p < .05, η² = .002. Icon search was significantly faster in a stable menu (M = 1.36, SD = 0.29) than in a menu that changed with every trial (M = 1.56, SD = 0.22), F(1, 20) = 5.47, p < .05, η² = .16. Distance showed no significant main effect, F < 1. The difference between layouts was not significant as well, F(1, 20) = 3.09, p = .09, η² = .01.

The ANOVA showed no significant interactions between all conditions. However, a detailed analysis of the interaction between layout, distance, distractors and practice was performed in order to test the hypotheses of the study.

Layout, Distance and Distractors
We predicted that the effect of layout and distance are dependent on the number of colour distractors; such that layout and distance between icons only affects searches when there are five colour distractors present. In order to test this hypothesis, the data was split into one and five colour distractors. A mixed-design ANOVA was performed to judge the significance of effects. The statistics software did not allow a post-hoc pairwise comparison on a repeated-measures ANOVA, thus requiring to compare means individually.

Separating the analysis of one colour distractor from five colour distractors revealed two-way interaction between distance and layout. For one colour distractor there was no interaction between distance and layout, F < 1. Interestingly, for five colour distractors there was a significant two-way interaction between distance and layout, F(1, 20) = 4.71, p = .04, η² = .02. Thus, the data was further split into close distance and wide distance. The layout had no significant effect in the wide distance condition, F < 1, participants were significantly faster in the close distance condition when searching in a hexagonal layout (M = 1.40, SD = 0.22) than when using the grid layout (M = 1.48, SD = 0.27), F(1, 11) = 10.46, p = .01, η² = 0.03. When icons were further apart there was no significant difference between layouts, F(1, 11) = 0.06, p = .82, η² = .0004.

This effect also showed without distinguishing between colour distractors. The analysis was split into close and wide distance and repeated-measures design ANOVA was performed. For both wide distance and close distance there were no interactions between distractors and layout detected, F’s < 1. Interestingly, when icons were closer together participants were significantly faster when using the hexagonal layout (M = 1.40, SD = 0.22) than when using the grid layout (M = 1.48, SD = 0.27), F(1, 11) = 10.46, p = .01, η² = 0.03. When icons were further apart there was no significant difference between layouts, F(1, 11) = 0.06, p = .82, η² = .0004.
calculated comparing the frequency of positions. No equally distributed a Chi-square test of independence was performed. The results confirmed the hypothesis only partially. Layout and distance showed no significant effects in both conditions. Interestingly, the results indicate that the effect of colour distractor diminishes with practice. The effect of colour distractors was not significant in the stable condition, $F(1, 10) = 4.53, p = .06, \eta^2_p = .07$. However, in the random condition participants were significantly faster with one colour distractor present ($M = 1.48, SD = 0.17$) than with five present ($M = 1.64, SD = 0.23$), $F(1, 10) = 9.81, p = .01, \eta^2_p = .15$. For the stable condition there were no interactions detected, $F's < 1$. The three-way interaction of layout, distance and colour distractors was near the significant level in the random condition, $F(1, 10) = 4.09, p = .07, \eta^2_p = .04$.

**Display Size**

The display size was unevenly distributed across condition because participants used their own devices. Nine participants used a 4" display, seven a 5.8" display, five a 4.8" display, two a 9.7" display and one a 5.5" display. A mixed-design ANOVA with display size as a random effect and a mixed-design ANOVA that does not account for display size were performed. A comparison of the results showed no notable difference. The display size was therefore excluded from the analysis. The distribution of devices with different display sizes can be seen in Figure 10.

**Target Description Position**

The target description was randomly varied between three positions (see Materials). To assess if the positions were equally distributed a Chi-square test of independence was calculated comparing the frequency of positions. No significant interaction was found, $\chi^2(2) = 4.06, p = 0.13$, indicating an equal distribution of positions. Thus, the target description position was incorporated into the analysis as a random factor. This had no significant impact on the results.

**Repeated Experiments**

The study was designed to repeat the experiment every day over a period of three days. However, only 15 participants completed all three experiments (62.25%). Sixteen participants (66.7%) completed two experiments. On average participants completed the next experiment after 36 hours ($SD = 57.29$). The longest time interval between two experiments was 14 days. Thus, the data of the second and third experiment was excluded from the analysis due to missing data and the significant deviation from the original study design.

**Discussion**

This study was set out investigate hexagonal layouts, an unexplored way of arranging circular icons, and to replicate previous studies on visual search factors using smartphones. The experiment confirmed findings on visual search factors colour and practice, while it showed unexpected effects of the interaction between layout and distance. The following paragraphs discuss the most interesting findings by contrasting results with hypotheses, offering explanations and showing how the findings fit into current literature.

**Icon Layout**

Contrary to expectations, results of the study indicate that at a high icon density hexagonal icon arrangements are searched faster than grid layouts. Participants were significantly faster using the hexagonal layout when five colour distractors were present and the distance between icons was zero. Based on [6,42] we expected hexagonal layouts to be slower than grid layouts regardless of the distance. A possible explanation for this finding may be that icons in a hexagonal layout are closer together, thus more icons can be assessed with one fixation. This is supported by the literature on text-based menu search.

Although, the results differ from some published studies on icon search [6,9,31,53], they are broadly consistent with most studies on text-based menus. Tarling and Brumby [52] demonstrated that more dense layouts are searched faster, even though sparser layouts were preferred by participants. Ojanpää et al. [39] reported similar results with more dense layouts being searched faster. They also found longer fixations for more dense layouts, suggesting that people adjust their gaze to grasp multiple items at once. This can be confirmed by Fleetwood and Byrne's [16] model on mixed icon/text displays which predicts that smaller spacing improves search performance. Opposing results in icon literature may have been found because they used visually more complex icons. Lindberg [31], for instance, used system icons of Windows 98 that were hardly distinguishable when arranged right next to each other (see Literature Review). Thus, further research needs to clarify how space between icons affects search performance.
Figure 11 provides a visual explanation for the finding that the effect of layout depends on the distance between icons. It shows all four layout and distance combinations used in the experiment. Similar to Kieras and Hornof’s model [27], black circles indicate the field of view containing the icons that can be assessed with one gaze. The figure reveals that, when there is no space between icons, seven icons can be assessed at the same time in the hexagonal layout while the grid layout only allows five icons to be assessed at the same time. When the icons are further apart, in both layouts only one icon can be assessed at the same time. The collected data reflects this explanation. In the close distance condition participants were faster searching the hexagonal layout, whereas, in the wide distance condition there was no difference between layouts. Thus, hexagonal icon arrangements are searched faster at high icon densities because more adjacent icons can be assessed with one fixation. An eye tracking study should be carried out in order to validate this hypothesis.

In addition, Figure 11 suggests that more densely arranged icon menus have an overall better search performance because more icons are within perceptual span to be assessed at the same time. However, the results do not support this claim. This inconsistency may be due to an increase in pointing time that mitigated the improved visual search time. Hornof [23] found that people are more careful at pointing when other objects are near the target. Thus, it is likely that the reduced space between icons slowed down pointing speed and lead to no significant difference between small and big icon space. A feature study that separates visual search time from pointing time could explain this hypothesis.

The study has been unable to explain why search in the hexagonal icon arrangement of the app launcher of the Apple Watch was found to be very slow [42]. Interestingly, the arrangement of icons is very similar to the layout that performed best in the experiment. Thus, it is likely that their reported finding is not caused by the hexagonal icon layout. A possible explanation for long search time may be the fact that the app launcher is scrollable in all directions. As a result, it may be more difficult for users to determine the absolute position of the icon and remember its location. Further work on the effects of scrolling and location learning is needed to find the reason for long search times of the app launcher on the Apple Watch.

Colour Distractors
Findings on colour distractors supports both our hypothesis and previous findings. The most obvious finding to emerge from the analysis is that icons are found faster when fewer colour distractors are present. These results mirror long standing research in the field of visual search [18,27,57]. People fixate the objects that are similar to the target. When more than one object are similar to the target people need to deploy their attention over multiple objects. Thus, a greater the number of colour distractors requires a greater number of fixations to assess each icon, which leads to longer search time.

Additionally, the analysis of colour distractors showed that the interaction between distance and layout was mostly pronounced when more colour distractors were present. These results are consistent with data obtained in [26,27,55] and with vision psychology [57]. Colour is the visual feature that can be distinguished the furthest away from the fovea [27]. Therefore, people detect icons that have the same colour as the target in their peripheral vision without adjusting their gaze. Consequently, fewer colour distractors mitigate the effect of other visual factors such as layout and icon space because people can directly gaze at the icons regardless of their position.

Location Learning
The current study showed that repeatedly searching the same icon in the same menu leads to faster search time. This result supports previous research on the learning effect [1,8,12,51]. They demonstrated that people directly move their finger to the learned location of the target when the menu remains the same. When the menu changes with every trial, people need to visually search through the menu because they do not know the location of the target. Surprisingly, we found little evidence to support the hypothesis that learning the location of an icon diminishes the effect of layout, distance and colour. No significant difference between distance levels and layout levels was found in both random and stable conditions. Solely colour distractors showed significant differences in the random condition while no significant difference was measured in the stable condition. This observation may support the hypothesis that colour distractors become less significant after the location of an icon was learned [7,25]. However, these data must be interpreted with caution because the difference between significant levels was very small. The lack of significant observations regarding the learning effect
may be attributed to the similarity of icons that inhibited learning (see Limitations). Further research is required to determine how colour distractors, layout and distance are affected by location learning.

Online Study
Conducting a visual search experiment online using a mobile phone application was partially successful. The use of Apple’s ResearchKit framework facilitated the implementation of key study aspects such as the sign-up process and the experiment. It provided a professional design that made the app both easy to use and trustworthy. Data collected by the app was transmitted securely and reliably to the researcher. However, participants did not complete the study as instructed which lead to incomplete data. Even though the experiment took less than five minutes to complete, a third of the participants did not repeat the experiment a second and third time. Other participants allowed far more than one day between each attempt of the experiment. As a result, long term learning effects could not be investigated.

These rather disappointing results could be attributed to insufficient incentive. When signing up for the experiment participants were informed they could win one of five £20 Amazon vouchers. Apparently this did not incentivise participants sufficiently. A potential alternative to monetary incentive is to provide insights to participants about their own behaviour. Wiseman et al. [56] found that people who took part for self-less reasons complete online studies with the same care as monetarily reimbursed participants. Instead of money, participants received insights about their own behaviour, emphasising that learning something about themselves is a strong incentive. The current study employed a graph showing the participant’s search times of the hexagonal and grid layout to provide a non-monetary incentive. A possible issue with this incentive was that for most participants the insights they gained did not facilitate any behaviour change. Providing concrete suggestions on how to arrange icons based on the participant’s results may raise motivation. In addition, participants did not know how their performance compared to other people. A possible solution to this may be to gamify the experiment by adding a leaderboard. This could potentially lead to more focused completion of the task but also increase the chance of people trying manipulate their results.

Another explanation for participants not completing the study is that participants forgot to complete the experiment because an insufficient number of reminders was sent. Participants were notified via a push notification eighteen hours after completing the first experiment. No further reminders were sent if a participant did not complete the experiment. As a result, participants may have forgotten to complete the experiment. Additional reminders paired with better non-monetary incentives are likely to have increased the number of participants completing all experiments. Despite missing data, results from the remaining data successfully replicated previous findings and contribute new findings to the field of visual search.

Implications and Future Work
The findings may help designers decide how to arrange icons on small devices such as smartwatches. The results suggest that the icon layout of app launchers affects search performance when there is little space between icons. Hence, it may be beneficial for search performance to use a hexagonal layout in cases where space is limited and icons need to be close together. One application area that fits these conditions are smartwatches. Studies emphasise that interactions with smartwatches are more brief than on smartphones [42,54]. Therefore, it seems that hexagonal layouts are ideal for smartwatches because they facilitate brief interactions with displays that have limited space. However, this is only the case when icon density is very high. For cases with more space between icons, designers can choose the layout that suits best their requirements. Additionally, it is unclear how the findings apply to smaller displays with smaller icons. To develop a full picture of the effect of hexagonal layout additional studies will be needed that take icon size into account.

Findings on colour could be relevant to both software designers and computer users who seek to optimise search time in app launchers. Böhmer and Krüger [2] show the multitude of ways in which people arrange icons on their smartphone app launcher, e.g. based on relatedness or aesthetics. According to the current study it is beneficial to have fewer icons with the same colour on the same screen. More importantly, however, people should keep the location of the icons the same in order to benefit from the learning effect. As a consequence they can form a spatial memory of the menu and do not need to visually search for an icon. Besides, findings indicate that by keeping the location of the icon the same the number of colour distractors becomes less important. Thus, for experienced users it may not be important how icons are arranged as long as the location remains the same. Future work is needed to assess how location learning is affected by visual features such as colour.

Limitations
Conducting the visual search experiment online may have biased the results in multiple ways. First, participants may have used different fingers to select an icon. For example, some icons are harder to reach when using the thumb compared to using the index finger, leading to longer search times. Second, the use of different devices may have influenced the results. iPhones with a 5.5" display scales content differently because of their higher pixel density. As result, the icons appear slightly bigger than on smaller iPhones models. iPad intentionally had bigger icons as well (see Materials). According to Fitt’s Law pointing speed on these devices was therefore faster which may have resulted in shorter search times. However, an analysis of screen size as random effects showed no significant difference. Third, participants may not have been independent. The iCloud account was use to create a unique identifier for each participant. This account could have been changed to participate in the experiment multiple time. However, this is unlikely because participants did not know how they were uniquely identified. Fourth, participants may have been less
focused because they were not controlled. This could also explain the less pronounced learning effect.

Although observed data confirmed the location learning effect, the effect was not as strong as expected. Searches were only 11.5% faster than in the random condition, while other studies such as [51] showed up to 80% improvement in search performance. There are several possible explanations for this result. First, participants searched the same target six times, whereas previous studies used twelve to twenty repeated searches [1,51]. Therefore, participants had more trials to learn the location, which lead to improved search time. However, according to the power law of practice [37] most learning takes place within the first trials. Thus, this may not be the reason for the weak learning effect. Another cause that could have mitigated the learning effect was cognitive load. Participants searched nine other targets before searching the same target again. As a consequence, the location could not be memorised because participants' short-term memory was occupied with remembering too many targets at the same time. In addition, icons were very similar. Apart from the background colour, icons could only be distinguished by a white shape. According to the von Restorff effect, people tend to remember stimuli better that differ from the rest [40]. The similarity of icons therefore could have inhibited learning the location of an icon. Employing real world icons in future studies could both solve this issue and improve ecological validity. In conclusion, the number of repeated searchers, cognitive load and the similarity of icons may have diminished the effect of learning.

CONCLUSION

The current study was designed to investigate search time of hexagonally arranged icons using a smartphone app. Results show that search time in hexagonal icon layouts depends on the icon density. At a high icon density hexagonal icon layouts are searched faster than grid layouts. At lower icon density levels no difference between layouts was detected. These findings can be explained by the fact that at a higher screen density hexagonal icon layouts allow assessment of more icons at the same time. Additionally, the study confirmed the effects of colour distractors and location learning. Icons are found faster when fewer colour distractors are present. Practicing with the same layout significantly improves overall search performance and is likely to mitigate the effect of colour distractors. The results are based on search data collected using a custom smartphone app that participants obtained from the Apple App Store. Findings imply that designers should employ hexagonal icon layouts on small devices such as smartwatches to safe space and improve search performance. Due to the novelty of the investigation, findings should be interpreted cautiously. Further research needs to validate the effect of layout for different icon densities and sizes. Overall, the study introduces a novel approach to conducting visual search experiments online and contributes new findings on icon arrangements.

ACKNOWLEDGEMENTS

I would like to thank Duncan Brumby for his supervision and advice throughout the project. I would also like to thank all my participants for their time and effort during my study, as well as fellow classmates, family and friends for their continuous support, for which I am truly grateful.

REFERENCES


54. Aku Visuri, Zhanna Sarsenbayeva, Niels van Berkel, Jorge Goncalves, Reza Rawassizadeh, Vassilis Kostakos and Denzil Ferreira. Quantifying Sources and Types of Smartwatch Usage Sessions.


