



**HAL**  
open science

## Interacting with Autostereograms

William Delamare, Junhyeok Kim, Daichi Harada, Pourang Irani, Xiangshi Ren

► **To cite this version:**

William Delamare, Junhyeok Kim, Daichi Harada, Pourang Irani, Xiangshi Ren. Interacting with Autostereograms. MobileHCI '19: 21st International Conference on Human-Computer Interaction with Mobile Devices and Services, 2019, Taipei, Taiwan. pp.1-12, 10.1145/3338286.3340141 . hal-03030557

**HAL Id: hal-03030557**

**<https://hal.science/hal-03030557v1>**

Submitted on 30 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Interacting with Autostereograms

**William Delamare\***  
Kochi University of Technology  
Kochi, Japan  
University of Manitoba  
Winnipeg, Canada  
William.Delamare@acm.org

**Junhyeok Kim**  
University of Waterloo  
Ontario, Canada  
University of Manitoba  
Winnipeg, Canada  
junhyeok.kim@uwaterloo.ca

**Daichi Harada**  
Kochi University of Technology  
Kochi, Japan  
r190362r@gmail.com

**Pourang Irani**  
University of Manitoba  
Winnipeg, Canada  
pourang.irani@cs.umanitoba.ca

**Xiangshi Ren**  
Kochi University of Technology  
Kochi, Japan  
ren.xiangshi@kochi-tech.ac.jp

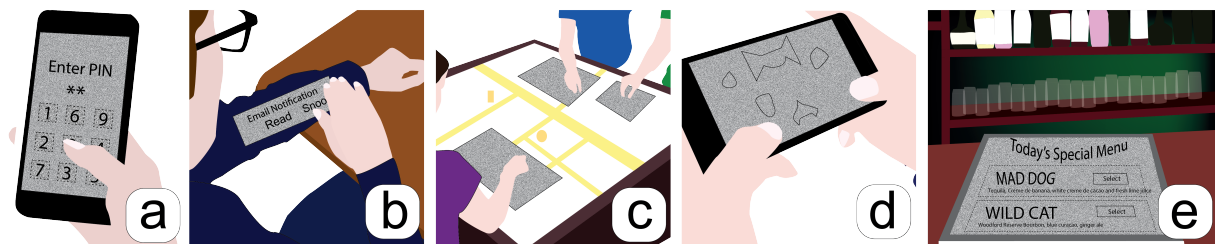


Figure 1: Illustrative examples using autostereograms. a) Password input. b) Wearable e-mail notification. c) Private space in collaborative conditions. d) 3D video game. e) Bar gamified special menu. Black elements represent the hidden 3D scene content.

## ABSTRACT

Autostereograms are 2D images that can reveal 3D content when viewed with a specific eye convergence, without using extra-apparatus. We contribute to autostereogram studies from an HCI perspective. We explore touch inputs and output design options when interacting with autostereograms on smartphones. We found that an interactive help (i.e. to control the autostereogram stereo-separation), a color-based feedback (i.e. highlight of the screen), and a direct touch input can provide support for faster and more accurate interaction than a static help (i.e. static dots indicating the stereo-separation), an animated feedback (i.e., a ‘pressed’ effect), and an indirect input. In addition, results reveal that participants learn to perceive smaller and smaller autostereogram content faster with

\*JSPS International Research Fellows

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*MobileHCI '19, October 1–4, 2019, Taipei, Taiwan*

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6825-4/19/10...\$15.00

<https://doi.org/10.1145/3338286.3340141>

practice. This learning effect transfers across display devices (smartphone to desktop screen).

## CCS CONCEPTS

• **Human-centered computing** → **Graphics input devices; Interaction techniques; Graphical user interfaces;**

## KEYWORDS

Autostereograms; Interactive; Hidden 3D; Magic Eye; SIRDS; Help; Stereopsis; Input; Output; Mobile; Direct; Indirect; Touch

## ACM Reference Format:

William Delamare, Junhyeok Kim, Daichi Harada, Pourang Irani, and Xiangshi Ren. 2019. Interacting with Autostereograms. In *21st International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '19), October 1–4, 2019, Taipei, Taiwan*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3338286.3340141>

## 1 INTRODUCTION

Stereopsis - the ability to perceive depth and 3D stereoscopic images - comes from the horizontal disparity of the two images captured by our eyes [31]. This disparity is then processed by the brain to create a depth perception. Since the development of random-dot stereograms (RDS) by Julesz in 1960 [22], stereograms have been used as a stimulus in several stereopsis experiments to understand human [10, 28, 31] and primate [13, 28] stereo-vision and neurophysiological mechanisms of

depth perception. Autostereograms - or Single Image Random Dot Stereograms (SIRDS) - are 2D 'flat' images that can reveal 3D content when perceived with a specific eye convergence [5, 19, 29]. In contrast to standard stereograms, autostereograms do not require any extra-apparatus such as polarized or shutter glasses.

Autostereograms, also called "Hidden 3D" or "Magic Eye", have also been used as an entertainment medium, such as in image-book collections [12, 32, 46]. Such entertaining and engaging aspect is partially due to the combination of the minimum oculomotor efforts required from the observer, and the surprising resulting 3D effect, making autostereograms an attractive gamified solution to view 3D scenes [42, 43].

However, autostereograms have received little attention from a human-computer interaction perspective. This work fills this gap by introducing the concept of *Interactive Autostereograms*. Indeed, both core concepts defining autostereograms, i.e. "hidden" and "3D", can be useful in various scenarios. For instance, autostereograms can solve the problem of privacy when inputting sensitive information by preventing shoulder-surfing attacks [49]. Yet, no previous attempt was made to interact with autostereograms.

We report on an empirical study investigating **how** users can interact with autostereograms. Results show that an interactive help leads novices to perceive autostereogram content faster than with other help methods. A color-coded feedback provides more accuracy than other feedbacks. A direct input leads to faster and more accurate interaction than an indirect input. Participants can learn to see smaller and smaller autostereogram content faster with practice. Lastly, this learning effect transfers across display apparatus.

By introducing the concept of *Interactive Autostereograms*, our contributions are twofolds: (1) the design of an interactive help method to perceive the hidden 3D content, and (2) an investigation into the impact of different touch input and output design options when interacting with autostereograms.

## 2 RELATED WORK

We briefly present results from studying the stereopsis process in the cognitive literature, and previous works linking computer sciences and autostereograms.

### Stereopsis

Front-facing eyes allow for slightly different captured images, while still allowing for an overlap between the two images. This horizontal disparity allows to derive depth information [21]. Random-dot stereograms (RDS) are pairs of monocular images in which some random dots from one image can be fused with random dots from the other image to create a depth perception [22]. RDSs do not contain any monocular cues such as perspective or shadowing. Thus, RDSs revealed that binocular depth perception was processed by a devoted

part of the brain, not by the eyes [8, 10, 21, 31]. Researchers have then used RDSs in numerous studies to better understand various aspects of stereopsis [14, 17, 19, 23, 30].

Perhaps more relevant to our HCI approach are results reporting a relatively large variability in depth perception with RDSs [12, 19, 21, 27–29]. It appears that around 68% of the population have good to excellent stereo ability, while 32% have poor to moderate stereo ability [21], and 5% of the population is said to be blind to the stereo effect [12]. This can result in a long initial latency to perceive depth [9].

Fortunately, the time to perceive depth and figure in RDSs decreases with repetitions [18, 28]. This learning effect appears to be position specific on the retinae, based on selective spatial attention, and stimulus specific to name a few, but not due to learning depth contours or a priori information about the hidden 3D [28]. Such perceptual learning can last weeks or month depending on the experimental procedure reported in previous work [16, 18, 27, 28]. In addition, learning also concern motion detection in RDSs [18, 34].

Single-image-random-dot-stereograms (SIRDS) are a specific type of RDS that display random dot patterns so that depth can be perceived without additional apparatus [5, 19, 42]. This requires users to explicitly look at the image so that the two eyes capture the dot patterns with the correct disparity to allow the brain to fuse the two images. In other words, users have to decouple eye convergence (crossing-point of the line-of-sight) from focus (depth of lens adjustment). This oculomotor skill is known to be difficult at first [5, 19, 32, 42–47]. Thus, perceiving depth in SIRDSs can take several minutes and several attempts [19]. Possible explanations include the ignorance of the proper viewing strategy [19], or the automatic rejection of the unusual depth impression [47]. However, once the correct decoupling is acquired, the percept is highly compelling [5], with convincing images with vivid depth [43], and can also be modified without any further eye adjustments [5, 32]. Like other RDSs, SIRDS benefit from a perceptual learning, but also from an additional kinesthetic learning of the convergence/focus decoupling oculomotor skill, which is also stable over time [5].

### Autostereograms in Computer Science

Previous work introduced [12, 42], and classified autostereograms [44, 45, 47], encouraging the community to find new scientific applications [12, 42, 47]. Classifications allow us to differentiate autostereograms according to their rendering texture properties (ASCII or random dots, colored or not, etc). Authors also propose an ordering of autostereograms according to how difficult they are to see [44]. However, we could not find any empirical evidence supporting this claim. Previous work also proposed algorithms to generate autostereograms [12, 32, 43, 45, 51]. These algorithms present solutions to create a random dot texture according to a 3D scene via depth map

[43], via ray-tracing [51], or hardware-accelerated via shader programming [32]. Some algorithms also include additional features such as animation support [32, 45], or color inclusion, such as the creation of subtle color effects (e.g. shading) [46], or to render 2D effects on the texture itself [44].

More related to HCI are previous works using eye convergence as an input method [25, 26, 36]. For instance, eye convergence can be used to determine the position a user is looking at in volumetric data to display the corresponding layer as an autostereogram with the correct disparity [36]. A solution to overcome the initial oculomotor skill required to see autostereogram is to rely on gaze tracking [26]. The system can display autostereograms with a disparity according to the current eye convergence. To help users converge their eyes in pre-computed depth levels, authors include physical markers between the desktop screen and the users. Users can then look at the markers' reflection in the screen, hence allowing them to look 'behind' the screen.

In this work, we also tackle the problem of the initial oculomotor skills. However, we consider an interactive approach that could be used in several contexts, not only with gaze tracking and physical markers in-between users and displays. In addition, we investigate input and output solutions when interacting with autostereograms.

### 3 ILLUSTRATIVE SCENARIOS

We provide illustrative examples to demonstrate how interactive autostereograms could find their place in our everyday life. This list complements what has been proposed so far, i.e. using autostereograms for cryptography [50], as an amusement medium [12, 32, 43, 46], for artistic compositions [32, 42], and for scientific visualization [12, 32, 43]. This list is also a response to previous work encouraging the community to further explore the use of autostereograms [12, 42, 47]. Note that these scenarios will be realistically possible once HCI researchers have investigated (i) solutions regarding the initial latency to perceive depth, as well as (ii) the impact of interaction options on performances (e.g., interaction time).

Autostereograms offer **privacy** since the content is only perceived according to a specific viewpoint, with a specific eye convergence. Thus, the content cannot be revealed if the stereo disparity cannot be interpreted by the brain [42]. On mobile devices, a straightforward application of interactive autostereograms concerns password input [26] (Figure 1, a). Autostereograms can prevent shoulder-surfing attacks [33, 49], but also smudge attacks (via oily residues on touch screen) [6] if GUI element positions are randomized at each use. This scenario concerns any access of sensitive information in public, such as checking a bank account balance on public transit for instance. Autostereograms can be perceived with very low resolution [42]. We can hence envision future fabric clothing enabling interactive autostereograms. This

solution could allow fast, easy, and subtle access to email notifications for instance (Figure 1, b). Proprioception offered by fabric input [20, 35] could in turn help for eye-free interactions, such as for quickly snoozing a reminder. Autostereograms can also be a software-oriented solution for accessing private information in a collaborative environment such as on a tabletop for instance [11, 39]. Users could display and interact with private information anywhere on the screen in front of them, without hardware restrictions, and without other users being able to see the content (Figure 1, c).

With some oculomotor efforts, autostereograms offer hidden 3D content with a vivid depth perception [43], making autostereograms a good candidate for **entertainment** purpose (e.g., [2]). On mobile devices, such 3D content could be applied to video games. For instance, a platform game could provide a secret path via the background texture. It could also be used to provide an actual 3D experience while gaming (Figure 1, d), as implemented by the game Quack II [1]. A pub could propose a menu on which special drinks and cocktails could be ordered, hence creating a special and unique experience for its customers (Figure 1, e). In this scenario, the oculomotor efforts and the challenging initial depth perception have the potential to initiate novel social interactions among groups of friends, and hence to positively contribute to the business publicity<sup>1</sup>.

### 4 EXPERIMENTAL FACTORS

We want to explore and evaluate the use of different help methods - to perceive the 3D - as well as different input and output solutions - to interact with autostereograms.

#### Help Methods

We want to provide a solution that could reduce the oculomotor effort required to initiate the interaction. We consider three help methods: **none** (baseline), **static** (standard method), and **interactive** (custom method).

A standard help consists in displaying two solid **static** dots at the top or at the bottom of the image [42, 43, 46, 47]. Users are then instructed to accommodate their eyes so that a third dot appears in the middle. They can then focus on the center dot until depth is perceived in the scene. The system displays two 10-pixels dots on the far plane of the image, positioned at 1/20th from the bottom edge of the screen (Figure 2, left, Appendix A). We created an **interactive** method based on results from previous stereopsis experiments [28]. More specifically, we use the fact that autostereograms can be interactive, and that repeated exposure can facilitate depth perception. We hence provide a slider to control the disparity level (i.e. the separation between corresponding pixels), ranging from 0 (no

<sup>1</sup>Some services exist to create on-demand custom autostereograms for advertisement purpose (e.g., <http://www.magiceye.com/samples.html>, <http://www.custom-stereograms.com/index.html>)

disparity, making the content impossible to see) to 2 (large disparity), with 1 corresponding to the original disparity level [43]. We hypothesize that it can solve the problem of users variability to perceive autostereograms [12, 19, 21, 27–29]. In addition, the system always displays a cube at the center of the scene (Figure 2, left). This can help users anticipate what they can see, and provoke a retinal position-specific learning effect [28] based on the selective spatial attention of the probable stimulus location [28]. The cube is rotated to fully benefit from the depth-within-figure effect [28], stating that depth-within-figure is easier to perceive than the figure contours.

### Touch Inputs

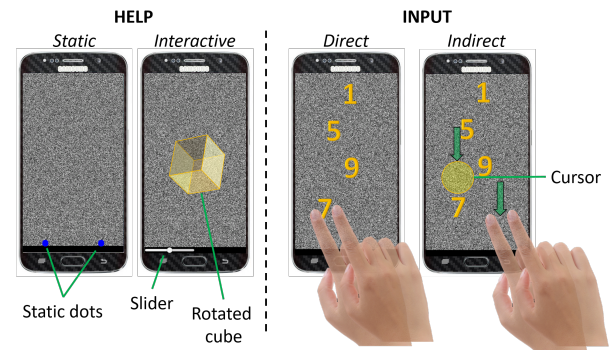
Touch interaction commonly uses **direct** input to provide an affordance close to real world manipulations. However, previous work comparing direct vs indirect touch input reveals that direct touch manipulation can be faster and more accurate during a 2D task [37], while indirect can be more accurate during a 3D task [24]. More closely related to our work, indirect input can lead to a better accuracy during 3D manipulations in a 3D stereoscopic context [40]. This result can be explained by the fact that direct touch aims for the middle of the two projected images [48] due to the Vergence-Accommodation conflict, an effect originating from the difference between the perceived depth and the actual screen depth. In our study, we are also interested in knowing if direct interaction can lead to the loss of the stereo-perception of the 3D scene. Indeed, with autostereograms, the eyes are not focusing on the screen plane, thus creating a double-vision of the finger performing a direct touch (Figure 2, right). If users try to focus on the finger, the stereo-perception of the scene will then be lost.

We hence also include an **indirect** input (Figure 2, right). The system displays a circle as a cursor. The 50-pixels circle, part of the hidden interface, is positioned on top of all 3D elements in the scene. We apply a gain of 2 between the finger movements and the cursor: a 1px finger movement creates a 2px cursor displacement. The cursor cannot go outside the boundary of the screen. Users can select an object by positioning the cursor on top of it, and then performing a tap action. The control of the cursor (slide and tap actions) can be performed wherever on the screen.

### Visual Outputs

Any user interfaces, hidden or not, should offer some feedback after a user action [7, 38]. We are interested in knowing if visual feedback is adapted when interacting with autostereograms, especially since visual feedback can be disruptive for the 3D perception. We consider three visual feedbacks: **none** (baseline), **color**, and **animated**.

For the **color** feedback, the system highlights the full screen when a selection is detected: every pixels displays a random green shade instead of a grey shade. Indeed, it is not



**Figure 2: Help methods (left) and touch inputs (right). Yellow elements represent hidden autostereogram content. Green elements represent picture annotations.**

possible - and perhaps not desirable for privacy reasons - to highlight only the selected 3D object [46]. We discarded the solution of highlighting an on-screen bounding area around the touch position, as it created a double-shaped color box under the interaction focus. For the **animated** feedback, we build on the fact that dynamic stereo motion can be perceived (and learnt) in random-dot-stereograms [18, 34]. We hence consider a motion-in-depth feedback: the depth of the selected object is animated to create a ‘pressed’ effect. The object moves backward (to the far plane) and forth (to its origin), with an animation of 0.05 distance unit every 0.1ms (with the near plane being at 1, and the far plane being at 0). To prevent geometrical limitations (e.g., unfortunate echo due to random pixel color), pixel color should be computed at each frame [43], hence creating a dynamic white noise effect. The animated feedback is then hidden by the dynamic white noise and does not reveal any content when seen normally.

While the color feedback might distract at the screen level, the animated feedback might distract in the 3D scene level. In contrast, the color feedback might prevent distraction by being outside the scope of focus, and the animated feedback might prevent distraction by being in focus.

## 5 EXPERIMENT

We consider two experimental tasks: one to assess interaction aspects, and one to assess the minimum size 3D objects should be to be perceived and recognized.

### Tasks

We chose to have two distinct tasks in order to avoid any interference between the size and the interaction factors.

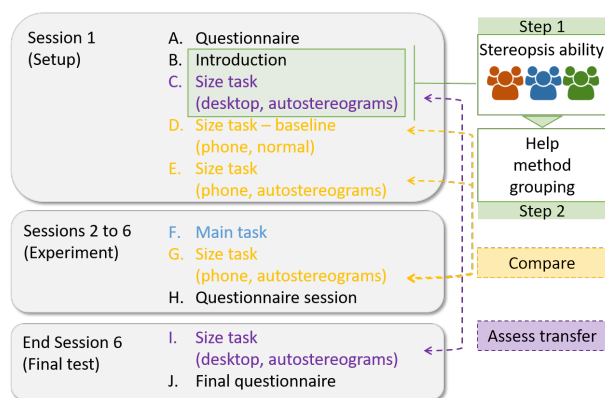
*Main Task - 4 Digits Selection.* The first task consisted in inputting a sequence of four single-digit numbers displayed as autostereograms. Having a sequence of numbers assured that participants could recognize and discriminate the 3D digit shapes. In addition, to capture any potential loss of 3D perception due to our experimental factors (*Input* and *Output*),

the system displayed the two first digits first, and added the remaining two digits after the selection of the second digit. This ensured that participants could not see and memorize the overall 3D scene before starting interacting with it and potentially lose the stereo-perception without us detecting it. Although spotting 3D objects before any interaction in case of stereo-perception loss could likely be the standard in real situations, we wanted to fully control and detect any loss of 3D perception for evaluation purpose. We used a column layout to avoid putting numbers too close to the screen edges. Each number ( $\sim 140 \times 140$  pixels) was slightly shifted, randomly, to the left or to the right from the screen middle vertical line to avoid participants learning finger movements. The numbers, their combinations, and their vertical position were randomly generated at each trial, but we ensured that a trial had only one occurrence of each number. A trial started with a label presenting the random sequence to input. When ready, participants could press a button to make the label disappear and the autostereogram appear. We set a timeout of 1 minute per trial. A trial ended after four correct or incorrect selections (wrong number or background), or when reaching the timeout limit. Participants performed 10 repetitions with each condition, resulting in: 3 Helps  $\times$  3 Inputs  $\times$  2 Outputs  $\times$  5 Sessions  $\times$  10 repetitions  $\times$  12 participants = 10,800 sequences acquisitions.

**Size Task - Single Digit Recognition.** The second task consisted in recognizing single digit numbers displayed as autostereograms without using any help method. Digits became smaller and smaller after each trial. First, the system displayed an autostereogram. Participants had to perceive its 3D content without time limit. When ready, participants could go to the next screen and chose the number they saw (among three other unique random single digit numbers), or the "I do not know" option. If the answer was correct, another autostereogram appeared with a new 3D random number with a scale reduced by 10%. The task finished in case of a wrong answer or if participants declared not knowing since the next trial would use an even smaller size.

### Procedure and Experimental Design

Due to a strong learning effect in perceiving stereograms [5, 18, 28], we chose a mixed-design with HELP as a between-subject factor, and INPUT and OUTPUT as within-subject factors. We counterbalanced INPUT and OUTPUT within groups with a Latin-Square design. We also ran the experiment over multiple SESSIONs to capture learning effects on performances and preferences. Thus, the overall procedure consisted of six sessions: a setup session, and experimental sessions 2, 3, 4, 5, and 6. Sessions were approximately 1h long (hence a total of 6h per participants), and separated by at least 2h and at most 24h from each other. We made sure that participants could take a break whenever they wanted to guarantee



**Figure 3: Experimental sessions.**

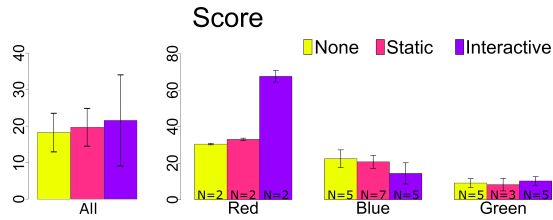
minimum or no fatigue. The 2h to 24h gap between sessions was to allow some rest while still ensuring a learning effect, like in the motor learning literature [4].

**Session 1: Setup.** The first setup session (Figure 3) allowed us to assess participants' ability to perceive autostereograms. Indeed, 30% of the population can find it difficult to perceive stereograms [21], and 5% is completely 'blind' to them [12]. Since HELP methods are evaluated with a between-subject design, we wanted to make sure that our groups were correctly balanced regarding participants' ability to see autostereograms. We assessed participants' ability to perceive autostereograms via an introduction (Figure 3, B) and a size task on a desktop computer (Figure 3, C). We used existing examples from the internet to introduce the concept. During this initial introduction, we instructed participants to try different methods to see their first autostereogram:

- Put the nose on the screen to prevent accommodation, and slowly go back.
- Look behind the screen (e.g., wall behind the display, or the on-screen face reflection [12, 43]).
- Daydream so that eyes diverge and 'look' behind the screen.

We then used the number of correct answers and the average time per trial in the size task on desktop to create three stereopsis ability groups (Figure 3, step 1). We finally assigned each participant to a HELP method group so that each HELP group had approximately the same stereopsis ability levels (Figure 3, step 2). Participants then performed a size task on smartphone without autostereograms (Figure 3, D). We use this condition to have a point of reference for the size task with autostereograms (Figure 3, E and G).

**Session 2 to 6: Experimental Parts.** Participants performed the main task (selection of sequences of digits - Figure 3, F) using the HELP method according to their group defined in Session 1. Participants then performed the size task (Figure 3, G) before finishing the session with a questionnaire to collect qualitative data (Figure 3, H).



**Figure 4: Grouping scores (the lower the better). Error bars represent 95% CI in all graphs.**

*End Session 6: Final Test.* The last experimental session finished with another size task on desktop (Figure 3, I), like the setup session, to assess a potential learning transfer from smartphone to desktop. The experiment ended with a final questionnaire about autostereograms in general (Figure 3, J).

### Participants and Apparatus

36 unpaid volunteers (18 females), ranging in age from 18 to 33 ( $M=23.4$ ,  $SD=3.5$ ) participated in the experiment. Four of them already heard about autostereograms, and nine had previous experience with 3D stereoscopic movies.

Both applications (size and main tasks) were implemented in C# with the Unity3D game engine and ran on a 3.6 GHz Intel Core i7 computer and a Samsung Galaxy S6 smartphone (screen size: 5.1 inches, weight: 138g, resolution:  $1440 \times 2560$  pixels 640 dpi). We based our rendering mechanism on a previous algorithm [43]. However, a C# implementation only lead to an application running at around 4 frame per seconds (fps) on the smartphone. We hence opted for a ray casting method to get the depth map. To fit the proposed algorithm [43], depth is then inverted and normalized so that  $Z = 0$  corresponds to the far plane:  $Z = 1 - (Z - Z_{min}) / Z_{max}$ . We also use two dictionaries: one for rendering purpose, and one for interaction purpose. Thus, one dictionary allows to render the correct shape (e.g., with a hole in the '0' shape), and the second one allows touch interaction (e.g., tapping on the '0' shape selects '0', even if tapping in the hole). The ray cast happens before the scene rendering, thus only creating a delay after pressing the 'start trial' button. The software ran at  $464 \times 824$  pixels (dpi=144), at around 25 fps.

## 6 MAIN TASK RESULTS

We consider three main independent variables: HELP (*None*, *Static*, *Interactive*), INPUT (*Direct*, *Indirect*), and OUTPUT (*None*, *Color*, *Animated*). We refine our findings considering two secondary independent variables: SESSION (1 to 6) to assess learning effects, and TAP (1, 2, 3, 4). Indeed, tap 1 informs on the initial viewing, and tap 3 informs on potential loss of 3D perception during the first two taps.

Instead of listing the impact of every factors per metric (i.e. "Which factors impact metric  $m$ ?"), we present our results according to every metrics per factor (i.e. "Which metrics does factor  $f$  impact?"). This allows us to cast a broad overview of

the complete set of effects the factors have, and hence ease the discussion. We consider 6 dependent variables: the trial completion time (the time to input all 4 digits), the tap time (the time to input one digit), the trial success rate (trials with all 4 digits corrects), the tap success rate (taps on the correct digit), the number of missed targets (taps on the background), and the number of timeouts (trials with less than 4 digits after 60s). We perform our analysis with non-parametric tests (Friedman and Wilcoxon tests for within-subject factors, Kruskal-Wallis and Mann-Whitney tests for between-subject factors, all post-hoc tests reporting Bonferroni corrected p-values) and report averaged values with their 95% confidence intervals (Mean, [CIlow, CIhigh]) and effect sizes.

### Participants Grouping and Rejection Rate

We computed a heuristic metric to create three stereopsis ability groups. The metric is based on the number of correct answers and the average time per trial in the size task on desktop:  $score = (Avg(time)) / 10 - \#_{correct} + 30$ . We divide the average time per trial by 10 to get the same order of magnitude as the number of correct answers, and add 30 in order to have only positive values (post-hoc choice to shift results for presentation purpose). Thus, low scores indicate a relatively good ability to see autostereograms. This heuristic is not meant to become a generalized stereo-ability metric, but is merely a convenient way to categorize our participants without optometric tests [19]. We use color-coded categories to represent participants with poor (Red, score > 30), medium (Blue, score in [10, 30]), and good (Green, score < 10) stereo-ability.

To validate our participants grouping, we analyzed the effect of HELP and GROUP on the average heuristic scores. We did not find any significant main effect of HELP [ $\chi^2(2) = 1.06$ ,  $p=0.59$ ], suggesting a fair global repartition of our participants (Figure 4, left). We found a large effect of GROUP [ $\chi^2(2) = 26.22$ ,  $p < 0.001$ ], suggesting an actual difference between all groups with a large effect size [all  $p < 0.001$ , all  $r > 0.70$ ]. The lack of significant difference between HELP is true for all groups [all  $p > 0.1$ ] (Figure 4, right). Note that there is a low statistical power in the Red group ( $N=6$ , with 2 per help system). Thus, the interactive help has one participant as an outlier, but this does not significantly impact the results. We can conclude that our participants grouping is fair for all HELP systems. Although this heuristic does not include other factors such as the learning ability of participants, it at least ensures that we do not perform a mixed-design analysis with a HELP system ending up with all red participants and another one with all green participants for instance.

Two participants could not do the experiment, and were replaced by new participants to keep a counter-balanced design. This rejection rate ( $2/(36+2)=5.26\%$ ) is in line with the fact that 5% of the population cannot see autostereograms [12], but lower than in previous works ( $2/16=12.5\%$  [27], and

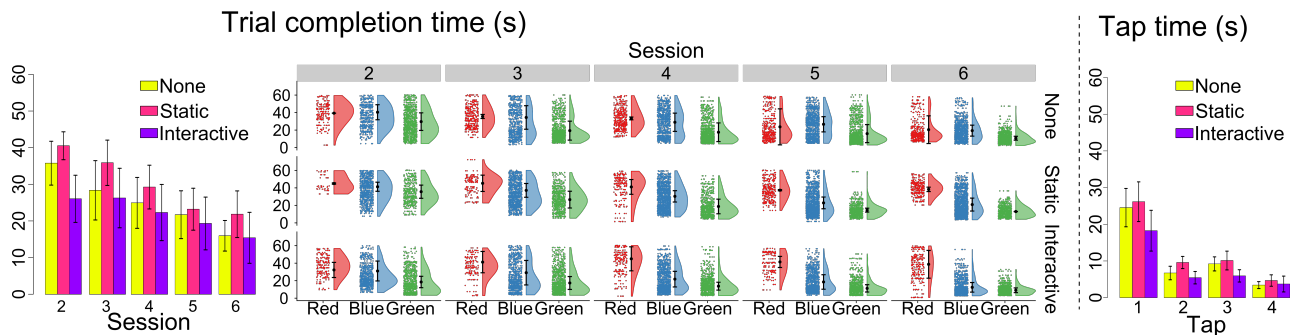


Figure 5: Effect of HELP methods and SESSION on trial completion time (left), and tap time during session 2 (right).

1/3=33% [28]). However, our rejection rate combined to our Red group is in line with these statistics:  $(2+6)/(36+2)=21\%$ . We decided to keep Red participants to provide a comprehensive analysis. In the remainder of this section, to move toward more transparent statistics [15], we provide raincloud plots [3] to display raw data, probability densities, and 95% confidence intervals for each group of participants.

### Help Methods

Help methods influence time-related metrics during the first experimental session (session 2). We found an effect on **trial completion time** (Figure 5, left) [ $\chi^2(2)=10.9$ ,  $p<0.01$ ], with *Interactive* ( $M=26.03s$ , [19.6, 32.5]) leading to faster times than *Static* ( $M=40.49s$ , [36.7, 44.3]) with a large effect size [ $p<0.01$ ,  $r=0.66$ ]. We also found a large effect of HELP on **tap time** during the first experimental session (Figure 5, right) [ $\chi^2(2)=15.9$ ,  $p<0.01$ ], with *Interactive* ( $M=6.4s$ , [5.2, 7.5]) faster than both *Static* ( $M=12.17s$ , [9.9, 14.4]) [ $p<0.01$ ,  $r=0.8$ ] and *None* ( $M=10.23s$ , [8.4, 12.1]) [ $p<0.01$ ,  $r=0.5$ ]. These differences are reflected for TAP 1 [ $\chi^2(2)=8.4$ ,  $p<0.05$ , all  $r>0.46$ ] and *Direct* INPUT [ $\chi^2(2)=13.1$ ,  $p<0.01$ , all  $r>0.5$ ]. As a side note, this effect disappears as participants reach SESSION 6, leading to a TAP 1 time of 4.9s, 6.4s, and 5.0s for *None*, *Static*, and *Interactive* respectively. The only difference between tap time with *Static* ( $M=13.23s$ , [10.6, 15.8]) and *None* ( $M=6.8$ , [4.8, 8.9]) emerges during SESSION 2, with an *Indirect* INPUT [ $\chi^2(2)=10.7$ ,  $p<0.01$ ] ( $r=0.72$ ).

*Discussion.* The *Interactive* help allows for faster trial completion times than the other help systems. Since this gain of time concerns only the first tap, we can conclude that this help concerns the initial depth perception only. However, differences exist only during the first session. This indicates that participants can quickly learn how to see autostereograms, even without help. Help can be beneficial for the first handling of such novel display type. The fact that *None* was faster than *Static* with an *Indirect* input was not expected. The static two dots are positioned at the bottom of the screen, while the cursor is in the center of the screen. This could explain difficulties faced by participants in this condition. Accuracy-related metrics are not impacted

by HELP. Thus, the quality of the depth perception is not impacted by how this percept is acquired in the first place.

### Touch Inputs

*Direct* ( $M=21.4s$ , [17.7, 25.1]) leads to faster **trial completion time** than *Indirect* ( $M=27.9s$ , [24.7, 31.1]) [ $W=30$ ,  $Z=7.4$ ,  $p<0.01$ ,  $r=0.9$ ] (Figure 6, left). This result is reflected in all visual outputs (all  $r>0.8$ ), in all help systems (all  $r>0.8$ ), and in all sessions (all  $r>0.8$ ) - although we note a strong learning effect, going from 31.7s and 36.8s (session 2) to 14.3s and 20.9s (session 6) for *Direct* and *Indirect* respectively. The same effect appears for **tap time**, with *Direct* ( $M=6.0s$ , [5.0, 7.0]) faster than *Indirect* ( $M=8.0$ , [7.0, 9.0]) and a large effect size [ $W=43$ ,  $Z=7.4$ ,  $p<0.01$ ,  $r>0.8$ ]. This result holds for all taps (all  $r>0.6$ ), all visual outputs (all  $r>0.8$ ), all help systems (all  $r>0.8$ ), and all sessions, with a learning effect leading *Direct* from 9s (session 2) to 4.5s (session 6), and *Indirect* from 10.1s to 6.8s. Regarding accuracy metrics, INPUT also influences the **trial success rate** with a large effect size [ $W=544$ ,  $Z=-7.4$ ,  $p<0.01$ ,  $r>0.8$ ]: *Direct* ( $M=57.0\%$ , [49.3, 64.7]) is more accurate than *Indirect* ( $M=44.0\%$ , [33.3, 54.7]), again reflected for all conditions (all  $r>0.8$ ) and a strong learning effect: from 39.8% to 71.3% for *Direct* and 27.2% to 56.3% for *Indirect*. The same general difference exists when considering the **tap success rate** (Figure 6, right), with *Direct* ( $M=79.5\%$ , [75.5, 83.5]) more accurate than *Indirect* ( $M=65.7\%$ , [57.6, 73.8]) [ $W=544$ ,  $Z=-7.4$ ,  $p<0.01$ ,  $r>0.8$ ]. There is also a strong learning effect: from 68.5% and 50.0% (session 2) to 87.2% and 75.6% (session 6) for *Direct* and *Indirect* respectively. It appears that *Direct* ( $M=7.6\%$ , [5.7, 9.6]) leads to less **missed targets** than *Indirect* ( $M=14.7\%$ , [10.1, 19.2]) for all conditions. However, we found a learning effect only for *Direct* [ $\chi^2(4)=30.9$ ,  $p<0.01$ ], going from 12.4% (session 2) to 4.2% (session 6) ( $r>0.4$ ). Lastly, *Direct* ( $M=6.0\%$ , [3.1, 8.9]) also leads to less **timeouts** than *Indirect* ( $M=11.6\%$ , [6.2, 17.0]) [ $W=73$ ,  $Z=-7.4$ ,  $p<0.01$ ,  $r>0.8$ ].

*Discussion.* INPUT influences all metrics, with a direct input leading to faster and more accurate results than an indirect input. The benefits of a direct input are also reflected in the **qualitative results**: participants preferred a direct input (85.1%) compared to an indirect input (14.9%). 14 participants



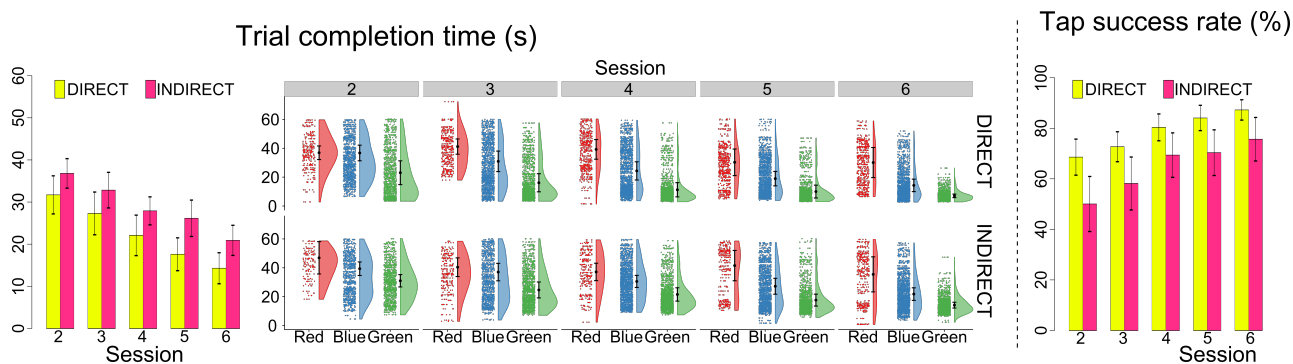


Figure 6: Effect of INPUT and SESSION on trial completion time (left), and tap success rate (right).

commented that the indirect input was cumbersome and felt like a "waste of time" compared to a direct input. In comparison, a direct input is considered as "intuitive" and "fast". Interestingly, 13 participants commented that the stereo cursor was "difficult to find", either at first or after dragging it. Seven participants reported difficulties to perceive both digits and the cursor at the same time. However, one participant noticed that when adjusting the slider (with the *Custom HELP*), this difficulty disappeared. Only four participants noticed that a direct input lead to an "uncertain tap location" because of a "double-vision" of the finger. Thus, current advantages of a direct input (fast and intuitive thanks to an everyday usage) overcome the problem of double-vision we anticipated.

### Visual Outputs

OUTPUT has a small significant effect on **trial success rate** (Figure 7, left) [ $\chi^2(2)=15.4$ ,  $p<0.01$ ], with *Color* ( $M=54.5\%$ , [45.6, 63.5]) leading to more accurate input than *None* ( $M=50.3\%$ , [42.1, 58.5]) ( $r=0.1$ ) or *Animated* ( $M=49.6\%$ , [41.2, 58.1]) ( $r=0.1$ ). We also note a learning effect, as accuracy goes from 39.9%, 39.1%, and 31.2% (session 2) to 60.1%, 66.9%, and 63.7% (session 6) for *None*, *Color*, and *Animated* respectively. These differences are reflected in the *Direct* condition only [ $\chi^2(2)=13.4$ ,  $p<0.01$ , all  $r = 0.1$ ]. A similar pattern appears for the **tap success rate** metric (Figure 7, right) [ $\chi^2(2)=16.4$ ,  $p<0.01$ ], with *Color* ( $M=76.5\%$ , [71.3, 81.6]) marginally more accurate than *None* (73.0%, [67.8, 78.2]) ( $r=0.1$ ) and *Animated* ( $M=72.5\%$ , [67.4, 77.6]) ( $r=0.1$ ). A learning effect takes place

as accuracy goes from 64.2%, 64.5%, and 59.8% (session 2) to 79.8%, 84.1%, and 81.1% (session 6) for *None*, *Color*, and *Animated* respectively. Note that this small difference is reflected for the *Dynamic* help method [ $\chi^2(2)=7.7$ ,  $p<0.05$ , all  $r = 0.1$ ] and the *Direct* input [ $\chi^2(2)=28.7$ ,  $p<0.01$ , all  $r = 0.1$ ].

*Discussion.* OUTPUT has a marginal effect on accuracy-related metrics, with *Color* leading to more accurate selections than *None* and *Animated* - which both lead to equivalent results. This effect is reflected in **qualitative results**, with participants ranking *Color* first more often (64.6%) than *None* (21.7%) and *Animated* (13.7%). Indeed, participants noted that *Color* was easier to perceive, while 21 participants could not see (at all or clearly) the *Animated* feedback. Of these 21 participants, one reported that it might be due to the loss of 3D perception after a tap on the screen, while two mentioned that they faced difficulties only in the *Direct* condition. In addition, four participants could see the animation after some practice and changed their preferences until realizing the delay it imposes on the interaction process. Two participants noticed that the animation impacted their 3D perception (while four reported the same effect with *Color*).

## 7 SIZE TASK RESULTS

The second task was meant to determine the size limit of autostereogram content compared to standard viewing. We consider 3 dependent variables: the time delay to perceive a digit, the total number of digits seen during the task,

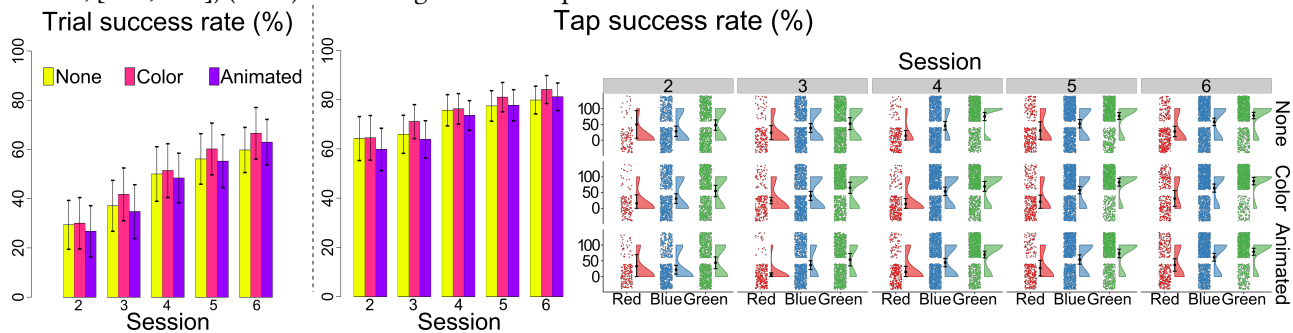


Figure 7: Effect of OUTPUT and SESSION on trial success rate (left), and tap success rate (right).

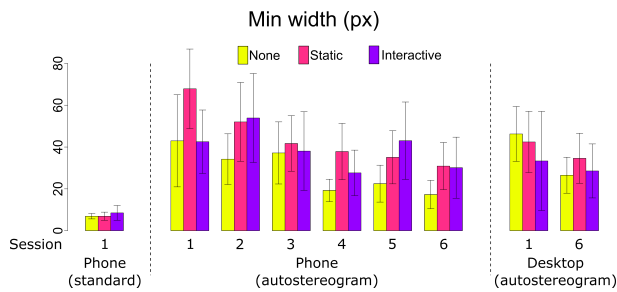


Figure 8: Minimum perceived width.

and the bounding box width of the last digit seen. We consider 2 independent variables: the session, and the help method participants used during the main task (recall that participants performed the size task without help method).

We found a significant large effect of learning on the **time delay**, going from 64.5s, [29.4, 99.5] (session 1 - no practice) to 8.5s, [6.4, 10.7] (session 6) [ $\chi^2(5)=56.9$ ,  $p<0.01$ ,  $r=0.6$ ]. Even after practice, this time is slower than without autostereograms ( $\sim 0.8s$ ). This learning effect is transferred to autostereogram viewing time on desktop [ $W=594$ ,  $Z=7.27$ ,  $p<0.01$ ,  $r=0.9$ ], with session 1 ( $M=33.1s$ , [20.6, 45.7]) slower than session 6 ( $M=9.4s$ , [6.3, 12.6]). We did not find any effect from the help system participants used during the main task [ $\chi^2(2)=1.6$ ,  $p=0.4$ ]. Learning occurs for the **number of digits** participants managed to recognize [ $\chi^2(5)=30.5$ ,  $p<0.01$ ], going from 9.3 [6.6, 12.0] in session 1 to 15.9 [13.7, 18.0] in session 6 ( $r=0.4$ ). For comparison, participants manage to recognize  $\sim 28$  digits without autostereograms. The learning effect is transferred to the desktop condition [ $W=171.5$ ,  $Z=7.3$ ,  $p<0.01$ ,  $r=0.9$ ]. We did not find any effect of the help method on the number of recognized digits [ $\chi^2(2)=3.9$ ,  $p=0.14$ ]. Lastly, we also found a significant effect of learning on the **width limit** (Figure 8) participants could identify (the height limit follows the same trend) [ $\chi^2(5)=29.2$ ,  $p<0.01$ ]: from 50.6px [39.9, 61.3] (session 1) to 26.1px [19.4, 32.8] (session 6) ( $r=0.4$ ). For comparison, the width limit without autostereograms is  $\sim 7.5px$ . The learning effect is transferred to the desktop condition, going from 40.9px [31.1, 50.6] in session 1 to 30.0px [23.7, 36.3] in session 6 [ $W=410$ ,  $Z=7.27$ ,  $p<0.01$ ,  $r=0.9$ ]. We did not find any effect from the help system on this width limit [ $\chi^2(2)=4.0$ ,  $p=0.14$ ].

## 8 LIMITATIONS

Our *Interactive* help was effective during the first experimental session only. Participants still took  $\sim 5s$  to perceive the hidden content during the last session. Although promising, it appears that we can continue our efforts to provide even more fruitful help designs, perhaps focusing on the learning aspect instead of the initial perception of novices. Our investigation is a first step toward the use of autostereograms and includes all groups of users (red, blue green). Future work could be more focused (e.g., with the RED group of

users only). To reach this goal, we provide our experimental material to the community [http://hci.cs.umanitoba.ca/assets/publication\\_files/Autostereo\\_PublicationSoftwareData.zip](http://hci.cs.umanitoba.ca/assets/publication_files/Autostereo_PublicationSoftwareData.zip). Indeed, since we employed a between-subject design, future works can come up with their own help methods, and compare their results to the three help methods we already analyzed. Assuming a fair and similar participants grouping (red, blue, green), future works can evaluate their methods without extra-costs and efforts regarding our baselines. Once interaction challenges identified and solved, prospective scenarios described in this work and new ones could then become common day-to-day interactive medium.

Participants reported difficulties to perceive dynamic elements: the cursor for the indirect touch input, and the ‘pressed’ effect of the animated feedback. We hypothesize that although learning occurs [18, 34], only a subset of our experimental conditions involved these dynamic elements. We hence consider that we did not evaluate the full impact of dynamicity, but rather its effects before and during learning, showing that direct input leads to more time-efficient and accurate selections than an indirect input.

## 9 KEY FINDINGS

**Learning** occurs:

- 1) Participants can learn to see (faster and smaller) autostereogram content, even without help.
- 2) The learning process transfers across apparatus (smartphone and desktop).

**Help** benefits novices time performance:

- 3) Providing help for the initial viewing of autostereograms can reduce the time it takes novices to perceive the hidden content.
- 4) The quality of the depth perception is not impacted by how this percept is acquired in the first place.

**Touch inputs** influence both time and accuracy:

- 5) A direct touch interaction leads to faster and more accurate input than an indirect one.
- 6) Current advantages of a direct input overcome the double-vision of the finger problem.

**Visual outputs** marginally influence accuracy:

- 7) A color-based feedback increases accuracy.
- 8) An animated feedback is difficult to perceive.

## 10 CONCLUSION AND FUTURE WORK

Autostereograms are 2D images that can reveal a 3D content when perceived with a specific eye convergence, without any extra apparatus. Autostereograms have received little attention from an HCI perspective. This work fills this gap with the concept of *Interactive Autostereograms*.

Given the minimum oculomotor-skills required to perceive the content of autostereograms, we first illustrate when such

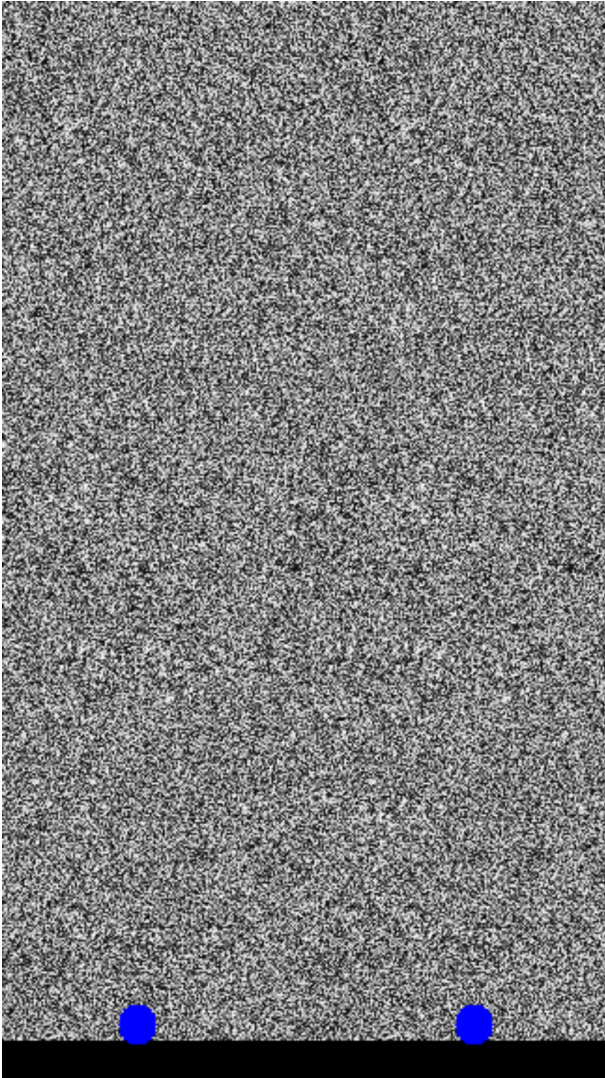


Figure 9: Autostereogram example using the *Static* two dots help method and an *Indirect* touch input.

medium could offer new use-case scenarios in our everyday life. We then perform an empirical investigation regarding how users can interact with autostereograms. We evaluate three different methods to help users perceive the hidden 3D content, two touch-inputs, and three visual feedbacks. Results show that our interactive help allows for faster content perception during the first experimental session. We also provide evidences that a color-based feedback can improve accuracy, while a direct touch input improves both interaction time and accuracy. Lastly, results reveal a strong learning effect, with participants being able to decrease the time it takes to perceive smaller and smaller hidden content. Such learning effect transfers across display apparatus (smartphone to desktop).

Beside pursuing efforts toward decreasing the initial time required to perceive autostereograms, we identify two

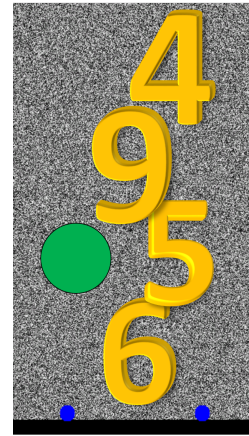


Figure 10: Explicit visualization of the autostereogram example in Figure 9.

research avenues. First, we need to assess if there is some concern about long-term exposure. There is no long term negative or positive impact of 3D stereoscopy display [41]. It might be the same for autostereoscopic display, but a formal experiment is required. Second, if autostereograms become pervasive and common in our everyday life, we might want to investigate if it has an impact at the subconscious level: Can we induce feelings, etc. via autostereograms even if the content is not explicitly perceived?

#### A APPENDIX: AUTOSTEREOGRAM EXAMPLE

Figure 9 shows a screen shot of our experiment main task with an *Indirect* touch input and the *Static* two dots help method. The scene displays four numbers (4, 9, 5, and 6) and a circle representing the cursor. Figure 10 shows a visualization of what readers should see in 3D in Figure 9.

In order to perceive the 3D content using the *Static* help method, readers should **converge** their eyes so that the two solid dots appears as four dots. To do so, slightly converge your eyes as if looking behind the image. Then, readers should accommodate their eyes so that the two images of the dots at the center merge to become one, hence creating only three dot images. The next step consists in controlling the **focus**. The difficulty resides in controlling the focus (i.e. in the image) without losing the correct convergence (i.e. behind the image) obtained in the previous step. For novices, a simple trick is to continue fixating the third central dot image steadily until the brain captures the stereoscopic image. Since the '6' number will be just above this third dot image, readers can slightly move their eyes into the scene while keeping track of their convergence (the third central dot) via their peripheral vision.

#### ACKNOWLEDGMENTS

We acknowledge the support from the Canada Research Chair program, as well as the JSPS fellowship program.

## REFERENCES

- [1] 2018. Quake II AbSIRD. <http://www.leweyg.com/download/SIRD/q2/>
- [2] 2019. IWISTER - the Interactive Wireframe Stereogram. <http://andygiger.com/iwister/>
- [3] Micah Allen, Davide Poggiali, Kirstie Whitaker, and Tom Rhys Marshall. 2018. Raincloud plots: a multi-platform tool for robust data visualization. *PeerJ Preprints* 6 (2018), e27137v1. <https://doi.org/10.7287/peerj.preprints.27137v1>
- [4] Fraser Anderson and Walter F Bischof. 2013. Learning and performance with gesture guides. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, New York, New York, USA, 1109–1118. <https://doi.org/10.1145/2470654.2466143>
- [5] Stephen R. Arnott and Judith M. Shedden. 2000. Attention switching in depth using random-dot autostereograms: Attention gradient asymmetries. *Perception and Psychophysics* 62, 7 (2000), 1459–1473. <https://doi.org/10.3758/BF03212146>
- [6] Adam J Aviv, Katherine Gibson, Evan Mossop, Matt Blaze, and Jonathan M Smith. 2010. Smudge Attacks on Smartphone Touch Screens. *USENIX conference on Offensive technologies* 10 (2010), 1–7.
- [7] J M Christian Bastien and Dominique L Scapin. 1993. *Ergonomic criteria for the evaluation of human-computer interfaces*. Technical Report RT-0156. INRIA. 79 pages. <https://hal.inria.fr/inria-00070012>
- [8] Stephen Becker, Christopher Bowd, Stephanie Shorter, Kaylin King, and Robert Patterson. 1999. Occlusion contributes to temporal processing differences between crossed and uncrossed stereopsis in random-dot displays. *Vision Research* 39, 2 (1999), 331–339. [https://doi.org/10.1016/S0042-6989\(98\)00110-2](https://doi.org/10.1016/S0042-6989(98)00110-2)
- [9] Mark F. Bradshaw, Brian J. Rogers, and Bart De Bruyn. 1995. Perceptual Latency and Complex Random-Dot Stereograms. *Perception* 24, 7 (jul 1995), 749–759. <https://doi.org/10.1068/p240749>
- [10] a Burgess. 2003. Changes in neural complexity during the perception of 3D images using random dot stereograms. *International Journal of Psychophysiology* 48, 1 (2003), 35–42. [https://doi.org/10.1016/S0167-8760\(03\)00002-3](https://doi.org/10.1016/S0167-8760(03)00002-3)
- [11] Li Wei Chan, Ting Ting Hu, Jin Yao Lin, Yi Ping Hung, and Jane Hsu. 2008. On top of tabletop: A virtual touch panel display. In *2008 IEEE International Workshop on Horizontal Interactive Human Computer System, TABLETOP 2008*. IEEE, 169–176. <https://doi.org/10.1109/TABLETOP.2008.4660201>
- [12] Jean François Colonna. 1995. Are autostereograms useful for computer graphics and scientific visualization. *The Visual Computer* 11, 7 (1995), 390–396. <https://doi.org/10.1007/BF01909879>
- [13] M.L.J. Crawford, G.K. von Noorden, R.S. Harwerth, and E.L. Smith. 1996. Judgments by monkeys of apparent depth in dynamic random-dot stereograms. *Behavioural Brain Research* 79, 1-2 (sep 1996), 219–225. [https://doi.org/10.1016/0166-4328\(96\)00036-8](https://doi.org/10.1016/0166-4328(96)00036-8)
- [14] Helen Davis, David Buckley, Ruth E.G. Jacobs, David A.A. Brennan, and John P. Frisby. 2002. Accommodation to large disparity stereograms. *Journal of American Association for Pediatric Ophthalmology and Strabismus* 6, 6 (dec 2002), 377–384. <https://doi.org/10.1067/mpa.2002.127916>
- [15] Pierre Dragicevic. 2016. Fair Statistical Communication in HCI. In *Modern Statistical Methods for HCI*. Springer, 291–330. [https://doi.org/10.1007/978-3-319-26633-6\\_13](https://doi.org/10.1007/978-3-319-26633-6_13)
- [16] John P Frisby and Jeremy L Clatworthy. 1975. Learning to See Complex Random-Dot Stereograms. *Perception* 4, 2 (jun 1975), 173–178. <https://doi.org/10.1068/p040173>
- [17] Liat Gantz and Harold E. Bedell. 2011. Variation of stereothreshold with random-dot stereogram density. *Optometry and Vision Science* 88, 9 (2011), 1066–1071. <https://doi.org/10.1097/OPX.0b013e3182217487> arXiv:NIHMS150003
- [18] Liat Gantz, Saamil S. Patel, Susana T.L. Chung, and Ronald S. Harwerth. 2007. Mechanisms of perceptual learning of depth discrimination in random dot stereograms. *Vision Research* 47, 16 (jul 2007), 2170–2178. <https://doi.org/10.1016/j.visres.2007.04.014>
- [19] Aurora Torrents Gómez, Núria Lupón, Genís Cardona, and J. Antonio Aznar-Casanova. 2012. Visual mechanisms governing the perception of auto-stereograms. *Clinical and Experimental Optometry* 95, 2 (mar 2012), 146–152. <https://doi.org/10.1111/j.1444-0938.2011.00664.x>
- [20] Florian Heller, Stefan Ivanov, Chat Wacharamanotham, and Jan Borchers. 2014. FabriTouch: Exploring Flexible Touch Input on Textiles. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers - ISWC '14*. ACM, ACM Press, New York, New York, USA, 59–62. <https://doi.org/10.1145/2634317.2634345>
- [21] Robert F Hess, Long To, Jiawei Zhou, Guangyu Wang, and Jeremy R Cooperstock. 2015. Stereo Vision: The Haves and Have-Nots. *i-Perception* 6, 3 (jun 2015), 204166951559302. <https://doi.org/10.1177/2041669515593028>
- [22] Bela Julesz. 1960. Binocular Depth Perception of Computer-Generated Patterns. *Bell System Technical Journal* 39, 5 (1960), 1125–1162. <https://doi.org/10.1002/j.1538-7305.1960.tb03954.x>
- [23] Frederick A.A. Kingdom, David R. Simmons, and Stéphane Rainville. 1999. On the apparent collapse of stereopsis in random-dot stereograms at isoluminance. *Vision Research* 39, 12 (1999), 2127–2141. [https://doi.org/10.1016/S0042-6989\(98\)00257-0](https://doi.org/10.1016/S0042-6989(98)00257-0)
- [24] Sebastian Knoedel and Martin Hachet. 2011. Multi-touch RST in 2D and 3D Spaces : Studying the Impact of Directness on User Performance. In *3D User Interfaces (3DUI), 2011 IEEE Symposium on*. IEEE, 75–78.
- [25] Shinya Kudo, Hiroyuki Okabe, Taku Hachisu, Michi Sato, Shogo Fukushima, and Hiroyuki Kajimoto. 2013. Input method using divergence eye movement. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems - CHI EA '13*. ACM, ACM Press, New York, New York, USA, 1335. <https://doi.org/10.1145/2468356.2468594>
- [26] Shinya Kudo, Ryuta Okazaki, Taku Hachisu, Michi Sato, and Hiroyuki Kajimoto. 2015. Personally supported dynamic random dot stereogram by measuring binocular parallax. In *Proceedings of the 6th Augmented Human International Conference on - AH '15*. ACM, ACM Press, New York, New York, USA, 195–196. <https://doi.org/10.1145/2735711.2735806>
- [27] Peter J. Maccracken, Joan A. Bourne, and William N Hayes. 1977. Experience and Latency to Achieve Stereopsis: A Replication. *Perceptual and Motor Skills* 45, 1 (aug 1977), 261–262. <https://doi.org/10.2466/pms.1977.45.1.261>
- [28] Alice J O'toole and Daniel J Kersten. 1992. Learning to see random-dot stereograms. *Perception* 21, 2 (1992), 227–243.
- [29] Alice J. O'Toole and Candice L. Walker. 1997. On the pre-attentive accessibility of stereoscopic disparity: Evidence from visual search. *Perception and Psychophysics* 59, 2 (1997), 202–218. <https://doi.org/10.3758/BF03211889>
- [30] SA Palmisano, RS Allison, and IP Howard. 2001. Effects of Decorrelation on Stereoscopic Surface Perception with Static and Dynamic Random-dot Stereograms. *Australian Journal of Psychology* 46, 1-2 (2001), 57–71.
- [31] Robert Patterson and Wayne L. Martin. 1992. Human Stereopsis. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 34, 6 (dec 1992), 669–692. <https://doi.org/10.1177/001872089203400603>
- [32] Christoph Petz, Bastian Goldluecke, and Marcus Magnor. 2003. Hardware-accelerated autostereogram rendering for interactive 3D visualization. In *Stereoscopic Displays and Virtual Reality Systems X*, Andrew J. Woods, Mark T. Bolas, John O. Merritt, and Stephen A. Benton (Eds.), Vol. 5006. International Society for Optics and Photonics, 359–366. <https://doi.org/10.1117/12.474120>
- [33] Vijay Rajanna, Adil Hamid Malla, Rahul Ashok Bhagat, and Tracy Hammond. 2018. DyGazePass: A gaze gesture-based dynamic authentication system to counter shoulder surfing and video analysis

- attacks. In *2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, Vol. 2018-Janua. IEEE, IEEE, 1–8. <https://doi.org/10.1109/ISBA.2018.8311458>
- [34] D Regan, C J Erkelens, and H Collewijn. 1986. Necessary conditions for the perception of motion in depth. *Invest Ophthalmol Vis Sci* 27, 4 (1986), 584–597.
- [35] T. Scott Saponas, Chris Harrison, and Hrvoje Benko. 2011. PocketTouch: through-fabric capacitive touch input. In *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11*. ACM, ACM Press, New York, New York, USA, 303–308. <https://doi.org/10.1145/2047196.2047235>
- [36] Michi Sato and Hiroyuki Kajimoto. 2012. Dynamic stereograms based on eye convergence for displaying multilayered images. In *SIGGRAPH Asia 2012 Emerging Technologies on - SA '12*. ACM, ACM Press, New York, New York, USA, 1–2. <https://doi.org/10.1145/2407707.2407714>
- [37] Dominik Schmidt, Florian Block, and Hans Gellersen. 2009. A Comparison of Direct and Indirect Multi-touch Input for Large Surfaces. In *IFIP Conference on Human-Computer Interaction*. Vol. 5726 LNCS. 582–594. [https://doi.org/10.1007/978-3-642-03655-2\\_65](https://doi.org/10.1007/978-3-642-03655-2_65)
- [38] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, and Steven Jacobs. 2009. *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (5th ed.). Addison-Wesley Publishing Company, USA.
- [39] Garth B. D. Shoemaker and Kori M. Inkpen. 2001. Single Display Privacyware: Augmenting Public Displays with Private Information. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '01*, Vol. 3. ACM, ACM Press, New York, New York, USA, 522–529. <https://doi.org/10.1145/365024.365349>
- [40] Adalberto L. Simeone and Hans Gellerseny. 2015. Comparing indirect and direct touch in a stereoscopic interaction task. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, IEEE, 105–108. <https://doi.org/10.1109/3DUI.2015.7131733>
- [41] Debi Stransky, Laurie M. Wilcox, and Robert S. Allison. 2014. Effects of Long-Term Exposure on Sensitivity and Comfort with Stereoscopic Displays. *ACM Transactions on Applied Perception* 11, 1 (2014), 1–13. <https://doi.org/10.1145/2536810>
- [42] Harold Thimbleby. 1997. 3D creativity for the Web. *The Internet in 3D* 3 (1997), 41–53. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.35.9945&rep=rep1&type=pdf>
- [43] H.W. Thimbleby, S Inglis, and I.H. Witten. 1994. Displaying 3D images: algorithms for single-image random-dot stereograms. *Computer* 27, 10 (oct 1994), 38–48. <https://doi.org/10.1109/2.318576>
- [44] Thomas Tonnhofer and M Eduard Gröller. 1996. Autostereograms - Classification and Experimental Investigations. In *12th Spring Conference on Computer Graphics*, Vol. 5. Citeseer, 133–144.
- [45] Son Tran Minh, Kalman Fazekas, and Andras Gschwindt. 2002. The presentation of three-dimensional objects with single image stereogram. *IEEE Transactions on Instrumentation and Measurement* 51, 5 (2002), 955–961. <https://doi.org/10.1109/TIM.2002.806035>
- [46] Yusuke Tsuda, Yonghao Yue, and Tomoyuki Nishita. 2008. Construction of Autostereograms Taking into Account Object Colors and Its Applications for Steganography. In *2008 International Conference on Cyberworlds*. IEEE, IEEE, 16–23. <https://doi.org/10.1109/CW.2008.86>
- [47] Christopher W Tyler and Maureen B Clarke. 1990. Autostereogram. In *Stereoscopic displays and applications*, Vol. 1256. International Society for Optics and Photonics, 182–198.
- [48] Dimitar Valkov, Frank Steinicke, Gerd Bruder, and Klaus Hinrichs. 2011. 2D Touching of 3D Stereoscopic Objects. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. ACM, ACM Press, New York, New York, USA, 1353–1362. <https://doi.org/10.1145/1978942.1979142>
- [49] Susan Wiedenbeck, Jim Waters, Leonardo Sobrado, and Jean-Camille Birget. 2006. Design and evaluation of a shoulder-surfing resistant graphical password scheme. In *Proceedings of the working conference on Advanced visual interfaces - AVI '06*. ACM, ACM Press, New York, New York, USA, 177–184. <https://doi.org/10.1145/1133265.1133303>
- [50] Feng Yi, Daoshun Wang, and Yiqi Dai. 2006. Visual Secret Sharing Scheme with Autostereogram. *IACR Cryptology ePrint Archive* 2006 (2006), 295.
- [51] Billy T. Yu and William W. Yu. 1997. Single-image random-dot stereogram by crosstalk. In *SPIE Vision Geometry IV*, Robert A. Melter, Angela Y. Wu, and Longin J. Latecki (Eds.), Vol. 3168. International Society for Optics and Photonics, 115–122. <https://doi.org/10.1117/12.279656>