Cue Impact On Visual Search Experiment

RUDY ZUPETZ, Colorado State University, USA
JACOB FAULKNER, Colorado State University, USA
JON EISENHOUR, Colorado State University, USA
AYDEN ADAIR, Colorado State University, USA

Visual and audio cues can increase the accuracy and speed in which users are able to interact with their computer. This paper analyzes the techniques used in previous experiments, and how they compare to our findings with audio and visual cues used in an "object search" game created in Unity.

CCS Concepts: • **Human Computer Interaction → Search**; • **Visual and Auditory cues**;

Additional Key Words and Phrases: Human Computer Interaction, Visual Search, 2D Environments, Visual and Audio Cues

ACM Reference Format:
Rudy Zupetz, Jacob Faulkner, Jon Eisenhour, and Ayden Adair. 2023. Cue Impact On Visual Search Experiment. 1, 1 (May 2023), 14 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Traversing through the maze of icons and objects in a UI can be a daunting experience for users unfamiliar with the software they're navigating. An overflow of information on the screen can lead to users' efficiency dropping when the visual workload becomes too high [4]. Increasing ease of use and providing the best interaction is a goal of every developer when creating their software, and in our experiment we aim to see how cues can impact this. One idea is to introduce spatial audio cues to help divide the workload from an individuals eyes to their other senses when the visual stimuli becomes too overwhelming. Another is the idea that a soft visual guidance around what users are looking for could help to act as a beacon in an array of clutter. Why audio cues? Primitively, hearing guides our vision towards any potential threats or potential prey [9]. There are many studies relating to visual and audio cues alone, but fewer attempting to connect the two. The purpose of the experiment is to test if audio cues, visual cues, or a combination of the two, can significantly reduce search times in a 2-D UI setting. The participants were tasked with searching through a three by seven grid of various shapes and images, looking for a specified shape in a 2-D environment. This search was completed four times by participants with no cues, visual cues, audio cues, and a combination of both visual and audio cues. Our object was to ultimately measure how said cues impacted their search time. The visual cue was a soft, transparent outline surrounding the target image, the audio cue had left, right, or center direction depending on the location of the shape in the grid.

Authors' addresses: Rudy Zupetz, zupetzrt@colostate.edu, Colorado State University, Fort Collins, Colorado, USA; Jacob Faulkner, Colorado State University, Fort Collins, Colorado, USA; Jon Eisenhour, Colorado State University, Fort Collins, Colorado, USA; Ayden Adair, Colorado State University, Fort Collins, Colorado, USA.

2 RELATED WORKS

Visual search is a constant feature of human life, from searching a fridge for ingredients to searching for an app on your phone. The human eye takes in a lot of information, and the brain must perform a search through this data to comb for importance and drive focus which can take a lot of dedicated brain power [16, 23, 26]. Visual search is universal for any animal with eyes from a survival standpoint. The ability to take in a lot of visual information and find the most important factors is critical to animals in the wild, just as it is to humans completing everyday tasks [11]. Different targets call for different levels of focus at times scrutiny is needed while others only require a visual pass over to complete a search [5].

Research has yielded significant evidence that cues can improve the time it takes for users to find targets in a specific environment. With research of cues going back as far as the 1960's the concept has evolved greatly and has been tested on a variety of devices and displays for a variety of tasks including: augmented reality, virtual reality, 2-D interfaces, and 3-D interfaces [1, 2, 12, 19, 20, 29]. There is also historical differentiation between types of visual and auditory cues. This includes dynamic versus static visual cues, simple static auditory cues vs complex or 3-D auditory cues, and directly mapped visual cues vs indirectly mapped visual cues [2, 20].

Since the 1960's, audio cues have been explored for assisting visual search in operating systems, and using static audio cues reduced the search time from 18.15s to just 10.49s [20]. A study conducted in 2008 by Van der Burg reaffirms this claim Subjects were given a non-spatial auditory signal to aid them in visual search, and it reduced search time drastically [8, 25, 35, 39]. Additionally, a study in 1997 found that audio cues can help visually impaired users more easily navigate 3-D environments [21, 36]. Research in 1989 tested auditory cues against visual cues and found that complex audio cues could replace traditional visual cues. Additionally, In cases where the visual field is already very full of stimulus it has been found that auditory cues are very valuable and are more useful than additional visual cues [27, 31]. This finding was especially accurate in tests where speed was not crucial with the two methods ranking similarly otherwise [4]. It has been found that the right audio can simplify complex images in a persons mind [7]. Audio cues were also shown to improve spatial neglect from a study in 2019 [24]. Another study supports this idea suggesting awareness cues can help the user map an environment around them [37]. Audio cues have been analyzed in a variety of ways from using user facing speakers, to in a car, to setting up full room 3d audio [3, 4, 15**?**]. In 1986, different audio cues were used to signify different notifications, and the familiarity was shown to significantly reduce search times [14]. Visual stimuli also were shown to provide significant improvements in search with even just a small faint light in the general area of the target [10]. Cues can range in their complexity as well as how obvious or subtle they are, in augmented reality devices it was found that subtle transparent cues could function as well as more obvious cues. [19] When comparing the use of visual and auditory cues in visual search there has been a mix of results for what is most effective though it is often found that both are significantly better than unaided search [13, 18]. One study comparing multiple visual and multiple auditory cue types found that visual cues were slightly stronger [2]. Another study found that visual cues lower mental, physical, and temporal demand, as well as improving performance overall and lowering effort required and frustration generated when compared to auditory cues [2].

Multi-modality or the mixing of visual and auditory cues has been found to in cases have a strong effect towards helping in search [6, 10, 29, 30, 38]. Compared to either of the cues by themselves multi-modality provides a best of both worlds solution that covers different types of people [17, 28]. Multi-modality has been found to lead to lessen frustration, mental demand, and time to complete tasks though the difference between visual and visual with audio is not significant for those categories [2]. With this being said, audio cues are still sufficient to work on their own without the assistance of any visual component [34].

3 METHODOLOGY

In our study, we performed within-subjects testing with three or more conditions. We had all of our participants try out four different search scenarios. Each scenario consisted of a two by seven grid of randomly organized shapes and symbols, one of the shapes would be chosen as the “target” shape that the participants would need to click on. Upon entering each scenario, participants would be informed of the target shape and be prompted to start the search. Once the search began, a timer would start counting down until the target shape was clicked on ten times. Every time a participant clicked on a shape the grid would re-randomize and participants would need to begin their search again. Only if the target shape was clicked could the counter increment, and the timer would stop once it hit ten. Between the four search scenarios different conditions persisted.

The conditions included:

(1) No visual or audio cues to help in the search of the target symbol.
(2) A visual cue to help the user find the target symbol.
(3) A directional audio cue to help the user find the target symbol.
(4) A visual and audio cue used in tandem to help the user find the target symbol.

3.1 Experiment Design

Below is the implementation of the experiment. Four different search tasks were created to measure the impact of visual and audio cues. Upon opening the experiment program, participants were presented a menu where they could select one of the four search tasks. A task with no aid, a task with a visual aid, a task with an audio aid, and a task with both a visual and an audio aid (Figure 5). Participants were then read the script in Section 3.2 and instructed which specific task to begin. Every participant did the tasks in a different order determined by a Latin square in order to minimize the effects of learning within the test. In the task itself participants would start a timer and have to search for a predetermined target symbol ten times, shown in the top right of the screen. After they found the symbol ten times the timer would stop and their time would be recorded. Each figure below showcases how each task looked, Figure 1 shows the No Aid search where, participants searched for the heart symbol in the grid without any aid. Figure 2 showcases the Visual Aid Search, the visual aid itself was an orange, translucent outline around the target shape. Figure 3 depicts the Audio Aid search where a positional audio cues would play out of the left, right, or both ears depending on the target shapes position within the grid. Finally, Figure 4 shows the Both Aid search, which included the orange highlight around the target from the Visual Aid Search, as well as positional audio from the Audio Aid search simultaneously.

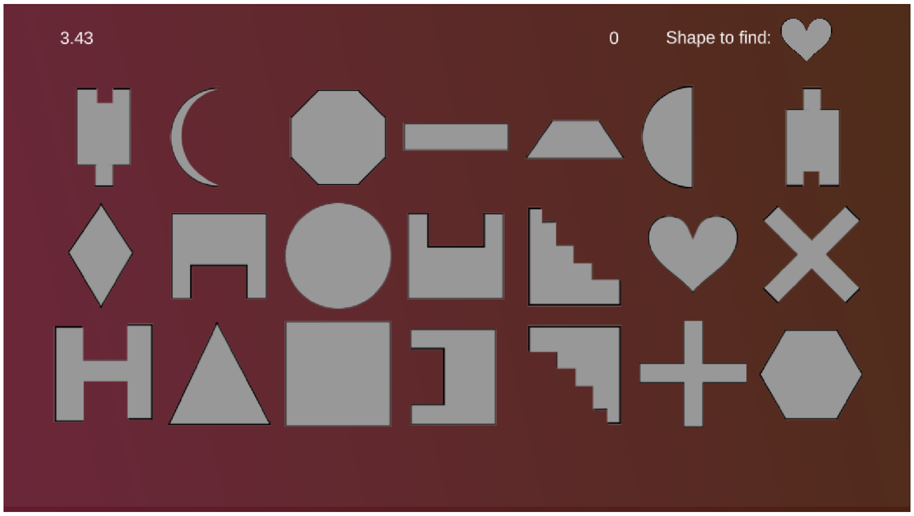

Fig. 1. No Audio or Visual Aid.

Fig. 2. Visual Aid Only.

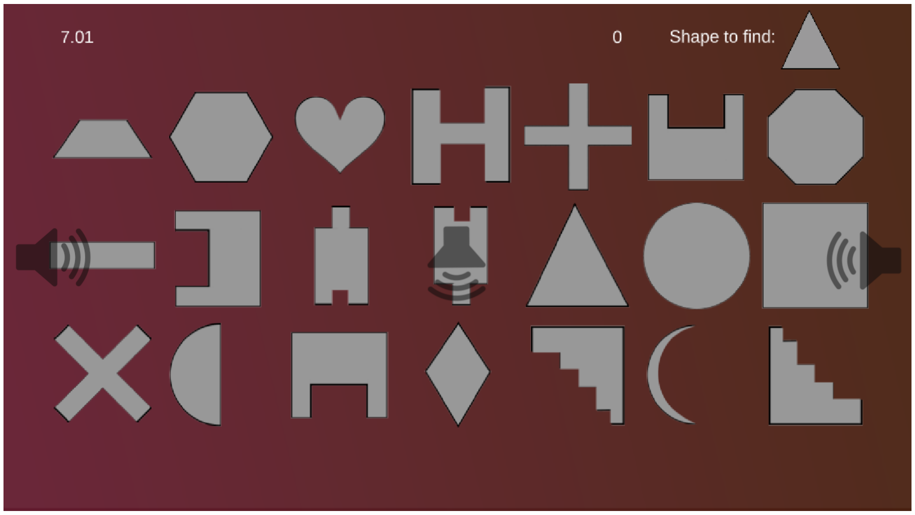

Fig. 3. Audio Aid Only.

Note: The position of the audio depends on the location of the shape to find. In this case the audio cue will be played in the participants right ear. Additionally, the speaker icons are not visible to the participant.

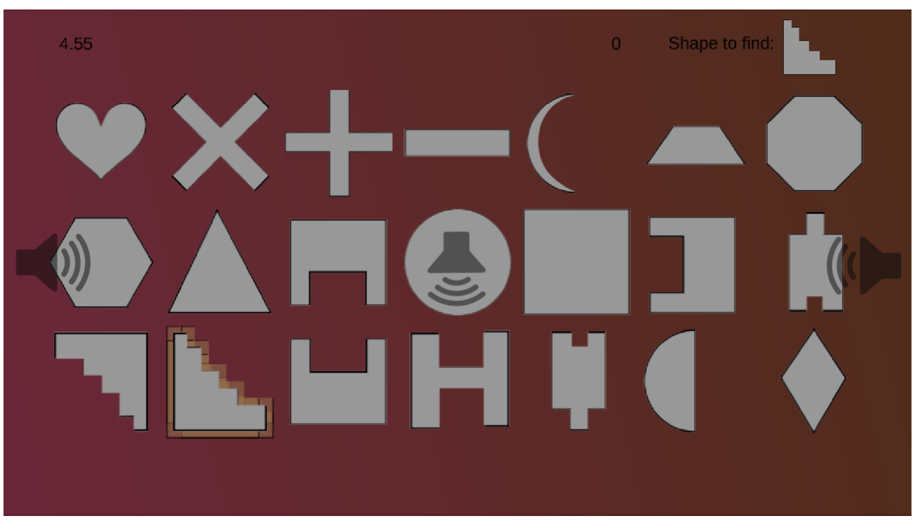

Fig. 4. Both Audio and Visual Aid.

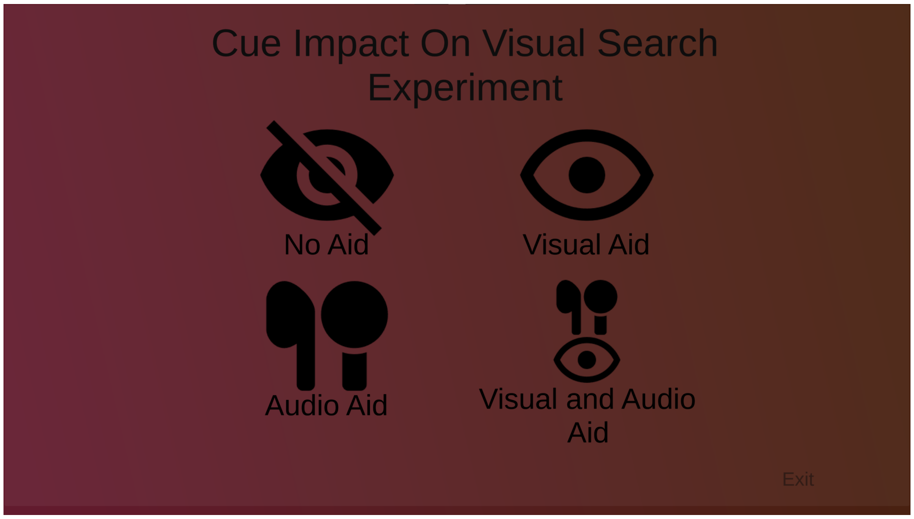

Fig. 5. Initial Menu.

3.2 Experiment Script

In order to reduce external variables in the testing process, a script was generated for the test administer to deliver to the subjects. This script gives a brief overview of the experiment, the four different cue types they'll be receiving, how to interact with the experiment software, and how to start the process. Additionally, safety information such as the expected risks and their ability to quit at any time were explained to each participant. While it does not reduce all external factors with administering the study, having a script helps ensure that no matter who is administering the test the participant still receives the same information.

3.3 Administration

The experiment was administered in one of two ways for a given participant.

(1) In person on a 17" laptop screen owned by the research team.
(2) Online via Discord, while the subject shares their screen with the researcher.

Ideally, all users would have participated in this experiment in the same room, using the same materials, the same size screen, and in the same environment. This would help prevent random variables that interfere with the testing conditions. However, due to the constraints of this study, it was deemed that a remote option would still provide useful information to the researchers, and would not introduce enough random variables to conflict with the results of the study. Additionally, allowing the use of personal computers from the participants provided more external validity to the experiment as a whole. It very well may have been the case where the comfort of the system the participant is using makes the results more applicable.

4 DATA

4.1 Collection

The experiment had 23 participants. Table 1 shows the raw data collected from these trials. Because the search scenarios were completed in a different orders for each participant based on the Latin Square. Each participant has a corresponding tests 1 through 4 in the table. The letter before their score indicates which search scenarios were completed during that test.

(1) A - No cues
(2) B - Audio cues
(3) C - Visual cues
(4) D - Visual and audio cues

Table 2 shows the data of each participant grouped into columns, with each column being one of the cue types. Because the tests were assigned in random order this table is purely for convenience to see the performance of each participant in each test. The final results must take into account the fact that learning occurs between each simulation.

4.2 Analysis

The average time taken with no assistance was 14.36 seconds. The average time taken with the audio cue assistance was 14.15 seconds. The average time taken with visual cue assistance was 15.56 seconds. The average time taken with visual and audio assistance was 11.48 seconds. These raw averages may indicate that cues don't give the user much of an advantage unless the audio and visual cues were paired together, but this does not represent the full story. Users learned how to better navigate the software in-between each experiment. When only looking at tests 3 and 4, where the user would already be familiar with the simulations, and removing the outliers, the new results show:

(1) No Cue - 13.961 seconds
(2) Audio Only - 13.341 seconds
(3) Visual Only - 12.32 seconds
(4) Audio and Visual - 11.34 seconds

Our table has a single independent variable with 4 modes, meaning we can use a single factor ANOVA to analyze the results. As displayed in Table 3, the ANOVA produces a P-value of 0.025. Because this is lower than the alpha value of 0.05, we are able to reject the null hypothesis of "There is no variation in time to complete a search task with or without visual and audio cues.". This demonstrates that visual and/or audio cues can create a significant decrease to the search time of a user.

Three additional single factor ANOVA tests were performed between the base case of "no cue" and the 3 other modes, to see which cues were able to reject the null hypothesis.

Table 4 showcases the relationship between the no aid search data and the audio aid search data. Utilizing the same ANOVA calculations as before, this ANOVA yielded a P-value of 0.756. Given that this is higher than the alpha value of 0.05 we are unable to reject the null hypothesis of "There is no variation in time to complete a search task with or without audio cues". This suggests that an audio aid on its own doesn't provide a significant decrease in the search time of a user.

Table 5's results were similar to Table 4's, running a single factor ANOVA on the no aid search data and the visual aid search data returned a P-value of 0.238. Resulting in a value larger than the alpha of 0.05 we again were unable to reject the null hypothesis of "There is no variation in time to complete a search task with or without visual cues.". Portraying that an visual aid on its own doesn't provide a significant decrease in the search time of a user.

Table 6 tells a different tale, however; The relationship we examined this time was between the combination of both visual and audio cues compared to no cues. Running a single factor ANOVA on the data resulted in a P-value of 0.022. Being less that our alpha value of 0.05 we are able to reject the null hypothesis of "There is no variation in time to complete a search task with a combination of visual and audio cues or without cues". This highlights that the combination of the audio aid playing the location of the target image, and the outline visual aid in tandem, did provide a significant decrease in the search time of a user.

4.3 Sources of Error

4.3.1 Search Object Shape. Each test had a static search shape. This means that every participant was looking for a heart shape for no-aid, a triangle for audio-aid, and so forth. We did this to reduce the source of error that some shapes were easier to find than others. However, it was still possible that among the shapes selected for our experiment, certain simulations had naturally easier to find shapes than another. For example, due to the unique shape of a heart compared to the other geometric shapes, many participants reported that it was easier to find in comparison to the tests that had a triangle.

4.3.2 Search Object Randomization. Every time an object is clicked during a simulation, the shapes are randomly reassigned a position on the screen. This means that it is possible a user got a "lucky" randomization that made it easier for their particular shape to be found. This may mean that the shape was placed near the cursor, in the top left corner, or somewhere else where the user is likely to start their search. Without having randomized seeds that ensures every user gets the shapes in the same location after each click, a source of error is introduced to the experiment.

4.3.3 Hardware Differences. This experiment was hosted online for a significant portion of participants. The subject shared their screen via Discord so the researcher could monitor and administer the test, as well as record the time information. This means that almost every participant completed the experiment with a different mouse, monitor, headphones, and general environment. While this gives a unique benefit of the user being familiar with their controls, potentially expanding the external validity of the test, it also introduces error as we are unable to account for the differences in hardware response time and monitor size impact on the search.

4.3.4 Administration. The experiment was administered by one of the four researchers for any given participant. While the researchers were given a list of information to provide to the participants before each experiment, the delivery of the information, the atmosphere the researcher produced, or even the tone of the researchers voice could create variance in how participants responded to the experiment.

4.4 Random Variables

The ability in which a user can control the mouse to quickly find a a target on the monitor may be correlated to the user's experience in gaming or general computer usage. The subjects in this experiment were not questioned about their prior experience or daily usage of a computer. Because of this, it is possible that this acted as a confounding variable in the study. It may be that cues are more beneficial for less apt computer users than they are for gaming veterans.

4.5 External Validity

The subject pool for this study consisted of primarily men in their early 20s. Many of these individuals would also self identify as "gamers". In the same way gamers may have less of a need for cues compared to other demographics, young adult males are only a small subsection of the population that uses a computer. This study failed to target middle aged people, senior citizens, teenagers, and women as a whole. Due to the restrictions in the study's sample size and breadth, this study lacks external validity. We can not confidently say that cues would impact demographics outside of men in their young 20s.

5 CONCLUSION

Often times screens can be littered with unnecessary clutter that gets in the way of what is needed to be done. Our data suggests that visual and audio cues used alone and in unison can help give direction in an otherwise confusing environment, as well as reduce the search time significantly. This decrease in search time signifies a decrease in the mental workload as less time is needed to find a target. Over the course of a career this can increase efficiency by days if not weeks. For the everyday user, this can make using a computer more seamless and more intuitive. Supported by studies dating back to the 1960s, we can drastically increase the search speed in a visual search and make simple tasks on a desktop more intuitive to the everyday person [22]. Integrating audio cues with visual cues will be a critical next step in human-computer interaction to make computers less overwhelming and more natural for people as a whole. Adapting technology to work well with our natural biology is an important part of human computer interaction and visual search and auditory cuing are hardwired into the human experience [11, 32].

6 FIGURES AND TABLES

Participant	Test 1	Test 2	Test 3	Test 4	Notes
1	A 12.87	B 11.78	D 11.47	C 13.25	23 y/o woman
2	B 7.82	C 19.22	A 11.01	D 8.15	22 y/o man
3	C 36.55	D 9.19	B 8.52	A 12.37	22 y/o man
4	D 11.95	A 13.53	C 12.78	B 8.09	24 y/o man
5	A 12.46	B 13.7	D 10.71	C 9.78	22 y/o man
6	B 23.04	C 8.4	A 11.91	D 7.75	20 y/o man
7	C 9.19	D 9.44	B 19.72	A 9.95	24 y/o man
8	D 9.91	A 12.56	C 9.71	B 18.91	21 y/o man
9	A 14.37	B 20.82	D 10.99	C 13.85	25 y/o man
10	B 8.35	C 26.65	A 13.28	D 10.22	23 y/o man
11	C 8.55	D 8.43	B 12.15	A 13.49	21 y/o man
12	D 7.30	A 10.02	C 9.34	B 13.35	21 y/o man
13	A 18.96	B 14.95	D 10.32	C 9.10	21 y/o man
14	B 11.83	C 8.52	A 11.31	D 8.47	20 y/o man
15	C 33.65	D 14.82	B 15.70	A 16.89	21 y/o man
16	D 23.13	A 20.10	C 17.42	B 17.81	21 y/o man
17	A 17.25	B 13.04	D 13.12	C 26.50	22 y/o woman
18	B 21.50	C 22.10	A 24.09	D 10.95	20 y/o man
19	C 19.14	D 13.13	B 11.59	A 18.09	21 y/o man
20	D 23.53	A 15.34	C 15.65	B 11.91	22 y/o man
21	A 19.41	B 14.72	D 13.30	C 18.06	23 y/o man
22	B 19.67	C 29.37	A 18.45	D 18.8	19 y/o woman
23	C 14.47	D 9.86	B 13.09	A 12.81	22 y/o man

Table 1. Raw Data From Experiment

Participant	No cues (A)	Audio cues (B)	Visual cues (C)	Visual and Audio cues (D)
1	12.87	11.78	13.25	11.47
2	11.01	7.82	19.22	8.15
3	12.37	8.52	36.55	9.19
4	13.53	8.09	12.78	11.95
5	12.46	13.7	9.78	10.71
6	11.91	23.04	8.40	7.75
7	9.95	19.72	9.19	9.44
8	12.56	18.91	9.71	9.91
9	14.37	20.82	13.85	10.99
10	13.28	8.35	26.65	10.22
11	13.49	12.15	8.55	8.43
12	10.02	13.35	9.34	7.30
13	18.96	14.95	9.10	10.32
14	11.31	11.83	8.52	8.47
15	16.89	15.70	33.65	14.82
16	20.10	17.81	17.42	23.13
17	17.25	13.04	26.50	13.12
18	24.09	21.50	22.10	10.95
19	18.09	11.59	19.14	13.33
20	15.34	11.91	15.65	23.53
21	19.41	14.72	18.06	13.30
22	18.45	19.67	29.37	18.80
23	12.81	13.09	14.47	9.86

Table 2. Raw Data From Experiment Sorted By Type

	Groups	Count	Sum	Average	Variance
SUMMARY	No Aid	24	355.22	14.80	13.01
	Audio	24	346.50	14.44	19.50
	Visual	24	408.26	17.01	69.01
	Audio-Visual	24	289.97	12.08	18.47

	Source of Variation	SS	df	MS	F	P-value	F crit
ANOVA	Between Groups	293.22	3	97.74	3.26	0.0251	2.70359
	Within Groups	2759.7224	92	29.996			
	Total	3052.938816	95				

Table 3. Single Factor Anova On All Groups

	Groups	Count	Sum	Average	Variance
SUMMARY	No Aid	24	355.22	14.80	13.01
	Audio	24	346.50	14.44	19.50

	Source of Variation	SS	df	MS	F	P-value	F crit
ANOVA	Between Groups	1.585	1	1.585	0.0975	0.7562	4.0517
	Within Groups	747.657	46	16.253			
	Total	749.24	47				

Table 4. Single Factor Anova Between No Cue and Audio Cue

	Groups	Count	Sum	Average	Variance
SUMMARY	No Aid	24	355.22	14.80	13.01
	Visual	24	408.26	17.01	69.01

	Source of Variation	SS	df	MS	F	P-value	F crit
ANOVA	Between Groups	58.61	1	58.61	1.429	0.2380	4.051
	Within Groups	1886.47	46	41.01			
	Total	1945.08	47				

Table 5. Single Factor Anova Between No Cue and Visual Cue

	Groups	Count	Sum	Average	Variance
SUMMARY	No Aid	24	355.22	14.80	13.01
	Audio-Visual	24	289.972	12.082	18.469

	Source of Variation	SS	df	MS	F	P-value	F crit
ANOVA	Between Groups	88.696	1	88.696	5.635	0.0218	4.051
	Within Groups	723.99	46	15.739			
	Total	812.687	47				

Table 6. Single Factor Anova Between No Cue and Audio-Visual Cue

ACKNOWLEDGMENTS

To Professor Ortega, for his assistance and guidance.

REFERENCES

[1] 1997. *The Effect of Auditory Cues on the Haptic Perception of Stiffness in Virtual Environments*. ASME International Mechanical Engineering Congress and Exposition, Vol. Dynamic Systems and Control. https://doi.org/10.1115/IMECE1997-0372 arXiv:https://asmedigitalcollection.asme.org/IMECE/proceedings-pdf/IMECE97/18244/17/6857199/17_1_imece1997-0372.pdf

[2] Amit Barde, Matt Ward, Robert W. Lindeman, and Mark Billinghurst. 2019. Less is More: Using Spatialized Auditory and Visual Cues for Target Acquisition in a Real-World Search Task. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 340–341. https://doi.org/10.1109/ISMAR-Adjunct.2019.00-16

[3] Robert S. Bolia, William R. D'Angelo, and Richard L. McKinley. 1999. Aurally Aided Visual Search in Three-Dimensional Space. *Human Factors* 41, 4 (1999), 664–669. https://doi.org/10.1518/001872099779656789 arXiv:https://doi.org/10.1518/001872099779656789 PMID: 10774135.

[4] M. L. Brown, S.e L. Newsome, and E. P. Glinert. 1989. An Experiment into the Use of Auditory Cues to Reduce Visual Workload. 20, 1, Article 5 (March 1989), 8 pages. https://doi.org/10.1145/67450.67515

[5] Louis K.H. Chan and William G. Hayward. 2013. Visual search. *WIREs Cognitive Science* 4, 4 (2013), 415–429. https://doi.org/10.1002/wcs.1235 arXiv:https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcs.1235

[6] Wu Yi-Shiun Chen, Taizhou and Kening. Zhu. 2018. Investigating different modalities of directional cues for multi-task visual-searching scenario in virtual reality. In Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology. (2018). https://doi.org/10.1145/3281505.3281516

[7] Yanxiang Chen, Tam V. Nguyen, Mohan Kankanhalli, Jun Yuan, Shuicheng Yan, and Meng Wang. 2014. Audio Matters in Visual Attention. *IEEE Transactions on Circuits and Systems for Video Technology* 24, 11 (2014), 1992–2003. https://doi.org/10.1109/TCSVT.2014.2329380

[8] Kathleen Brown Thomas Z. Strybel David R. Perrott, Kourosh Saberi. 1990. Auditory psychomotor coordination and visual search performance. *Perception Psychophysics* 48, 3 (05 1990), 214–226. https://doi.org/10.3758/BF03211521

[9] Tricia Kwiatkowski Thomas Heil Douglas S. Brungart, Sarah E. Kruger and Julie Cohen. 2019. The Effect of Walking on Auditory Localization, Visual Discrimination, and Aurally Aided Visual Search. *Human Factors* 61, 6 (2019), 976–991. https://doi.org/10.1177/0018720819831092

[10] Brody Downs, Aprajita Shukla, Mikey Krentz, Maria Soledad Pera, Katherine Landau Wright, Casey Kennington, and Jerry Fails. 2020. Guiding the Selection of Child Spellchecker Suggestions Using Audio and Visual Cues. In *Proceedings of the Interaction Design and Children Conference* (London, United Kingdom) *(IDC '20)*. Association for Computing Machinery, New York, NY, USA, 398–408. https://doi.org/10.1145/3392063.3394390

[11] Miguel P. Eckstein. 2011. Visual search: A retrospective. *Journal of Vision* 11, 5 (12 2011), 14–14. https://doi.org/10.1167/11.5.14 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/933487/jov-11-5-14.pdf

[12] Juan Carlo M. Figueroa, Raul Alberto B. Arellano, and Janeen Mikee E. Calinisan. 2018. A Comparative Study of Virtual Reality and 2D Display Methods in Visual Search in Real Scenes. In *Advances in Human Factors in Simulation and Modeling*, Daniel N. Cassenti (Ed.). Springer International Publishing, Cham, 366–377.

[13] Patrick Flanagan, Ken I. McAnally, Russell L. Martin, James W. Meehan, and Simon R. Oldfield. 1998. Aurally and Visually Guided Visual Search in a Virtual Environment. *Human Factors* 40, 3 (1998), 461–468. https://doi.org/10.1518/001872098779591331 arXiv:https://doi.org/10.1518/001872098779591331 PMID: 9849104.

[14] William W. Gaver. 1986. Auditory Icons: Using Sound in Computer Interfaces. *Hum.-Comput. Interact.* 2, 2 (jun 1986), 167–177. https://doi.org/10.1207/s15327051hci0202_3

[15] Cristy Ho and Charles Spence. 2005. Assessing the Effectiveness of Various Auditory Cues in Capturing a Driver's Visual Attention. *Journal of Experimental Psychology: Applied* 11, 3 (2005), 157–174. https://doi.org/10.1037/1076-898x.11.3.157

[16] James E. Hoffman. 1979. A two-stage model of visual search. (1979). https://doi.org/10.3758/BF03198811

[17] Kumar A Singh KD. Jain A, Bansal R. 2015. A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. (May 2015). https://doi.org/10.4103/2229-516X.157168

[18] Sreenivasan Kusumadevi. 2020. Effect of Computer Usage on Visual Reaction Time in Information Technology Professionals of Bangalore City. (2020). https://doi.org/10.5958/2320-608X.2020.00021.9

[19] Weiquan Lu, Dan Feng, Steven Feiner, Qi Zhao, and Henry Been-Lirn Duh. 2013. Subtle cueing for visual search in head-tracked head worn displays. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 271–272. https://doi.org/10.1109/ISMAR.2013.6671800

[20] John Paul McIntire. 2007. *Visual search performance in a dynamic environment with 3D auditory cues*. Ph. D. Dissertation.

[21] Stephen W. Mereu and Rick Kazman. 1997. Audio Enhanced 3D Interfaces for Visually Impaired Users. *SIGCAPH Comput. Phys. Handicap.* 57 (jan 1997), 10–15. https://doi.org/10.1145/250025.250029

[22] McCormick E.J. Mudd, S. A. 1960. The use of auditory cues in a visual search task. (1960). https://doi.org/10.1037/h0045878

[23] Hermann J. Müller and Joseph Krummenacher. 2006. Visual search and selective attention. *Visual Cognition* 14, 4-8 (2006), 389–410. https://doi.org/10.1080/13506280500527676 arXiv:https://doi.org/10.1080/13506280500527676

[24] S. Puschmann N. Turgut A. Kastrup C. M. Thiel H. Hildebrandt N. Schenke, R. Franke. 2020. Can auditory cues improve visuo-spatial neglect? Results of two pilot studies. *Neuropsychological Rehabilitation* 31, 5 (2020), 710–730. https://doi.org/10.1080/09602011.2020.1727931

[25] Cunningham JA Brickman BJ Haas MW McKinley RL Nelson WT, Hettinger LJ. 1998. Effects of localized auditory information on visual target detection performance using a helmet-mounted display. *Human Factors* 40, 3 (09 1998), 452–460. https://doi.org/10.1518/001872098779591304

[26] A.C. Nobre, J.T. Coull, V. Walsh, and C.D. Frith. 2003. Brain Activations during Visual Search: Contributions of Search Efficiency versus Feature Binding. *NeuroImage* 18, 1 (2003), 91–103. https://doi.org/10.1006/nimg.2002.1329

[27] Sadralodabai T. Saberi K. Strybel T.Z. Perrott, D.R. 1991. Aurally Aided Visual Search in the Central Visual Field: Effects of Visual Load and Visual Enhancement of the Target. Human Factors. (1991). https://doi.org/10.1177/001872089103300402

[28] James Merlo Jan B.F. Van Erp Peter A. Hancock, Joseph E. Mercado. 2012. Improving target detection in visual search through the augmenting multi-sensory cues. (May 2012). https://doi.org/10.1080/00140139.2013.771219

[29] Strybel T. Z. Rudmann, D. S. 1999. Auditory Spatial Facilitation of Visual Search Performance: Effect of Cue Precision and Distractor Density. (1999). https://doi.org/10.1518/001872099779577354

[30] Gideon. Shelton, Jose Kumar. 2010. Comparison between Auditory and Visual Simple Reaction Times. (Jan. 2010). https://doi.org/10.4236/nm.2010.11004

[31] Iyer Nandini Brungart Douglas S. Simpson, Brian. 2010. Aurally Aided Visual Search With Multiple Audio Cues. (June 2010). https://smartech.gatech.edu/bitstream/handle/1853/49859/SimpsonIyerBrungart2010.pdf

[32] CHARLES SPENCE and MAYA U. SHANKAR. 2010. THE INFLUENCE OF AUDITORY CUES ON THE PERCEPTION OF, AND RESPONSES TO, FOOD AND DRINK. *Journal of Sensory Studies* 25, 3 (2010), 406–430. https://doi.org/10.1111/j.1745-459X.2009.00267.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1745-459X.2009.00267.x

[33]]Sumikawa1985 K A Sumikawa, M M Blattner, K I Joy, and R M Greenberg. [n. d.]. Guidelines for the syntactic design of audio cues in computer interfaces. ([n. d.]). https://www.osti.gov/biblio/5972020

[34] M. Grohn T. Lokki. 2005. Navigation with auditory cues in a virtual environment. *IEEE MultiMedia* 12, 2 (2005), 80–86. https://doi.org/10.1109/MMUL.2005.33

[35] Erik Van der Burg, Christian N. Olivers, Adelbert W. Bronkhorst, and Jan Theeuwes. 2008. Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance* 34, 5 (2008), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053

[36] györgy wersényi. 2012. virtual localization by blind persons. *journal of the audio engineering society* 60, 7/8 (july 2012), 568–579.

[37] Jason Wuertz, Sultan A. Alharthi, William A. Hamilton, Scott Bateman, Carl Gutwin, Anthony Tang, Zachary Toups, and Jessica Hammer. 2018. A Design Framework for Awareness Cues in Distributed Multiplayer Games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173817

[38] Yuchong Zhang, Adam Nowak, Yueming Xuan, Andrzej Romanowski, and Morten Fjeld. 2023. See or Hear? Exploring the Effect of Visual and Audio Hints and Gaze-assisted Task Feedback for Visual Search Tasks in Augmented Reality. arXiv:2302.01690 [cs.HC]

[39] Heng Zou, Hermann J. Müller, and Zhuanghua Shi. 2012. Non-spatial sounds regulate eye movements and enhance visual search. *Journal of Vision* 12, 5 (05 2012), 2–2. https://doi.org/10.1167/12.5.2 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/933491/i1534-7362-12-5-2.pdf

Received 2 May 2023