Inspirational Stimuli Attain Visual Allocation: Examining Design Ideation with Eye-tracking

Henrikke Dybvik*[1]; Filip G., Abelson[1]; Pasi Aalto[1]; Kosa Goucher-Lambert[2]; Martin Steinert[1]
[1]*Norwegian University of Science and Technology, Norway*
[2]*University of California, Berkeley, United States*
**henrikke.dybvik@ntnu.no*

Inspirational stimuli may be used to support the design process. This article aims to elicit new insights on the impact of inspirational stimuli on design ideation with eye-tracking technology. We replicated a design ideation experiment's methodology [1] but collected eye-tracking data and a think aloud protocol. Preliminary results of eye-tracking data demonstrate that inspirational stimuli influence participants' eye movements and visual allocation. Specifically, participants examine inspirational words significantly more than neutral words—and participants examine design problem statements significantly more in absence of inspirational stimuli. We also observe distinct individual visual search strategies. Experimental procedure, data, and code are openly available to facilitate further replication efforts.

Introduction

Visual stimuli affect designers during concept generation [2], facilitating [1] or hampering [2, 3] design ideation depending on the visual stimuli's type and timing. To support design processes, visual stimuli may e.g., serve as "inspiration", which is often sought by designers. If visual stimuli are "inspirational," how does it affect designers' visual allocation? Will designers devote more visual attention to inspirational stimuli?

This work uses eye-tracking technology to obtain further insight into visual stimuli's effect on idea generation by replicating the task, stimuli, and experimental procedure of a design ideation experiment [1] (referred to as

Design Computing and Cognition DCC'22. J.S. Gero (ed), pp. 511-529.

"the original study" throughout this article) that investigated whether inspirational stimuli of varying analogical distance to the problem space influence design concept generation.

The original study tasked participants to generate concepts for open-ended design problems and found that inspirational stimuli increased idea fluency [1]. Compared to control stimuli, inspirational stimuli both nearer and farther from the problem space facilitated participants to produce more ideas, and the effect was most prominent after a period of time. Moreover, functional magnetic resonance imaging (fMRI) data suggested two different activation patterns of brain regions. Two search strategies were coined accordingly: The *inspired internal search* activates brain regions associated with memory retrieval and semantic processing—herein, participants likely recognize the inspirational stimuli as helpful or applicable to the design problem. The *unsuccessful external search* increase activation in brain regions associated with directing attention outwards and visual processing—participants likely continue searching the problem space for an inkling. Control stimuli generally produce unsuccessful external search while near stimuli produce inspired internal search. Far stimuli exhibit features from both strategies.

While inspirational stimuli evoked different brain regions and facilitated participants to generate a greater number of ideas, it remains unknown whether these ideas were different from ideas produced in control conditions and which words were most conducive to the "inspired ideas." These unknowns constitute the research objectives of this work, warranting the collection of two new data sources. To learn whether ideas are different with and without inspirational aid, we added a think aloud protocol to the experimental protocol. To determine the most conducive (or "inspirational") words and investigate if they impact the visual allocation of participants or participants' perception, we used eye-tracking.

This article investigates designers' eye movements during ideation; it addresses the following questions; if visual stimuli are "inspirational," how does it affect visual allocation; will more visual attention be devoted to inspirational stimuli?

This article aims to provide new insights from eye-tracking; it presents a preliminary analysis of eye-tracking data, specifically, the differences in visual allocation between stimuli. The results show that participants examined inspirational words significantly more than control words; and that design problem statements were examined significantly more in the absence of inspirational stimuli. We moreover observe distinct individual visual search strategies. Exhaustive analysis of eye-tracking data and transcription of the think aloud recordings will be presented in future publications.

Background

Scientific advancement relies on the core principles of reliability, repeatability, and ultimately reproducibility, as do experiments. However, more often than one might think, experimental results are not replicated, or they fail to replicate due to a lack of information or other unforeseen factors. In the Open Science Collaboration's replication efforts, 32% of original results yielded insignificant results in combination with new data [4, 5]. This is a part of science's broader problems, further exacerbated by publication bias [5, 6].

Replication is a challenge to design cognition's future, a future that simultaneously holds repeated testing of predictions (or results) as a key opportunity [7]. These are closely related, and when taken together, they actualize replicability. Replication is both an opportunity to- and necessary for improving reproducibility, since irreproducible results may occur even within studies of exemplary quality due to random or systematic error [4].

We advocate for minimizing a potential replication crisis in design research and for providing replication studies with a positive connotation (our impression is that replication studies are frowned upon). Thus, this work replicates previous research's methodology [1] while gathering new data sources, and is thus a replication and extension study.

Eye-tracking technology

Eye-tracking measures record eye movements and gaze location over time and task [8]. The first record of eye-tracking dates back to 1823, and it was until recent technology advancements an expensive and effortful method. Today, eye-tracking is more affordable and accessible due to video-based eye-trackers [8, 9]. There are two main types of video-based eye-trackers: table and head-mounted configurations [8].

These eye-trackers shine an infrared light at the eye, not visible to humans, and illuminates it. Eye-facing cameras record the infrared light's reflection, which produces a corneal reflection and the pupil center through a bright or dark pupil effect [10]. The corneal reflection appears as a glint on the eye. When the infrared light is aligned with the camera's optical axis, the pupil's reflection is directed towards the camera producing the bright pupil effect. When the infrared light is not aligned with the camera's optical axis, the pupil's reflection is directed away from the camera, thus producing the dark pupil. The gaze position can be calculated by using the location of the corneal reflection and the pupil center.

Eye-tracking data

Eye-tracking data are time-series data sampled at a given frequency yielding the gaze position [8]. When the gaze rests (or fixates) at the same target for a period of time, these gaze points can be aggregated into a *fixation*. Fixations consist, therefore, of both a duration and gaze position. Fixation lengths vary and are usually within the range of 180-330 milliseconds [11]. The rapid eye movements between fixations, occurring when scanning the visual space and moving the eyes, are called saccades. Herein, the visual input is suppressed [8, 11]. Some eye-trackers can also measure pupil dilation.

Eye-tracking in design ideation research

Design research uses eye-tracking to investigate visual reasoning in design activities [12]. Similar design ideation tasks with eye-tracking have explored differences between beginning and advanced design students during idea generation using stimuli of varying distance from the problem space, but used images as stimuli [13]. The Alternative Uses Test (AUT) has been used to investigate the relation between eye movements and idea output (creativity); participants were presented with images of 12 objects, and listed alternative uses of the object (i.e., ideated) for 2 minutes [14]. Eye-tracking and AUT have also been used to explore differences between designers and engineers in idea generation [15]. We have not found other studies investigating the effects of inspirational word stimuli during design ideation with eye-tracking.

Experimental method

This experiment differed from the original [1] in using head-mounted eye-tracking technology and the think aloud protocol as an additional task. Here, participants were seated at a desk in front of a monitor, equipped with a conventional computer mouse and keyboard to indicate new ideas and submit questionnaire ratings. Participants in the original experiment lay supine in the fMRI, used a response glove to indicate ideas and provide questionnaire ratings, and did not think aloud.

Participants were tasked to develop as many ideas as possible for 12 different open-ended design problems and instructed to “think aloud” by briefly explaining their idea in a think aloud protocol. Five words were presented along each problem in two blocks for 1 minute each, totaling 2 minutes of ideation time per design problem. The three first words were presented in the first block (called Wordset1), whereas the remanding two words were also presented in the second block (called Wordset2), i.e., the second block displayed all five words. A 1-back memory task was performed between blocks. Participants were exposed to three conditions: Near, Far, and Control. Words near or far from the problem space served as inspirational stimuli in the Near and Far conditions, while the Control condition reused words from the problem statement. See the original paper for an exhaustive description of the task, design problems, and word stimuli [1].

Participants were sequentially assigned to one of three counterbalanced groups of specific problem-condition pairs in the experiment's repeated measures design.

After each problem, participants rated the words' usefulness and relevancy, and the developed solutions' novelty (uniqueness) and quality. The number and timing of generated ideas were collected continuously.

Participants

None of the N = 24 healthy adults (18 male/6 female, 22 right-handed/2 left-handed ages 23-35, mean = 25.8 yrs., SD = 2.9 yrs.,) participating were native English speakers. Since glasses might interfere with the head-mounted eye-tracker, no participant wore glasses, but eight used lenses. We recruited through internal channels and contacts at the Norwegian University of Science and Technology (NTNU). Participants were graduate-level students or higher (minimum 4th year MSc, PhDs) at the Department of Mechanical and Industrial Engineering (MTP) and the Department of Design (ID) to ensure similar educational background as original participants. Monetary compensation was not given.

Experiment procedure and calibration

First, participants received general information in Norwegian about the experiment, its procedure, and the task, and gave informed consent. Then, after fitting participants with the eye-tracker, it was 3D calibrated according to manufacturers' "Best Practices" [16]. Thereafter, the experiment commenced by providing information, explaining the design ideation task, and the 1-back again. Finally, participants answered a demographic survey after completing the 1-hour experiment.

Hardware

The experiment ran on a conventional desktop computer with a 24-inch monitor, a conventional keyboard and mouse, and a head-mounted eye-tracker from Pupil Labs [17] with binocular setup (cameras on both eyes). See specifications below. Participants were seated in a chair approximately 70 cm from the monitor, see Fig. 1. A microphone was placed on a tripod in front of participants.

Higher accuracy in eye-tracking data may be acquired by using a chin rest. We were interested in areas, words, and patterns as a whole, which means sub-word accuracy was not necessary. We thought a chin rest might restrict participants and/or increase or induce a Hawthorne effect or other expectancy biases. A chin rest was therefore not used.

Hardware specifications:

- Desktop computer: Dell OptiPlex 7050, OS: Windows 10 Education 64-bit, CPU: Intel Core i7-7700 @ 3.60GHz, RAM: 32 GB
- Monitor: Dell UltraSharp U2412M, Size: 24” (61 cm), Resolution: 1920x1200 pixels, Refresh rate: 60 Hz
- Microphone: Zoom H1 Handy Recorder, fs: 48 kHz, Bit rate: 16 bit, Channels: 1 (mono recording)
- Eye-tracker: Pupil Core, World cam. Resolution: 1280x720 pixels, fs: 30 Hz, Field of view: 99 degrees x 53 degrees, Eye cam. Resolution: 192x192 pixels, fs: 120 Hz. Gaze accuracy 0.6 degrees, gaze precision 0.02 degrees.

Fig. 1 Experimental setup

Software

The experiment was recreated in an open-source software, PsychoPy v2021.1.4 [18], with wordsets presented as black text on a white background in font OpenSans. Letter height was set to 5 percent of the screen’s height in PsychoPy, which is 60 pixels on the monitor, or approximately 17 mm, which corresponds well with the fovea’s 1.5-2-degree visual field at 70 cm viewing distance [19] and the eye-tracker's 0.6 degree accuracy. Pupil Capture collected and recorded eye-tracking data. To synchronize eye-tracking data, audio data, questionnaire responses, timestamped ideas, and stimuli annotations, we used Pupil Network API. By using this API, we sat Pupil's clock to the global experiment clock in PsychoPy, and thereby ensuring time synchronization of PsychoPy and Pupil Capture. This API was also used to implement automatic data recording, ensuring that Pupil Capture began recording once the PsychoPy experiment was launched. Pupil Player[1], Pupil Labs' software, exported eye-tracking recordings from Pupil Capture.

Surface tracking

To recorded participants' gaze relative to the monitor and not only the video frame we used Pupil's Surface Tracker plugin in combination with AprilTags (small binary markers) fastened on the monitor's bezel. We designed and 3D printed custom monitor mounts to ensure no changes in marker setup during the experiment period. The planar monitor's surface was mapped out with Pupil's Surface Tracker and the exact size of the monitor was marked in the recording software.

Processing and analysis of eye-tracking data

To summarize, the following data modalities were recorded: eye-tracking data, audio recordings, number and timing of generated ideas, and subjective ratings via a questionnaire. The two latter have been preliminary analyzed; these results largely corroborated the original and are presented in its entirety elsewhere [20]. This article's scope is a preliminary analysis of eye-tracking data. A comprehensive analysis of all data will be published later.

Data processing

Exporting eye-tracking data

Raw eye-tracking data were exported to CSV files with Pupil Player and stored in participant-specific folders.

Apart from fixations, all eye-tracking data export without selecting and setting any parameters. Duration and dispersion thresholds must be selected before exporting since fixations spread out temporally and spatially. Pupil

---

[1] https://docs.pupil-labs.com/core/software/pupil-player/\#raw-data-exporter

Player uses a dispersion-based algorithm [21] that maximizes the fixation duration within the given parameters and outputs non-overlapping fixations. Pupil Capture calculates the gaze position using the dark pupil effect [17].

Our assumptions: The aim is to obtain an overview of which words were examined, not fixations within the words themselves. We recognized that participants could potentially fixate on a word for several seconds, i.e., a long fixation; we therefore wanted to prevent long fixations from being separated into a series of fixations. On the other hand, participants could also pay little attention to a word, e.g., recognize a control word, not find it interesting or helpful, and thus not spend any more time fixating on it. Such fixations have a short duration, but we want to capture them nevertheless.

The dispersion threshold was set to Pupil Player's maximum of 4.91 degrees. Pupil Labs states that there is no gold standard for setting fixation thresholds[2]. By exporting fixation data with different thresholds, we found that setting maximum duration too low caused a considerable number of fixations passing the threshold, which split long fixations into one or more shorter fixations. We found that we could capture fixations of a wide range of lengths by setting maximum duration to 4000 milliseconds. Although similar ideation research used a lower bound of 150 milliseconds [22], we set the lower fixation bound to 100 milliseconds (see e.g., Wass et al. [23]) to include potential short fixations.

Fixation files with other parameters can be exported since the raw data is publicly available.

Data concatenation

To ease data handling, a script using the Pandas library [24, 25] iterated across exported files of the same type (e.g., annotations, gaze, fixations) and concatenated the files into one larger file per data type. Assigning participant ID to each row in the concatenated data ensured each row's uniqueness.

Data analysis: Are participants paying more attention to inspirational words?

This article aims, as mentioned, to provide a preliminary analysis of eye-tracking data. We sought "the bigger picture; an overview of which words were examined; and an investigation of whether there are differences in visual allocation for the different wordsets. We hypothesize the following; participants spend significantly more time examining inspirational words (i.e.,

---

[2] See the following section in their "Best Practices": https://docs.pupil-labs.com/core/best-practices/\#fixation-filter-thresholds

words presented in Near and Far) than neutral words (i.e., compared to Control). We, therefore, evaluated eye-tracking metrics mostly on an aggregated level.

For data analysis and statistics, we used open-source Python libraries Pandas [24, 25], NumPy [26], SciPy [27], and Pingouin [28]—for visualization methods and plotting we used Seaborn [29] and Matplotlib [30].

Data quality

The eye-tracking software Pupil Capture appends a confidence score between 0 and 1 for each data point based on the quality of the pupil detection. To ensure high-quality data, we included only data with a confidence score above 0.8, thus discarding data with low confidence scores, e.g., blinking.

Heatmaps

Heatmaps provide a visual overview of gaze positions. First, our custom code implementation made two-dimensional histograms with each gaze data point binned into bins similar to the pixels of the monitor (1920 x 1200 pixels)—then, we smoothened the values with a Gaussian filter. Afterward, to increase visual differences between heatmaps, we filtered out values below a lower bound. The lower bound was the mean of the histograms' non-zero values, divided by 2. Fig. 2 illustrates the heatmap generation process. The color indicates the relative gaze distribution from green to red with increasing gaze density.

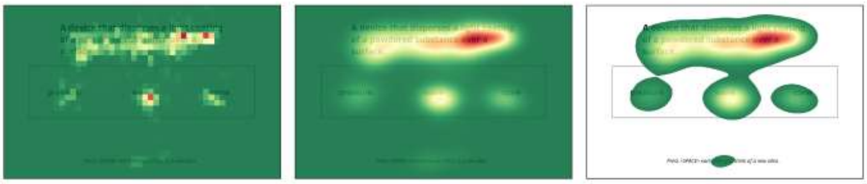


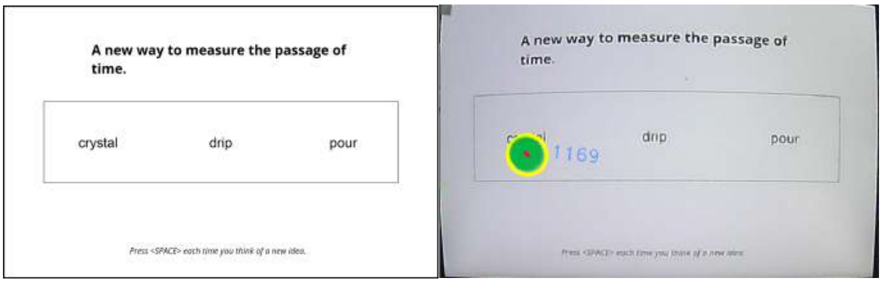
Fig. 2 Heatmap generation process from left to right: Left) 2D histogram binned to the monitors' pixels (here visualized with large bins for visibility); Center) Heatmap binned at each pixel with a Gaussian filter; Right) Lower values filtered out of center heatmap.

Fixation distribution

To obtain descriptive data of fixation distribution to test differences between stimuli, we split the monitor into four areas of interest (AOI) and calculated the ratio of time participants spent examining each area. The AOIs were:

problem, *words*, *off-screen* and *other*. *Problem* represents design problem statements, and *words* represent wordsets; these AOIs indicate how interesting the words are and how one might draw inspiration from the problem statement itself; they were selected to investigate differences between words and problems. We included *off-screen* since we observed some participants gazing outside the monitor when ideating in the experiment. *Other* represents the remaining parts of the monitor.

The Surface Tracker plugin does not map the monitor perfectly (as seen in Fig. 3) due to distortion from the world camera's fisheye lens. Upon preliminary inspection of heatmaps, we noticed a slight vertical offset relative to the text on the screen. We, therefore, extended the boundaries for the box encompassing the words, particularly for wordset 1, see Fig. 4. This distortion explains offsets when plotting heatmaps and scanpaths over a screenshot.



A new way to measure the passage of time.

crystal drip pour

Press <SPACE> each time you think of a new idea.

Fig. 3 Left) Frame as shown on monitor. Right) Monitor as mapped out by Pupil software.

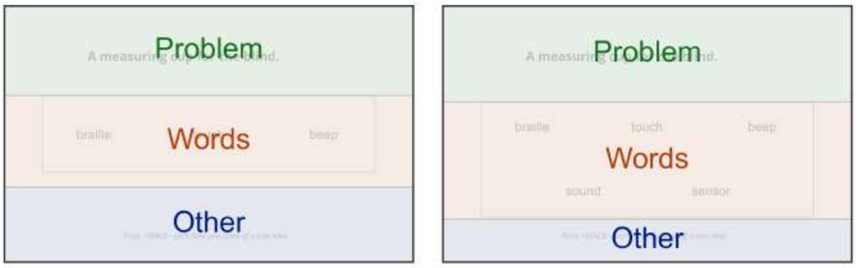


Fig. 4 Areas of interest (AOI) borders defined within the monitor for Wordset1 (left) and Wordset2 (right). Off-screen is outside the monitor.

Fixation distribution data was made by first assigning a label with corresponding AOI to each fixation based on the fixation's position on the monitor. Second, AOI distribution ratio was calculated by summing up fixation duration for each label and wordset (totaling 100%), and dividing by each label and wordset's fixation duration sum, respectively.

Statistical analysis of fixation distribution data

Friedman's test (a non-parametric test) assessed differences in fixation distribution between conditions due to violation of ANOVA's normality assumption for several subsets. Wilcoxon signed-rank test with a Bonferroni correction for multiple comparisons assessed pairwise comparisons post hoc. Hedges' g is used as effect size [6]. A significance level of $p<0.05$ was selected for all tests.

Scanpaths

Scanpaths visualize fixation data in a scatterplot where the dots are connected by lines. A dot indicates a fixation; the dot's size varies with the fixation's duration, i.e., larger dots indicate longer fixations. The lines connecting the fixations indicate saccades. The first and last fixation is indicated by a green and cyan point, respectively. This custom scanpath implementation plots the line between each fixation with chronologically varying opacity from transparent to opaque, meaning that the visualization retains the temporality in the eye-tracking data, whereas heatmaps only aggregates the position of gaze data.

Due to the temporal aspect of scanpaths, they are difficult to compare directly on an aggregated basis, which is possible with heatmaps.

Results

Heatmaps

Aggregated heatmaps for all combinations of conditions and wordsets are presented in Fig. 5. Firstly, we observe a strong tendency of gaze allocation towards the monitor's center for all conditions, which we attribute to the *central fixation bias* [31]; the "marked tendency to fixate the center of the screen when viewing scenes on computer monitors." In other words, the monitor's center is a natural place to rest the gaze when not actively scanning for new visual input.

Despite the *central fixation bias*, there is a clear visual difference between the inspirational stimuli and control words in the time participants spent looking at the different AOIs, as seen in Fig. 5. Gaze is allocated more to the problem statement in control conditions, whereas the gaze distributes more evenly over the entire monitor and more on the wordsets in Near and Far condition.

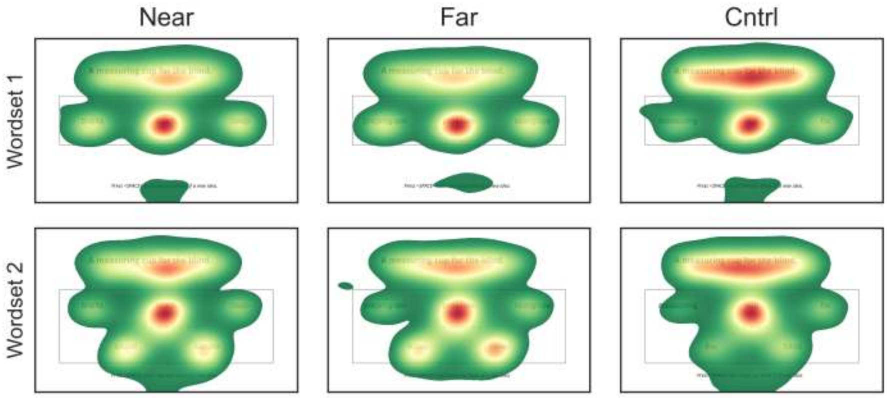


Fig. 5 Aggregated heatmap across all conditions

Fixation distribution

Friedmans test evaluated differences in fixation distribution between conditions and wordsets, i.e., objectively testing whether participants spent more or less time in any AOI. Table 1 presents the results, which were significant for AOI *words* and *problem* for both wordsets.

Table 1 Friedman test of AOI ratio

Wordset	AOI	DOF	χ^2	p
1	Problem	2	7.583	0.023*
	Words	2	18.250	<0.001**
	Off-screen	2	4.750	0.093
	Other	2	2.333	0.311
2	Problem	2	7.583	0.023*
	Words	2	14.333	0.001**
	Off-screen	2	2.333	0.311
	Other	2	0.083	0.959

*: $p<0.05$, **: $p<0.01$

Pairwise comparisons with Wilcoxon tests are presented in Table 2. For AOI *problem,* there were significant differences between Control and Near for Wordset1 and between Control and Far for Wordset2. Moreover, the difference between Control and Near for Wordset2 obtained a p=0.051, close to the significance threshold and thus noteworthy. Participants spent more time examining problem statements in control conditions compared to inspirational conditions.

For AOI *words,* there were significant differences between Control and Near, and Control and Far for both Wordset1 and Wordset2. Participants spent more time examining the inspirational words than control words. These findings align with our heatmap observations: when receiving control stimuli, participants spend time on the problem and less time on the words, compared to receiving inspirational stimuli, in which case participants spend more time on the words and less on the problem. This effect is apparent in Fig. 6 as well.

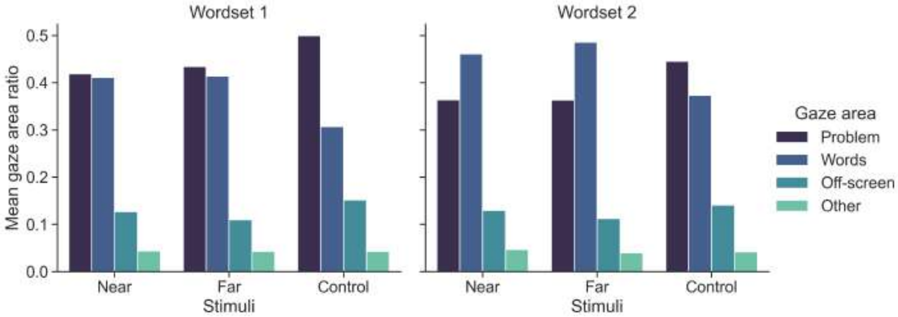


Fig. 6 Fixation distribution of AOIs for all conditions

Table 2 Wilcoxon signed-rank test with Bonferroni correction

WS	AOI	Between		W	p	Corr. p	Hedges' g
1	Problem	Control	Far	79.00	0.044	0.132	0.378
		Control	Near	45.00	0.003	0.008^{*}	0.482
		Far	Near	138.00	0.742	1.000	0.117
	Words	Control	Far	35.00	0.001	0.003^{**}	-0.608
		Control	Near	19.00	<0.001	0.001^{**}	-0.620
		Far	Near	143.00	0.853	1.000	-0.001
	Off-screen	Control	Far	75.00	0.033	0.100	0.311
		Control	Near	92.00	0.100	0.301	0.178
		Far	Near	111.00	0.271	0.814	-0.146
2	Problem	Control	Far	47.00	0.003	0.010^{*}	0.592
		Control	Near	66.00	0.017	0.051^{n}	0.570
		Far	Near	147.00	0.943	1.000	0.000
	Words	Control	Far	22.00	0.000	0.001^{*}	-0.695
		Control	Near	56.00	0.008	0.023^{*}	-0.510
		Far	Near	111.00	0.271	0.814	0.149

	Off-screen	Control	Far	107.00	0.225	0.674	0.256
		Control	Near	148.00	0.966	1.000	0.092
		Far	Near	107.00	0.225	0.674	-0.188

*: $p<0.05$, **: $p<0.01$, ^{n}: noteworthy, AOI *Other* is not included since its Friedman test was insignificant.

Discussion

The effect of inspirational words on participants visual allocation was significant. Inspirational words both near and far from the problem space received greater visual attention, i.e., participants spent more time visually fixating on the inspirational words, compared control words, throughout the entire ideation session. Further, participants visually examined the problem statement significantly more in control ideation sessions' second halves (Wordset2) compared to far inspirational ideation, and significantly more in control ideation sessions' first halves (Wordset1) compared to near inspirational ideation. The difference between control and near inspirational ideation sessions' second half (Wordset2) obtained a $p=0.051$, close to the significance threshold. It may have turned out significant with a larger or slightly different participant pool. Because its effect size ($g=0.570$) is comparable to that of the Control-Far (Wordset2) ($g=0.592$), we take this as an indication of the effect also occurring in the second half.

To summarize, this preliminary analysis of eye-tracking data yield two main findings/conclusions. One, participants allocate more visual attention (time) to word stimuli when receiving inspirational stimuli of any kind (both near and far from the problem space) compared to neutral (control) words throughout the entire ideation session; two, in the absence of inspirational stimuli (control condition) participants devote more visual attention to problem statements, an effect whose magnitude might depend on the inspirational words' distance to the problem space.

Finding one may be related to the *inspired internal search* from the original study, which suggested that participants found/recognized the inspirational stimuli as helpful or applicable to the design problem. Participants did rate inspirational stimuli as more useful both in the original study [1] as well as in this replication [20], which we suppose is why participants spent more time examining inspirational words than neutral words.

The second finding can be related to the strategy *unsuccessful external search* employed in the absence of inspirational stimuli where, originally, it is suggested that participants continue to search for clues in the design problem space [1]. The eye-tracking data confirm that participants continue trying to use the problem statement as a source of inspiration when they are not provided with any inspirational stimuli.

Scanpaths

The scanpaths presented here visualize a difference in how participants move their gaze around, possibly using different strategies during ideation. We selected an example illustrating participant 3 (in Fig. 7) versus participant 6 (in Fig. 8) for all problems in Wordset2 (both from group C)[3]. Participant 3 stays fairly central at all times, exhibiting the central fixation bias to a greater extent than participant 6, who moves vigorously around the visual space, looking for inspiration in almost every stimuli word, to us in a pattern strikingly similar to a hexagon. Although both participants show lacking interest in control stimuli words for problem 10, it appears that individual participants have different search strategies; this will be investigated further in future work. Further conclusions regarding search strategies are therefore not drawn here.

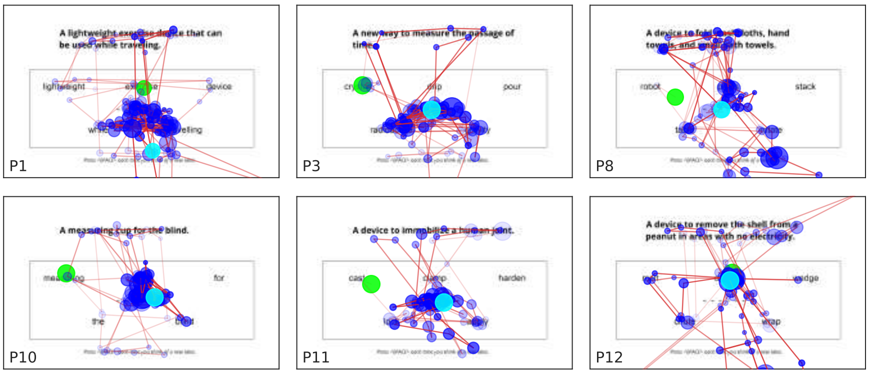


Fig. 7 Scanpaths for participant 3 for selected problems in Wordset2

---

[3] While only presenting a selection in this article, scanpaths were generated for all problems and participants.

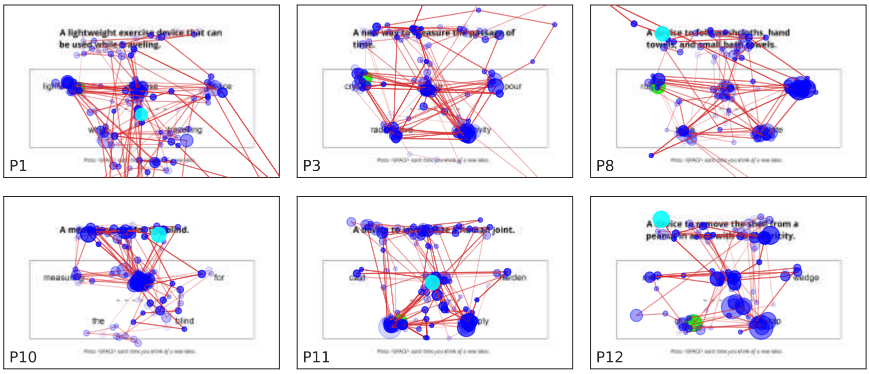


Fig. 8 Scanpaths for participant 6 for selected problems in Wordset2

Limitations

The study is limited by the central fixation bias, which could have been corrected for by randomizing the visual stimuli's position (problem statement and words etc.) on the monitor. However, randomization of positions was not possible since the study employed existing stimuli. Therefore, results are presumably influenced by the central fixation bias with *words* receiving disproportionately greater visual attention than other AOIs, as seen in Fig. 5. If we assume that the central fixation bias says consistent across conditions it is not affecting statistical results; if, however, it varies from condition to condition the statistical results are influenced.

Since this paper presents a preliminary analysis of eye-tracking data only the research objectives are not exhaustively answered; for this an exhaustive analysis is necessary. Further studies—collecting new and additional data modalities, in settings with higher ecological validity or in situ—are necessary to fully understand design ideation, visual- and inspirational stimuli.

Future work

Future work intends to present an exhaustive joint analysis of eye-tracking data, the think aloud protocol's transcription, the behavioral-, and subjective data measured.

The illustrated scanpaths appear to indicate that there may exist individual visual search strategies amongst participants, although we do not draw any conclusions here. Scanpaths are interesting descriptive data that we will use to inform future analysis. Currently, we do not know how nor if scanpaths will be useful or not, but this will be investigated in future work.

Conclusion

We used eye-tracking and a think aloud protocol in a replication and extension of a design ideation experiment with and without inspirational stimuli [1]. This article provided a preliminary analysis of eye-tracking data and aimed to provide new insights from eye-tracking technology. Results show clear influence from inspirational stimuli on visual allocation; participants examine (or gaze) significantly more on inspirational words than neutral words; in inspirational stimuli's absence, participants examine design problem statements significantly more. Finally, we facilitate further replication with the openly available experimental procedure, data, and code.

Published code repository and data

The code, raw data, and results from this study are publicly available:

- Code repository [32]: https://doi.org/10.5281/zenodo.5130090
- Pre-processed data [33]: https://doi.org/10.18710/PZQC4A
- Raw eye-tracking data [34]: https://doi.org/10.21400/7kq02wjl

References

1. Goucher-Lambert K, Moss J, Cagan J (2019) A neuroimaging investigation of design ideation with and without inspirational stimuli—understanding the meaning of near and far stimuli. Design Studies 60:1–38. https://doi.org/10.1016/j.destud.2018.07.001
2. Tseng I, Moss J, Cagan J, Kotovsky K (2008) The role of timing and analogical similarity in the stimulation of idea generation in design. Design Studies 29:203–221. https://doi.org/10.1016/j.destud.2008.01.003
3. Jansson DG, Smith SM (1991) Design fixation. Design Studies 12:3–11. https://doi.org/10.1016/0142-694X(91)90003-F
4. Open Science Collaboration (2015) Estimating the reproducibility of psychological science. Science 349:. https://doi.org/10.1126/science.aac4716
5. Shrout PE, Rodgers JL (2018) Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. Annu Rev Psychol 69:487–510. https://doi.org/10.1146/annurev-psych-122216-011845
6. Field A (2018) Discovering statistics using IBM SPSS statistics, 5th edition. SAGE Publications, Thousand Oaks, CA
7. Hay L, Cash P, McKilligan S (2020) The future of design cognition analysis. Design Science 6:. https://doi.org/10.1017/dsj.2020.20

8. Carter BT, Luke SG (2020) Best practices in eye tracking research. International Journal of Psychophysiology 155:49–62. https://doi.org/10.1016/j.ijpsycho.2020.05.010
9. Wade P of VPN, Wade N, Tatler BW, Tatler L in PB (2005) The Moving Tablet of the Eye: The Origins of Modern Eye Movement Research. Oxford University Press
10. Duchowski AT (2017) Eye Tracking Methodology, 3rd ed. Springer International Publishing, Cham
11. Rayner K (2009) Eye movements and attention in reading, scene perception, and visual search. Quarterly Journal of Experimental Psychology 62:1457–1506. https://doi.org/10.1080/17470210902816461
12. Gero JS, Milovanovic J (2020) A framework for studying design thinking through measuring designers' minds, bodies and brains. Design Science 6:. https://doi.org/10.1017/dsj.2020.15
13. Cao J, Xiong Y, Li Y, Liu L, Wang M (2018) Differences between beginning and advanced design students in analogical reasoning during idea generation: evidence from eye movements. Cogn Tech Work 20:505–520. https://doi.org/10.1007/s10111-018-0477-z
14. Kwon E, Ryan JD, Bazylak A, Shu LH (2019) Does Visual Fixation Affect Idea Fixation? Journal of Mechanical Design 142:. https://doi.org/10.1115/1.4045600
15. Colombo S, Mazza A, Montagna F, Ricci R, Monte OD, Cantamessa M (2020) NEUROPHYSIOLOGICAL EVIDENCE IN IDEA GENERATION: DIFFERENCES BETWEEN DESIGNERS AND ENGINEERS. Proceedings of the Design Society: DESIGN Conference 1:1415–1424. https://doi.org/10.1017/dsd.2020.161
16. Pupil Labs (2021) Best Practices - Tips for conducting eye tracking experiments with the Pupil Core eye tracking platform. In: Pupil Labs. https://docs.pupil-labs.com. Accessed 27 Apr 2021
17. Kassner M, Patera W, Bulling A (2014) Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. Association for Computing Machinery, New York, NY, USA, pp 1151–1160
18. Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, Kastman E, Lindeløv JK (2019) PsychoPy2: Experiments in behavior made easy. Behav Res 51:195–203. https://doi.org/10.3758/s13428-018-01193-y
19. Holmqvist K (2011) Eye tracking: a comprehensive guide to methods and measures. Oxford University Press, Oxford ; New York
20. Dybvik H, Abelson F G, Aalto P, Goucher-Lambert K, Steinert M (2022) Inspirational Stimuli Improve Idea Fluency During Ideation: A Replication and Extension Study with Eye-Tracking. In: Proceedings of the Design Society: DESIGN Conference. Cambridge University Press
21. Salvucci D, Goldberg J (2000) Identifying fixations and saccades in eye-tracking protocols

22. Vendetti MS, Starr A, Johnson EL, Modavi K, Bunge SA (2017) Eye Movements Reveal Optimal Strategies for Analogical Reasoning. Front Psychol 8:. https://doi.org/10.3389/fpsyg.2017.00932
23. Wass SV, Smith TJ, Johnson MH (2013) Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. Behav Res 45:229–250. https://doi.org/10.3758/s13428-012-0245-6
24. McKinney W (2010) Data Structures for Statistical Computing in Python. Austin, Texas, pp 56–61
25. Reback J, Jbrockmendel, McKinney W, Van Den Bossche J, Augspurger T, Cloud P, Hawkins S, Gfyoung, Sinhrks, Roeschke M, Klein A, Terji Petersen, Tratner J, She C, Ayd W, Hoefler P, Naveh S, Garcia M, Schendel J, Hayden A, Saxton D, Gorelli ME, Shadrach R, Jancauskas V, McMaster A, Fangchen Li, Battiston P, Skipper Seabold, Attack68, Kaiqi Dong (2021) pandas-dev/pandas: Pandas 1.3.0. Zenodo
26. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, Oliphant TE (2020) Array programming with NumPy. Nature 585:357–362. https://doi.org/10.1038/s41586-020-2649-2
27. Virtanen P, Gommers R, SciPy 1.0 Contributors, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17:261–272. https://doi.org/10.1038/s41592-019-0686-2
28. Vallat R (2018) Pingouin: statistics in Python. Journal of Open Source Software 3:1026. https://doi.org/10.21105/joss.01026
29. Waskom M (2021) seaborn: statistical data visualization. JOSS 6:3021. https://doi.org/10.21105/joss.03021
30. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. Comput Sci Eng 9:90–95. https://doi.org/10.1109/MCSE.2007.55
31. Tatler BW (2007) The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. Journal of Vision 7:4–4. https://doi.org/10.1167/7.14.4
32. Abelson FG (2021) Code Repository for Design Ideation Experiment (v1.0). Zenodo. https://doi.org/10.5281/zenodo.5130090
33. Abelson FG, Dybvik H, Steinert M (2021) Dataset for Design Ideation Study. DataverseNO. https://doi.org/10.18710/PZQC4A
34. Abelson FG, Dybvik H, Steinert M (2021) Raw Data for Design Ideation Study. https://doi.org/10.21400/7KQ02WJL