Similarity in Visual Designs: Effects on Workload and Performance in a Railed-Shooter game

David Milam[1], Magy Seif El-Nasr[2], Lyn Bartram[1], Bardia Aghabeigi[1], Perry Tan[1]

[1] School of Interactive Arts + Technology (SIAT), Simon Fraser University Surrey
250 -13450 102 Avenue, Surrey BC V3T 0A3 CANADA
{david_milam, lyn, baa17, pta13} @sfu.ca

[2] Northeastern University, Colleges of Computer and Information Sciences,
Arts, Media and Design,
Boston, MA, 02115 USA, 1 617.373.3982, magy@neu.edu

Abstract. Games are a popular form of digital entertainment and one elusive question is how complex visual designs affect the player experience. We address one aspect of this topic in terms of similarity of visual features, explored both as an organizing principle in Gestalt psychology and as a theory in visual attention. To address this issue, we developed a 3D railed shooter game with adjustable visual features of size, speed, and density of targets and non-targets. Based on these features we evaluate 105 player's performance in 4 visual conditions. In addition, we employ a cognitive workload assessment as a means to understand the perceived demands on players. Results show effects of expertise on performance and cognitive workload, per visual condition. Our methods and implications on game design are discussed.

Keywords. Game Design, Visual Design, Cognitive Load, User Research

1. Introduction

3D video games engulf users in sensory rich environments made possible through the visual design. The visual design helps distinguish the importance of elements among less important elements. Often without our awareness, activities, such as searching for artifacts or shooting enemies are comprised of visual search tasks. Visual search tasks require both perception of the environment and attention to elements with expressive features. This work considers visual search in situations, where the environment is too busy or contains many features at once causing a player to misperceive what is important. This disruption can lead to mistakes or a *breakdown in performance* [1]. Game designers remedy this problem by making game elements tied to goals clearly visible [2] and allude to the player's focus of attention [3, 4]. However, these works have not been empirically validated and are disconnected from a formal theory of attention. The goal of this paper is to apply a theory of visual attention to visual game elements and then analyze differences in players' performance and perception. We argue that this approach can improve game difficulty settings and accessibility.

Researchers within the fields of attention and perception proposed different theories through controlled experiments. Among the most foundational are *feature inte-*

gration [5] and *guided search* [6], whereby bottom-up (stimuli driven) and top-down (goal driven) approaches work in parallel in a given search task. Tests such as these are comprised of target and non-target elements combined with expressive features, such as color, shape, and motion. Attention researchers tested expert gamers and non-gamers in these tests [7], and found the experts have greater attentional resources than novices in spatial attention acuity, such as peripheral vision. Researchers in game approachability [8] reached similar conclusions, that inexperienced players perceive and perform differently in the face of challenge. While these insights are not surprising, the take away design lessons are not clear since the experiments either occur outside the context of a game or do not consider the visual relationships between elements in sufficient detail.

Visual search in sensory rich 3D games is complex due to the continuous change of visual stimuli, such as player's viewpoint and presentation of elements, over time. Gestalt organizational principles [9] support one well known approach to perceptually characterize this kind of complexity. The principles include proximity, similarity, common fate, continuity, and closure. Regardless of the underlining goal, the principles only consider visual relationships between element groups and the background.

We focus on *similarity*, defined as the degree of similarity between target and non-target elements sharing a common feature. Similarity is also chosen due to its application as a theory in attention [10], whereby a target search task decreases in efficiency and increases in reaction time as the non-targets become similar or dissimilar in appearance to targets. In this effort, we evaluate 4 experimental visual conditions according to the theory in a simple game called EMOS (*Expressive MOtion Shooter*). Only the presentation of target or non-target elements change, specifically speed, size, and density. Our research question is based on an analysis of players' play performance and self report between novice and expert players.

- **RQ1:** What design variables (speed, size, and density) applied to game target and non-target elements, tested in 4 different conditions (produced based on designers' choices), produce performance differences perceived and actual? Based on previous work, we know that expertise will have an effect on both performance measurements, thus we also look at expertise as a factor.

The previous work section introduces relevant work in visual attention and game user research methods. This is followed by our methodology including the experimental railed-shooter game developed, definition of visual conditions, and analysis procedure. Results are organized respective to each experimental condition.

2. Relevant Work

Many empirical approaches investigated visual design in games, such as dynamic lighting [11], color [12], and visual cues [13–15], to influence emotion or performance. To our knowledge, with the exception of our work [16], no one addressed multiple elements in motion within the context of a game. Within the field of visual perception, there is abundant research on the expressive properties of motion including *speed*, *shape* (the path a motion follows), *phase* (periodic motion), *flicker* (flashes), *smoothness*, and *direction* [16, 17]. For instance, many of these properties are found to influence affective impressions, such as valence, comfort, urgency, and intensity [17]. Another work evaluated the similarity of properties [18], specifically direction of motion and density of targets, on visual search.. Results found that reaction time increases not only with a higher density of targets, but also due to specific patterns of motion between target and non-target elements. The limitation with this work is that it is not situated within the context of a game. For instance, the random ordering removes any coherent escalation of difficulty, lack of task reward, or visual feedback, which is common in a game. We addressed these limitations in EMOS.

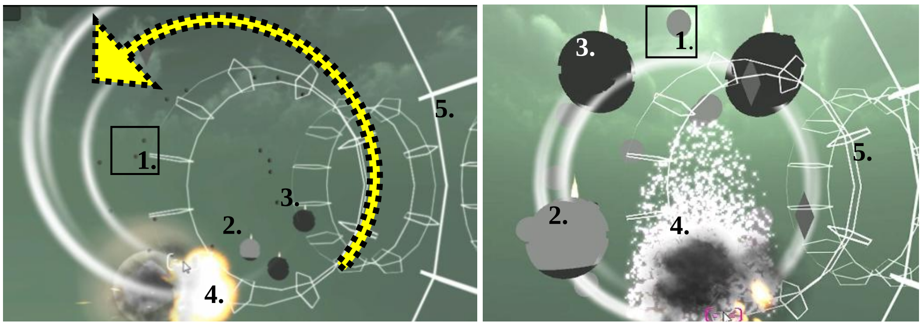
Game user research is a growing network of industry and academic groups [19, 20] with the goal to improve the play experience. Methods, such as triangulation and heuristic evaluation, have been applied to identify performance breakdowns. Triangulation methods include multiple data sources, such as the player's performance [20, 21] and attitudinal questionnaires [22, 23]. The most common metrics include time on task, task success, errors, and learnability (performance change over time) [20]. Others applied heuristic evaluation [8] to improve games for inexperienced players. For instance, game approachability principles [8] incorporate the player's self-efficacy [24] beliefs since the same level of challenge can be interpreted positively or negatively depending on expertise. These methods allow designers to gain in depth understanding *why* players feel the way they do, so corrective action can be taken.

3. Methodology

3.1. EMOS Railed Shooter Game

We investigate the impact of visual features on users' performance in the EMOS railed shooter game, developed in the Unity™ engine. Only one button is used to point and shoot. Additionally, there is no punishment as levels auto advance after 11 or 13 seconds. Like the perceptual tests, EMOS is comprised of 6 game elements; 2 types of targets, 2 non-targets, and 2 types of visual feedback. *Speed*, *size*, and *density* variables attach to the targets and ambient non-targets. The boss targets (TB) restrict advancement through levels and the minions (TM) increase the score. The non-target elements include ambient (NTA) and background ring (NTR) distractor elements. The ambient non-targets are the same color as boss targets, but do not contribute to the score or advance levels. The non-target rings rotate and are only variable by *speed*. Their purpose is to convey an illusion of camera movement. Finally, visual feedback for each shot fired (NTS) and on target explosions (NTE) is adjustable by *size*.

The game elements are abstract since we did not want the art-style to inform a strategy. To this extent, we also used simple geometric shapes, monochromatic color, basic sound effects, and periodic motion trajectories to the elements. Figure 1 shows an example of the game and how elements can visually change. The left image has small targets and many small ambient non-targets, while the targets on the right are larger with fewer ambient non-targets. In these levels, the boss and minion targets and ambient non-targets move along a circular trajectory (yellow arrow).



1. Ambient (NTA) 2. Boss (TB) 3. Minion (TM) 4.Visual feedback (NTE) 5. Rings (NTR)

Figure 1: Adjustable visual balance (top and bottom)

3.2. Visual Conditions

In our previous work [25], 8 expert game designers manipulated the variables: *speed*, *size*, and *density*, associated to target and non-target elements, in a toolset to manipulate the design of EMOS. Designers identified suitable designs for novice, intermediate, and expert levels. We then developed patterns from these designs. The study described in [26] discusses in depth the methodology, the study and the results. For the study presented here we will assume that the patterns used to construct the 4 conditions for the experiment are valid designs confirmed by eight expert game designers. We thus describe the 4 conditions here. Condition 1 is designed for intermediate difficulty level, and conditions 2-4 are designed for expert difficulty level. To preserve ecological validity as a game (as this is how designers indicated the design should proceed), condition 1 is played first by all players, followed by a random assignment to one of the expert difficulty conditions.

Due to the amount of elements that could change using the toolset, all four conditions modify *multiple* elements once. This approach is different from traditional perception experiments, which typically change one variable at a time. However, these conditions are more in sync with how designers approach level design. Designers typically change many elements between play tests. Therefore, to produce valid design lessons, the conditions need to adhere to designers' philosophy of how they typically manipulate the designs. The conditions produced are:

- **Condition 1: Increase T Density, Increase NT Speed:** Increases the density of targets, speed of targets, ambient, and ring non-targets. Boss and minion target density increases from 1 to 3, and 3 to 5, respectively. Target size is 3.5 times larger than the ambient non-target size. Target and ambient non-target speed increase 14% and 19% respectively, and ring speed increase 52%.
- **Condition 2: Increase T Density and Increase in NTR Speed:** Similar to condition 1 except target speed is held constant since fast speeds will be physically

harder to shoot for novice players. Boss and minion target density starts from 2 and 4 respectively and end with 3 and 7, respectively. Non-target ring speed is *twice* the speed in comparison to condition 1.

- **Condition 3: Increase similarity between T NT Size by Decreasing T Size:** Target and ambient non-target size become more similar with all remaining values are held constant. Target size decreases 24% respective to condition 1 and is now only 12% larger in comparison to the ambient non-target size.
- **Condition 4: Increase NT Density, NTE Size, and NTS Size:** Increases the ambient non-target density, visual feedback explosion and sparks size**.** In comparison to condition 1, the ambient density is 2.5 times higher, explosion size is 2 times larger, and sparks size is 28% larger.

3.3. Data Collection, Analysis, and Study Design

We collected 3 game metrics per level: level time (seconds), number of mouse clicks (#), and enemies shot (#). From these metrics, we generate the mean and rate of change, based on the slope of the linear regression line through data points in the known y axis (performance metric) and the known x axis (across 5 levels in each condition). Also, we administered two post-play surveys: 1) the task load index as a measure cognitive workload [26], and 2) five questions gauging gaming habits [22] as a basis for expertise. The workload survey contains questions regarding: mental demand (defined as following the target and ignoring distractions), physical demand (moving and clicking the mouse), temporal demand (feeling hurried or rushed), effort, success, and frustration. Each question is rated on a 7-point Likert scale (1=lowest and 7=highest). A response to an expertise questionnaire was also collected through surveys at the end so as not to bias participants [27]. The questionnaire gauged expertise based on the types and time spent playing games on a weekly basis. Players are novice if they play less than the average, prefer casual games, and non 3D genres.

4. Results

All analysis are within conditions, where expertise is evaluated in relationship to the independent variables (metrics and self report) using T-tests. Due to space limitations, we present values for statistically significant results only. The 105 participants include 48 females, 57 males, 55 experts, and 55 novices with an average age of 22.

- **Condition 1: Increase T Density, Increase NT Speed:** Mental, temporal, and effort are rated higher for the novice player in comparison to experts ($p < .001$, .004, and 002 respectively). Conversely success is perceived at a higher rate ($p < .019$) for expert players (mean rating is 5.4 vs. 4.9). Regarding performance, experts completed levels on average 28% faster with a lower rate of change ($p < .001$ and .007). Experts also fired 19% fewer shots with a lower rate of change ($p < .004$ and .005). As expected, experts perceived higher success in comparison to novices, and novices perceived higher cognitive load than experts. These perceptions suggest that even the intermediate difficulty was not as easy as we thought and may be problematic for novices. This is an important baseline.

- **Condition 2: Increase T Density and Increase in NTR Speed:** 20 experts and 15 novices. The success rating was perceived higher for experts ($p < 0.029$) in comparison to novices (mean 4.1 and 3.5 respectively), even though no significant change in performance was found within the condition. Thus, even though novices demonstrated improved performance from condition 1, and their performance was comparable to experts, they did not perceive high success rates in comparison to experts. In comparison to condition 1, all subjects improved performance, increased the enemies shot rate, while the level time and shots fired rate decreased.
- **Condition 3: Increase similarity between T NT Size by Decreasing T Size:** of 19 experts and 16 novices. No difference in self report ratings were found, even though on average experts finished levels 16% faster and shot 29% more enemies ($p < 0.004$ for both) in comparison to novices. In comparison to condition 1, all mean performance values increased, except for the enemies shot, which decreased for novices and remained constant for the expert players.
- **Condition 4: Increase NT Density, NTE Size, and NTS Size:** 16 experts and 19 novices. Success is perceived again higher for the expert ($p < 0.005$) in comparison to the novice player (mean 5.5 and 4.2 respectively) even though no significant change in performance was found within the condition. One possible explanation is that the large visual feedback size was perceived to be a higher reward even though the rules never change.

5. Discussion

Supported by a theory of stimuli similarity [10] in an experimental game, these findings are based on a definition of visual features attached to target and non-target, elements. In regard to the research question, we found that most differences in self reports and performance occur in the intermediate difficulty condition 1, rather than the more difficult conditions 2-4. This finding not only underscores the importance of training and preparation, but also sets a baseline in the perception of performance for the difficult levels that follow. Within conditions 2 and 4, novices demonstrated *improved* performance on par with expert players, yet report less success in comparison to experts. For novice players, feelings of success are dampened by the higher mental, temporal, and effort ratings in condition 1. In contrast, expert players report high success as a response to increased density of targets in condition 2, or the increased non-target visual feedback explosion size in condition 4. Within condition 3 by contrast, performance differences are found, yet no difference in self report ratings are found.

6. Conclusion and Future Work

This work investigated intermediate and complex visual designs informed by the similarity theory of visual attention, in an experimental railed shooter game. Our analysis found effects of expertise on performance and self report of performance in the intermediate difficulty condition. The same analysis in the complex visual settings found consistent *disagreement* in performance and perception of performance. Expert players report higher success in response to an increased density of targets or size of visual feedback, even though no performance differences are found. By contrast, no self report differences are found in the condition with an actual performance difference. Our contribution lies in the discussion of these results as well as the methods used to uncover them. Our future work incorporates pupillometry (pupil size) metrics as a continuous physiological metric and identifies additional implications on design.

7. Acknowledgements

This research is supported by the Natural Sciences and Engineering Research Council of Canada and Graphics, Animation, and New Media Network (GRAND).

8. References

1. Ryan, W., Siegel, M.A.: Evaluating Interactive Entertainment using Breakdown: Understanding Embodied Learning in Video Games. Breaking New Ground: Innovation in Games, Play, Practice and Theory. DIGRA, West London, UK (2009).
2. Smith, R.: Helping Your Players Feel Smart: Puzzles as User Interface. Game Developers Conference. , San Francisco, CA (2009).
3. Lemarchand, R.: Attention NOT Immersion: Making your games better with psychology and playtesting, the Uncharted Way. Game Developers Conference. , San Francisco, CA (2012).
4. Björk, S., Holopainen, J.: Games and Design Patterns. In: Zimmerman, E. and Salen, K. (eds.) The Game Design Reader. MIT Press, Cambridge, MA (2006).
5. Treisman, A.M., Gelade, G.: A feature-integration theory of attention. Cognitive Psychology. 12, 97–136 (1980).
6. Wolf, J.: Guided Search 2.0: A revised model of visual search. Psychonomic Bulletin & Review. 1, 202–238 (1994).
7. Green, S., Bavelier, D.: Effect of Action Video Games on the Spatial Distribution of Visuospatial Attention. Journal of Experimental Psychology: Human Perception and Performance. 32, 1465–1478 (2006).
8. Desurvire, H., Wiberg, C.: User Experience Design for Inexperienced Gamers: GAP - Game Approachability Principles. Evaluating User Experiences in Games. Springer-Verlag, London (2010).
9. Wertheimer, M.: Untersuchen zur lehre von der Gestalt [Laws of Organization in Perceptual Forms]. Psychologische Forschung [Psychological Research]. 4, 301–350 (1923).
10. Duncan, J., Humphreys, G.W.: Visual search and stimulus similarity. Psychological Review. 96, 433–458 (1989).

11. Seif El-Nasr, M., Vasilakos, T., Rao, C., Joseph, Z.: Dynamic Intelligent Lighting for Directing Visual Attention in Interactive 3D Scenes. IEEE Transactions on Computational Intelligence and AI in Games. 1, (2009).
12. Niedenthal, S.: Complicated Shadows: The Aesthetic Significance of Simulated Illumination in Digital Games, (2009).
13. Hoeg, T.: The Invisible Hand: Using Level Design Elements to Manipulate Player Choice, (2008).
14. Milam, D., Seif El-Nasr, M.: Analysis of Level Design "Push & Pull" within 21 games. Foundations of Digital Games. ACM, Monterey, CA (2010).
15. Samarinas, A.: Illuminating Minotaur's Lair: Light Design and Navigation in Gameworlds, (2009).
16. Bartram, L., Ware, C.: Filtering and Brushing with Motion. Information Visualization. 1, 66–79 (2002).
17. Lockyer, M., Bartram, L., Riecke, B.: Simple Motion Textures for Ambient Affect. Computational Aesthetics. ACM, Vancouver, Canada (2011).
18. Kingstone, A., Bischof, W.F.: Perceptual grouping and motion coherence in visual search. Psychological Science. 10, 151–156 (1999).
19. Isbister, K., Schaffer, N. eds: Game Usability: Advice from the Experts for Advancing the Player Experience. Morgan Kaufmann, San Francisco, CA (2008).
20. Tullis, T., Albert, B.: Measuring the User Experience: Collecting Analyzing, and Presenting Usability Metrics. Morgan Kaufmann, Burlington, MA (2008).
21. Pagulayan, R., Keeker, K., Thomas, F., Wixon, D., Romero, R.: User Centered Design in Games. In: Sears, A. and Jacko, J. (eds.) Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications. Lawrence Erlbaum Associates, New York (2008).
22. Erfani, M., Seif El-Nasr, M., Riecke, B.E.: The Effect of Previous Gaming Experience on Game Play Performance. Presented at the International Conference on Advances in Computer Entertainment Technology (2010).
23. IJsselsteijn, W., Poels, K., de Kort, Y.A.W.: The Game Experience Questionnaire: Development of a self-report measure to assess player experiences of digital games. Fun of Gaming (FUGA) workshop, (2008).
24. Klimmt, C., Hartmann, T.: Effectance, self-efficacy, and the motivation to play video games. In: Vorderer, P. and Bryant, J. (eds.) Playing video games: Motives, responses, and consequences. pp. 132–145. Lawrence Erlbaum Associates, Mahwah (2006).
25. Milam, D., Seif El-Nasr, M., Bartram, L., Lockyer, M., Feng, C., Tan, P.: Toolset to explore visual motion designs in a video game. SIGCHI, Play Experience Workshop. , Austin, TX (2012).
26. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, A. and Meshkati, N. (eds.) Human Mental Workload. North Holland Press, Amsterdam (1988).
27. Boot, W., Blakely, D., Simons, D.J.: Do action video games improve perception and cognition? Frontiers in Psychology. 2, (2001).