Augmentation Impacts Strategy and Gaze Distribution in a Dual-Task Interleaving Scenario

Olga Lukhova-Sanz a, Siegfried Wahl a,b, and Katharina Rifai a,b

Institute for Ophthalmic Research, University of Tübingen, Tübingen, Germany; Carl Zeiss Vision International GmbH, T&I, Aalen, Germany

ABSTRACT
When interleaving multiple tasks, people are confronted with a decision of how to distribute a finite amount of time between several tasks, which defines the task-interleaving strategy. In some challenging task interleaving scenarios where accurate timing is essential, people perform worse than they could have. With the growing advancement of technology, such as augmented reality, it became possible to impact people’s strategy and improve their performance. However, when augmenting visual input with additional visual content, the augmentation not only introduces the possible benefit but can also capture attentional resources. It is, thus, important to investigate how visual augmentation affects people’s performance in cases when otherwise people underscore in their performance. In the current study, using a psychophysics approach, it was investigated how visual augmentation impacts the task-interleaving strategy and, thus, performance in a dual-task setting with unequal task importance. In a simple dynamic 3D environment, four visual augmentations were generated aiming to prompt the user when it is more beneficial score-wise to switch from one task to another. The mean duration on one task before the task switch, as well as the resulting total performance, were evaluated in combination with the gaze direction distribution. In terms of the strategy and the total performance, all augmentations showed an advantage compared to when augmentation was not present. Furthermore, an abrupt augmentation onset based on the individual response time of the participant was more beneficial score-wise for the strategy compared to a constantly present visual augmentation. However, it affected the natural gaze direction distribution indicating the allocation of attentional resources to the augmentation. The results of this study provide an insight into potential visual augmentation designs aiming to improve user’s performance in a challenging dual-task interleaving setting.

1. Introduction
In daily life, people are inevitably faced with performing multiple tasks at the same time. One important type of multi-tasking is task interleaving. Given a finite amount of time, a person performs one task at a time, alternating between attending to several concurrent tasks. Task interleaving is very prominent in our society ranging from office workers (Dabbish et al., 2011) to air traffic controllers (Lee & Taatgen, 2002), and therefore, has been extensively studied in various disciplines (see Janssen et al., 2015). When interleaving between multiple tasks, people are essentially confronted with a decision of how to distribute a fixed amount of time between the given tasks, how much time to spend on one task before switching to the other, and at which moment in time to return to the first task. This scheduling decision is referred to as the task-interleaving strategy. The way how people distribute their time when interleaving between tasks is essential for different aspects of people’s personal and professional lives such as safety (Brumby et al., 2009; Janssen et al., 2012) as well as their productivity (Farmer et al., 2018; Janssen et al., 2019). Numerous studies investigated how efficiently people interleave between tasks (see Janssen et al., 2015). There is much evidence that in many dual-task scenarios people are capable to adjust and optimize their strategy to make their performance more efficient (Farmer et al., 2018; Janssen & Brumby, 2015). However, there are some challenging dual-task settings where people appear to underscore in their performance (Janssen et al., 2019). From multiple studies it is apparent that a particular adopted task-interleaving strategy can depend on numerous factors including complexity and characteristics of the tasks, personal motivation and experience with the tasks, and many others (Howes et al., 2009; Janssen & Brumby, 2010, 2015; Mone & Shelley, 1995). Among the factors affecting people’s decision on when to switch from one concurrent task to another while task interleaving, is the payoff function of the tasks (Farmer et al., 2018; Janssen et al., 2019; Janssen & Gray, 2012). The payoff function reflects the relative importance of each task and describes a reward level for each of the tasks in a task interleaving scenario. In one recent study it was shown that even though generally people appear to be flexible and adjust their strategy toward the higher performance during task interleaving, in a scenario when one task is significantly more important relative to the other, people tend to averse to...
the risk of failing the more important task and switch to it from the less important one more often than they could have (Farmer et al., 2018). Specifically, in a dual-task scenario with the unequal importance of two tasks, given a final amount of time, people tend to fall short on the favorable performance-wise mean duration on the less important task, leading to a decrease of the total performance. These findings well relate to everyday life observations when the accurate estimation of time is necessary for more efficient performance. One plain example is when a researcher has to interleave between, first, correcting a closed-answer test of his students, and, second, monitoring his or her manual coffee maker while preparing the morning coffee in the adjacent kitchen corner. Even though the researcher might be familiar with how long it takes for coffee to be ready, in the absence of additional information, he or she is likely to interrupt the test correction and go and check the coffee maker more often than necessary to averse to the risk of missing the critical point and switch it off in time. The total performance in such a scenario, however, would be lower than it could have been as the time spent on the test correction is shorter than it could have been, given a finite total time for both tasks.

With rapidly developing technology such as augmented reality, additional visual cues became a potentially powerful tool to purposefully guide user’s attention and improve user’s performance (Booth et al., 2013; Coughlan & Miele, 2017; Dey et al., 2018; Lukashova-Sanz & Wahl, 2021; Zarraonandia et al., 2014). On the other hand, augmenting visual input with additional visual content has also its drawbacks. In particular, it introduces a trade-off of the potential benefit for the performance and, at the same time, captures attentional resources, which essentially becomes a bottleneck for the design of visual augmentation (Akoumianakis & Stephanidis, 2005; Favela et al., 2010; Raja & Calvo, 2017). It is, therefore, of practical importance to investigate how visual augmentation impacts people’s performance on occasions when they perform inefficiently. In the case of a task-interleaving scenario with unequal payoff functions of each task, an augmentation can potentially support the user’s task-interleaving strategy by prompting him or her, when it is more beneficial performance-wise to switch from one task to another. It is not clear, however, how much in advance the augmentation should become informative before the suggested switch moment in time.

In this study, using a psychophysics approach, it was investigated how visual augmentation affects the task-interleaving strategy and, thus, performance in a dual-task setting with unequal task importance. The experimental paradigm was implemented in a simple dynamic 3D virtual environment. A set of visual augmentations was generated aiming to prompt the user when it is potentially better score-wise to switch between two tasks. In particular, the time-lead of the augmentation onset before the predefined suggested switch moment was modulated, ranging from a constantly present visual augmentation during the task performance, to an abrupt onset of the visual augmentation based on the individual response time of the participant. As a measure of the task-interleaving strategy, the mean duration on one task before switching to the other, as well as the resulting total performance over both tasks, were evaluated. As a measure of efficiency, a more efficient strategy is expressed in a longer mean duration spent on one task before switching to the other, resulting in higher total performance. In the current study, even though the total performance can possibly be improved using the augmentation, additional visual information can also draw attentional resources, which can be an issue in some real-life scenarios (Wickens, 2021). We addressed this potentially negative effect of visual augmentation by evaluating another performance parameter – the distribution of the gaze direction during the task performance. Specifically, the hypothesis was that visual augmentations, including the abrupt response-time-based and constantly present, would improve the participant’s task-interleaving strategy, but to a different extent. We expected a higher performance expressed in the total score when the participant was prompted with an exact switch moment (response-time-based augmentation) than when the participant, even though having an additional piece of visual information as a reference, still had to decide when it is long enough before the switch (continuous augmentation). Furthermore, an impact of augmentation on the gaze distribution was expected with a bias of gaze direction toward the location of augmentation when augmentation was present.

In the present study, four experimental augmented conditions were tested, namely, four time-leads of a motion onset of a head-contingent visual peripheral augmentation stimulus are compared. In a dual-task interleaving scenario, first, the impact of the time-lead of the augmentation on the mean duration on one task before switching to the other was evaluated. Second, the total performance was evaluated across various time-leads. Finally, the gaze direction distribution for all experimental conditions was evaluated.

2. Methods

2.1. Participants

In this study, 14 participants with an average age of 24.4 ± 3.7 years were tested. All procedures conformed to Standard 8 of the American Psychological Association’s “Ethical Principles of Psychologists and Code of Conduct (2010).” The study was approved by the ethics committee of the Faculty of Medicine at the University of Tübingen. Signed informed consent was obtained from each participant before the measurements. All participants except one had had previous VR experience, however, the prior experience with virtual reality was not a requirement to participate in the experiment.

2.2. Apparatus and stimuli

2.2.1. Apparatus

A dual-task setting was implemented in a 3D virtual environment. The visual content was displayed to the participant using HTC Vive Pro Eye (HTC Corporation, Taoyuan, Taiwan) virtual reality headset at a refresh rate of 90 Hz running on a Windows 10 PC with NVIDIA GeForce GTX 1070 graphics card (NVIDIA Corporation, Santa Clara,
California, USA). The horizontal and vertical fields of view of the headset reported by the manufacturer are 100° and 90°, respectively. The interaction with the environment was conducted by the participant via the HTC Vive controller. The eye-tracking data was collected using a built-in eye tracker at a rate of 90 Hz. The head rotation data were collected using tracking base stations 2.0 of HTC Vive Pro setup. The experimental paradigm was generated using the Unity Game engine version 2019.3.15.f1 (Unity Technologies, 2019). The data analysis was performed using Python 3.6 packages NumPy (Van Der Walt et al., 2011) version 1.19.1, SciPy (Virtanen et al., 2020) version 1.5.2 and Pandas (Mckinney, 2010) version 1.1.3. The statistical analysis was conducted using R (R Core Team, 2020) version 3.6.1. The data visualization was performed using Python packages Matplotlib (Hunter, 2007) and Seaborn.

2.3. Experimental procedure

2.3.1. Dual-task-interleaving setting: Configuration
A dual-task-interleaving scenario was designed and implemented in a virtual 3D environment. In contrast to a 2D paradigm, a 3D virtual reality environment offers a possibility to rotate the head freely while tracking the eye movements. This, in turn, enables the recording of a more natural gaze behavior compared to a screen-based experiment. The experimental setting is depicted in Figure 1. The environment consisted of two virtual rooms. Three blue spheres were placed in one room, and a red sphere in the other room. During the experiment, all of the spheres were continuously expanding in size until a maximum of five times the original size. Once a sphere reached its set maximum size, it “exploded” and reset to its original size followed by an expansion again. At the beginning of each trial, the participant had a starting capital of points. Each time a sphere exploded, the participant received a penalty of a certain amount of points (see Section 2.3.3). The ultimate goal of the participant was to keep as many total points in each trial as possible. To prevent a sphere from an explosion, the participant could reduce its size by touching an expanding sphere with a virtual stick operated by a mobile manual controller. As long as the participant was touching the sphere with the stick, the sphere was reducing in size at a constant velocity. If the sphere reached its minimum size, the size reduction was terminated. Once the participant stopped touching the sphere, the sphere continued to expand again. The participant was able to instantly switch between two rooms by pressing a trigger button of the manual controller. The participant could perform a task only in one virtual room at a time. Thus, during each trial, the participant interleaved two tasks, the blue and the red, by switching between two rooms: with the blue and the red spheres. Upon each expansion of the red sphere, its expansion speed was slightly varying between ±5% of its mean value. Doing so we intended to prevent adaptation of the participant to an exact time duration. In case the participant switched to the red room too late and the red sphere had previously exploded, it was indicated by the red sphere changing its color to black, and then turning red again once the participant reduced the size of it via touching it with the virtual stick. The participant was additionally provided with audio feedback upon each

![Figure 1. Experimental paradigm: a dual-task interleaving setting with the unequal importance of each sub-task. The unequal task importance was implemented through 1:10 penalty ratio for the explosion of each blue and red sphere, respectively. During each 60 sec trial the participant interleaved between the blue and the red tasks performing one sub-task at a time. The participant switched between two tasks by pressing a button on the HTC Vive controller. The example scene views demonstrate the case when augmentation was present, thus, the head-contingent ring is overlaid with the visual scene of the blue task. In the control baseline condition, the ring was not present. Note, that the example scene views demonstrate a larger field of view than the actual one in the headset – when performing the experiment in the virtual environment with the headset, the outer edge of the augmentation ring was not visible. For more details on the experimental paradigm see the description in the text.](image-url)
sphere explosion with distinct sounds for the blue and red spheres, respectively. Also, haptic feedback was provided to the participant via a vibration impulse of the manual controller to indicate a collision of the virtual stick and a sphere. The height at which spheres were presented was adjusted before the start of the experiment according to the height of each participant. Namely, the vertical coordinate of the headset was used as a reference to set the vertical position of the spheres. The vertical coordinate of the blue spheres as well as of the red sphere was set to a value 0.3 Unity-meters lower than the headset recorded height. To prevent a patterned hand movement on the blue task, the vertical coordinate of each blue sphere was set to a random value between \( \pm 0.2 \) Unity-meters of its set value after each sphere explosion as well as each new entry to the blue room. The horizontal coordinate of each of the blue spheres was set to \(-0.6, 0, \) and \(0.6\) Unity-meters, respectively, where 0 corresponds to the center of the room. The horizontal coordinate of the red sphere was set to a random value corresponding to \( \pm 45^\circ \) of visual angle relative to the headset to make the red task more challenging. The depth at which each sphere was located was set to 1 Unity-meter from the center of the room. Each time the participant switched to a room, the direction of the headset was set toward the middle of the corresponding room. The events of a collision of the virtual stick with any of the spheres, as well as the explosion of each sphere were constantly tracked and correlated with the timeline of the experiment. The spatial configuration of the experiment including augmentation is schematically shown in Figure 2. The details on the augmentation design are described in Section 2.3.2.

2.3.2. Visual augmentation: A head-contingent peripheral motion onset stimulus

As the visual augmentation, a subtle semi-transparent head-contingent ring was used. The term “head-contingent” indicates that the ring was always concentric with the forward direction of the VR headset, meaning, it was always in the same position within the visual field of the participant regardless of the head movement. The informative part of the augmentation was represented by a motion onset of an additional ring minor sector with a lower transparency level (Figure 3). The augmentation ring was designed as a head-contingent to be always fully visually available for the user irrespective of his or her heading direction. The peripheral location of the ring was selected to keep the augmentation unobtrusive for the tasks at hand. The inner and outer radii of the ring were arbitrarily set to 30° and 50° relative to the center of the visual field of view, respectively. Considering the field of view of the head-mounted device used to display the visual content, the outer border of the ring was not visible to the participants, whereas the inner border of the ring was laid in the peripheral part of the visual field of view. The size of the ring sector was arbitrarily fixed to 20° of the ring circumference resulting in 18° angle made by the arc of the sector at the center of the circle. To maintain the color scheme of the virtual environment when displaying augmentation, the color of the background ring and moving ring sector was set to gray, namely, RGB values (127, 127, 127). The transparency for the background ring was set to a constant value of 100. To avoid the ambiguity of which part of the ring sector to attend to when it is moving, a gradient was applied to the transparency. Specifically, the transparency of the moving ring sector was set to a gradually decreasing value starting from a value of 100 to match the background ring, reaching a minimum value of 50 to be well discriminable on the background ring. Upon the motion onset, the ring sector was moving clockwise at a constant speed of 68°/sec resulting in a full 360° turn in approximately 5 sec. The starting position of the moving ring sector was set to match its most right and least transparent (most opaque) edge with the top middle position of the background ring (see Figure 3). The time-lead of the ring sector motion onset was varied depending on the experimental condition. More details on the informative content of the augmentation and different experimental conditions can be found in Section 2.3.3.

2.3.3. Unequal importance of the tasks implementation: Unequal penalty

The unequal importance of the tasks was introduced via the game score. For each trial, the participant had a starting

---

Figure 2. The spatial configuration of the experiment. As an example, the blue room configuration is shown. The head-mounted device represents the head position of the participant, the rest of the figure schematically represents the virtual environment designed for the experiment. The details about the augmentation stimulus are described in Section 2.3.2.
capital of 200 points. Each time a sphere exploded, the participant was given a penalty of 1 point for a blue sphere and 10 points for a red sphere. At the end of each trial, the participant was provided with cumulative feedback with the total amount of points that were left from the starting capital, as well as how many points were lost on the blue and red spheres separately. Before each trial, the starting capital was reset to 200 points.

The penalty ratio is an important factor for the task-interleaving strategy (Farmer et al., 2018). Here, the selected ratio was used to capture a challenging scenario where the performance could be improved. Beyond the scope of the present study, a wider range of penalty ratios could be evaluated.

2.3.4. Estimation of the suggested switch moment

Before the main experiment, the best score-wise switch moment from the blue to the red task was estimated. The following independent parameters were considered: the penalty ratio for each respective blue and red sphere expansion, the trial duration, and the expansion speed of the blue and red spheres. The basic assumption was made that none of the spheres were saved from an explosion during the trial, which surely does not correspond to the reality, but is sufficient for the initial estimate. It was estimated that the minimum total penalty, or the maximum total score in a trial is achieved when the mean duration on the blue task is just below the time necessary for one red sphere to expand from its minimum to its maximum size. This estimate indicates that in the current paradigm the most efficient strategy in terms of performance for the participants was to stay on the blue task for such a period before switching to the red task just enough to keep the red sphere from an explosion. The individual parameters reflecting how well and how fast the participant performs the blue and the red task, namely, the rate of blue spheres’ explosions when performing the blue task, and the mean duration on the red task per visit were not considered for the estimation. For the current experiment, it was sufficient to know that the most efficient strategy in terms of performance for the participants was to keep as many red spheres from an explosion as possible.

2.3.5. Augmented task-interleaving: Experimental conditions

The idea of the augmentation was to prompt the participant when it is better to switch from one task to another to achieve a higher score, here: from the blue to the red task. In the

| Condition | Time-lead | Description |
|-----------|----------|-------------|
| Condition 0 | – | Control condition, no augmentation was present. |
| Condition 1 | 5000 ms | The motion onset of the ring sector was initiated instantaneously once the participant entered the blue task. The motion was continuous throughout the whole time spent on the blue task. |
| Condition 2 | 2200 ms | The motion onset of the ring sector was initiated later while the participant was performing the blue task, however, there was sufficient time before the sphere explosion for the participant to switch not immediately after detecting the motion. |
| Condition 3 | 1200 ms | The motion onset of the ring sector was abruptly initiated with a constant time-lead of 1200 ms until the red sphere explosion. This value was estimated from the response time test (see Section 2.3.6), where the response time is defined as the time that the participant needed to detect the motion onset, switch to the red room, and touch the red sphere. The value of 1200 ms was selected as approximately an average among the participants. |
| Condition 4 | Individual response times | The motion onset of the ring sector was abruptly initiated with an individual time-lead previously recorded in the response time test (see Section 2.3.6). The response time is defined as the time that the participant needed to detect the motion onset, switch to the red room, and touch the red sphere. |
experimental conditions where augmentation was present, the participant was presented with the visual augmentation whenever performing the blue task (for details on the augmentation design see Section 2.3.2). The participant was suggested to use the augmentation which aimed to optimize the participant’s mean duration of time spent on the blue task. Doing so, it was intended to increase the total score over each trial. The ring sector motion onset was based on how much time was left before the explosion of the red sphere. As described in Section 2.3.1, the expansion speed of the red sphere was slightly varied, and, thus, the duration of the sphere expansion from its minimum size to its maximum was also moderately varied. To account for this, and to ensure that augmentation was always presented on time, we considered the minimum duration of the red sphere expansion as a reference. In the present study, this minimum duration was set to 5 sec. The specific time-lead of the ring sector motion onset essentially defined the experimental conditions of the study. In total, five experimental conditions were tested, which are described in Table 1. In particular, Condition 0 was a control condition when no augmentation was presented, whereas in Conditions 1 to 4 the motion onset of the ring sector was initiated with respective time-leads before the red sphere explosion. In Condition 1, the motion onset was initiated instantaneously whenever the participant was performing the blue task, thus, the motion was continuous with no abrupt onset. In Condition 2, the motion started later during the performing of the blue task, however, with a sufficient time-lead for the participant to detect the motion and still have some time on the blue task before the switch. Specifically, the time-lead in Condition 2 was set to 2200 ms. In Condition 3, a short fixed time-lead of 1200 ms was used. The value was selected based on the results from the response time test (see Section 2.3.6), where the response time is defined as the time that the participant needed to detect the motion onset, switch to the red room, and touch the red sphere. The value of 1200 ms was selected as approximately the average response time among the participants. Finally, in Condition 4, individual response times were used as the time-leads, which were obtained from the response time test (see Section 2.3.6).

2.3.6. Response time test
In this study, the individual response time is referred to as the time necessary for the participant to notice the motion of the augmentation ring sector while performing the blue task, switch to the red task, and touch the red sphere with the virtual stick. To estimate individual response times, each participant performed a response time test. During this test the participant was instructed to follow the general procedure of the main experiment, except the score was not of importance in contrast to the main experiment, but rather the speed of response mattered. All of the experimental parameters were identical to the main experiment setting. During each trial of the response time test, the participant was, first, focusing on the blue task, then after a 2 sec period the motion of the ring sector was initiated. Before the test, the participant was instructed to switch to the red task via controller and touch the red sphere with the virtual stick as soon as possible after detecting the motion of the ring sector. Thereafter, the participant switched back to the blue task by pressing the trigger button and the cycle started again. Each participant continued this procedure for 5 trials of 60 sec, resulting in a total of approximately 75 response time estimates for each participant.

2.3.7. Main experiment: General procedure
Each participant performed three experimental sessions. In the first session, the participant performed two experiments: first, the baseline task interleaving where augmentation was not present (Condition 0), followed by a separate response time test. During the second session, another two experimental conditions were performed. Finally, during the third session, the remaining two experimental conditions were executed. The order of the experimental conditions in the second and third sessions was randomized. Each participant in each experimental condition performed 15 trials of duration 60 sec. The expansion speed of spheres was set such that if a participant would not touch any of the spheres during the whole trial, he or she would lose 30 blue spheres and 12 red spheres in each trial. The number of switches from the blue to the red task varied individually between participants and experimental conditions being approximately 15 switches to the red task per trial. Each condition was preceded by a practice session of five trials of 60 sec with the procedure identical to the main experiment. During the practice session all participants familiarized with the dynamics of the tasks.

2.3.8. Gaze direction data
To evaluate gaze behavior, the gaze direction data in the headset coordinates, as well as the world coordinates, were collected at a frequency of 90 Hz. Before each experimental session, for each participant, the built-in calibration procedure of the eye tracker was performed. The gaze position data was accessed using a customized written Unity script utilizing the HTC SRanipal SDK package functions (SRs) (2021). To prepare the data for further processing, first, similar to (Imaoka et al., 2020), the raw data were filtered based on the eye data validity bitmask value, which represents the bits containing all validity for the current frame. After the filtering, only the data where the eye data validity bit mask had value 31 for both eyes, were selected. Doing so, the data where the eye tracker partly or completely lost the pupil (including blinks) was filtered out. Next, the gaze position was calculated in spherical coordinates. In particular, the polar φ and azimuthal θ angles were computed using Equations (1) and (2). In Unity, the z-axis corresponds to the depth dimension.

\[
\phi = \arctan \frac{x}{2}
\]

\[
\theta = \arctan(2, \sqrt{x^2 + z^2})
\]

where \((x, y, z)\) are coordinates of normalized gaze directional vector in headset coordinates. Note that SRanipal returns the gaze direction vector in the right-handed coordinate system. To convert the coordinates to the left-hand coordinate system (same as Unity world coordinate system), the x-coordinate
was multiplied by \(-1\). To compute the gaze position in Unity world coordinate system, the gaze position in headset coordinate system was multiplied by the head rotation quaternion.

To summarize, the gaze direction in visual angle was computed in visual angle relative to the heading direction (headset coordinates), as well as relative to the forward direction in the virtual environment (world coordinates).

2.4. Analysis

2.4.1. Response time test

To estimate the individual response time of each participant from the response time test, the time duration between the onset of the ring sector motion and the moment of collision of the virtual stick and the red sphere was computed.

2.4.2. Main experiment: Behavioral data

The variable of interest was the mean duration on the blue task which reflects the task-interleaving strategy of the participant. Another variable of interest was the total score in points over both tasks in each trial, which is suggested to be directly correlated to the mean duration on the blue task. Both variables were normalized to the corresponding baseline value in Condition 0. Specifically, the mean duration on the blue task in Condition 0 was subtracted from those in the remaining four conditions. Similarly, the total score in Condition 0 was also subtracted from the respective scores in the augmented conditions. The main effect of the augmentation time-lead was statistically evaluated by one-way repeated-measures analysis of variance (ANOVA), followed by post-hoc analysis with Tukey correction to reveal significant differences between the conditions. The ANOVA assumptions were checked using the Shapiro–Wilk test for the normality assumption and the Levene test for the homogeneity of variances assumption. In case the ANOVA assumptions were not met, an alternative Kruskal–Wallis test as an alternative to one-way ANOVA was performed, with a subsequent Dunn test for multicomparison. The mean duration on the blue task and the total score were two separate dependent variables, and the augmentation condition was the independent variable. Furthermore, the correlation between the total score and the previously measured response time was evaluated. Specifically, the difference in performance in Condition 3 and Condition 4 as a function of the individual response time was computed and evaluated using Pearson correlation test.

2.4.3. Gaze direction data

The gaze direction in the headset coordinates, as well as the world coordinates, was evaluated for every participant in each experimental condition. Specifically, the gaze direction during performing the blue task was of interest as augmentation was present only when the participant was performing the blue task. The gaze direction distribution was evaluated by plotting the variance of the horizontal visual angle distribution of the gaze direction. The variance of the gaze distribution for each experimental condition was normalized to the variance of the gaze distribution in Condition 0 by taking the ratio. The main effect of augmentation time-lead on the gaze distribution was evaluated by one-way ANOVA, followed by post-hoc analysis with Tukey correction. The ANOVA was performed separately for the gaze direction in the headset and world coordinates, respectively. In case the ANOVA assumptions were not met, an alternative Kruskal–Wallis test was performed with a subsequent Dunn test for multicomparison.

3. Results

3.1. The mean duration on the blue task and the total score

In Figure 4, the normalized mean duration on the blue task as well as the normalized total score, averaged over 14 participants are presented. The average between-the-participants’

![Figure 4](image-url)
mean duration on the blue task in the baseline condition (Condition 0) was 3210 ms with a standard error of the mean of 112 ms. The corresponding total score value was 150 points with a standard error of the mean of 2 points. For the individual data see Supplementary information.

From Figure 4, the mean duration on the blue task as well as the total score in the baseline Condition 0 are the lowest compared to other conditions. This indicates that the experimental paradigm captured well the sub-optimal in terms of the total score performance of the participants, leaving some room for improvement for the augmented experimental conditions 1, 2, 3, and 4. The ANOVA assumptions for the mean duration on the blue task were not met, thus, a Kruskal–Wallis test was performed as a non-parametric alternative to one-way ANOVA. The main effect of the augmentation condition on the mean duration on the blue task was significant with $p < 0.001$. The total score follows a similar trend as the mean duration on the blue task, being higher for conditions where augmentation was present. The ANOVA assumptions for the total score were met. From ANOVA with the total score as the dependent variable, the main effect of the augmentation time-lead was significant with $p < 0.001$. From the post-hoc analysis (Dunn test), the difference in mean duration on the blue task between Condition 0 and Condition 1 was not significant. Comparing Condition 0 with the remaining three conditions, the difference was found to be significant with $p < 0.001$ for all conditions (Condition 2, Condition 3, and Condition 4). Furthermore, there was a significant difference in the mean duration on the blue task between Condition 1 and Condition 3, as well as between Condition 1 and Condition 4 with $p < 0.001$ and $p < 0.01$, respectively. The multicomparrison of the total score revealed significant differences between Condition 0 and all other conditions with $p < 0.001$ for all conditions. General learning of the task was evaluated on a trial-by-trial basis for each participant and was not significant (see Supplementary information).

Furthermore, we evaluated the performance in the condition with short fixed time-lead (Condition 3), and the condition with individually set response-time based time-leads (Condition 4). In Figure 5, the difference in the total score between Condition 3 and 4 as a function of individual response time is shown. Note, that the score in Condition 3 was subtracted from the score in Condition 4. The dependency on Figure 5 shows a negative trend, however, there is no significant correlation found. The absence of significance is highly influenced by a data point for one participant with the largest response time. One possible explanation of this outlier is a general improvement in the task performance for this participant as the response time was measured during the first experimental session, Condition 4 was performed during the second session, and Condition 3 – during the third session.

3.2. Gaze direction distribution

In Figure 6a,b an example of a gaze direction dataset for one participant in one experimental condition is shown. The plot represents only the data when the participant was in the blue room. The data is visualized through a joint histogram plot with hexagonal bins where the depth of the color represents the count of data points in the corresponding region. Figure 6a demonstrates the gaze direction in degrees of visual angle in the headset coordinates. The gray ring represents the interval of visual angles in the headset coordinates corresponding to the augmentation ring in experimental conditions with augmentation present. In Figure 6b the same data is combined with the head orientation and represents the gaze direction in the world coordinates, thus, no head-contingent augmentation ring is depicted (for details on computing the gaze direction in the headset coordinates please refer to Section 2.3.8). The variance of the horizontal visual angle of the gaze direction distribution in all experimental conditions.

![Figure 5](image-url)  
Figure 5. The difference in total score in Conditions 3 and 4 as a function of individual response time. Note, that the score in Condition 3 was subtracted from the score in Condition 4. The blue line corresponds to the best linear regression estimate with a confidence interval of 95%. The black dashed vertical line indicates the fixed value of the time-lead 1200 ms in Condition 3.
was compared. To illustrate the trend of the shape of the gaze distribution in different experimental conditions, in Figure 6c, d the gaze direction distribution in horizontal visual angle is demonstrated, combined for all participants, and grouped by the augmentation condition. The ANOVA assumptions were not met for the gaze distribution in the headset coordinates as well as the world coordinates. Thus, a Kruskal–Wallis test was performed as a non-parametric alternative to one-way ANOVA. A significant main effect of the augmentation time-lead on the variance of the horizontal visual angle distribution of the gaze direction in the headset coordinates was found with $p < 0.01$. From the post-hoc analysis of the gaze direction horizontal visual angle variance in the headset coordinates, a significant difference was found between the baseline Condition 0 and Condition 3, as well as between Condition 0 and Condition 4 with $p < 0.05$ and $p < 0.001$, respectively. Statistical analysis of the variance in the world coordinates did not reveal any significant differences between experimental conditions. These results are further discussed in Section 4.

4. Discussion

4.1. General discussion

In a dual-task interleaving scenario with unequal task importance using the psychophysics approach, the impact of visual augmentation on the strategy in combination with the gaze direction distribution was examined. It was investigated how the time-lead of the onset of the informative property of motion-based augmentation relative to the suggested switch moment impacts the user’s performance.

In a 3D virtual environment, a set of visual augmentations was generated aiming to prompt the user when it is more efficient to switch from one task to another, given a finite time
of each trial. As the informative property of augmentation, a motion of a head-contingent peripheral visual stimulus was used. The mean duration on one task before switching to the other task, and the total performance, as well as the gaze direction distribution, were evaluated.

Overall, the experimental paradigm well captured a sub-optimal performance in a dual-task interleaving scenario with unequal task importance. Similar to the results reported in (Farmer et al., 2018), in our experiment participants on average spent less time on the less important task before the switch which led to a lower total performance represented by the score. This is supported by significant differences between the baseline condition and the augmented conditions for both the mean duration on the blue task and the total score (see Figure 4).

4.2. Comparison of the continuous and the response-time-based augmentation (Condition 1 vs Condition 4)

A significant difference in the mean duration on the blue task between Condition 1 and Condition 4 demonstrates that the response-time-based augmentation can be more beneficial performance-wise compared to the continuously present augmentation. This result was anticipated, as one would expect the participant to feel safe to stay longer on the blue task if it is known that he or she would be prompted with an exact optimal moment of the task switch (Condition 4). In Condition 1, even though a visual reference of the process was available to the participants, it was more difficult for the participant to stay the same amount of time on the blue task, when they still had to decide when it is long enough but not too lengthy (Condition 1). On the other hand, the analysis of the gaze direction distribution revealed differences between experimental conditions. We initially expected a bias of gaze direction toward the location of augmentation when augmentation was present, particularly, for the response-time-based augmentation due to its abrupt motion onset nature (Abrams & Christ, 2003). In contrary to the original expectations, for Condition 1 and 4, no bias of the gaze toward the periphery was found, however, in Condition 4 the gaze direction was more confined toward the center of the visual field. The gaze direction distribution in the world coordinates was not found to be different in various conditions. This can be explained in the following way: when awaiting for an abrupt motion onset in the periphery in Condition 4, to gaze to a direction in the world coordinates further from the center of the room, participants preferred to rotate their head and keep the gaze closer to the center of the visual field instead of directly moving eyes and gazing further from the center. It is suggested, that as rapid detection of the augmentation motion onset was crucial in Condition 4, participants allocated more covert attention to the periphery of the visual field, thereby ensuring the fast motion detection by keeping the eccentricity of the visual stimulus stable. In contrast, in Condition 1, where instant motion onset detection was not crucial, participants pursued their natural gaze behavior similar to the baseline condition.

4.3. Comparison of the individual and the fixed time-lead based on response time (Condition 3 vs Condition 4)

The response-time-based augmentation was based on the individually measured response times for each participant, measured before the main experiment. However, keeping in mind a real-life application, a more universal augmentation design is more time-efficient avoiding excessive individual pre-measurements. Considering this, the third experimental condition was equivalent to the response-time-based condition where instead of using the individual response time as a reference for the motion onset of the ring sector, a fixed time-lead was used. In Condition 3, the mean duration on the blue task as well as the total performance was very similar to those in Condition 4. A negative trend of the correlation of the total performance in Condition 4 and 3, as a function of the individual response time, was found, however, it was not significant. Nonetheless, for the “slower” participants, who needed more time to detect the augmentation and switch to the red task, the performance was lower when the time-lead of augmentation was fixed to a value smaller than their response time, except one outlier with the largest response time (see Figure 5). The general performance in the main experiment of that outlier participant was, however, one of the best among the participants. It is possible, that the actual response time of that participant in the main experiment was smaller than the one measured in the original response time test. This would then result in a significant negative correlation in Figure 5. The gaze direction distribution in Condition 3 is similar to that in Condition 4, being more constrained toward the center of the visual field compared to conditions 0, 1, and 2. Even though in the current experiment no significant differences were found between the conditions with the fixed time-lead (Condition 3) and individual response-time-based condition (Condition 4), the individual abilities to use augmentation can play a role in how optimal in terms of performance the augmentation is. Thus, when designing augmentation, one should keep in mind potential individual differences.

4.4. Condition 2

Condition 2 was aimed to probe a time-lead of augmentation longer than the time-lead based on the response time, but shorter than the time-lead for the constantly present motion stimulus. In this condition, participants were expected to perform either similar to when the constantly moving augmentation was present (Condition 1), or a decrease in performance was expected. In particular, a lower performance was expected due to possible inhibition of return (Klein, 2000). This term refers to a phenomenon when attention is maintained in a certain location for a while, but if no relevant information is shown in that location, attention is driven away, and the visual performance in the previously attended location is inhibited for a short period of time. We suggest that in Condition 2 of the current study, the participant’s attention was likely to initially shift toward the augmentation. Thereafter, if the switch to another task hasn’t occurred, after some time attention is shifted away from the
augmentation, and the optimal moment of the switch could be potentially missed, leading to decreasing of the total score. From the results of Condition 2, few participants had a drop in performance possibly originating from the inhibition of return. However, on average the mean duration on the blue task as well as the total performance in Condition 2 did not significantly differ from other augmented conditions. It is suggested that the timescale of the inhibition of return effect is individually varied (Klein, 2000), and was potentially resolved only for few participants in the current experiment. Beyond the scope of the current study, further investigation of the impact of the inhibition of return is needed to test this hypothesis. The gaze direction distribution in Condition 2 also did not significantly differ from natural gaze behavior (Condition 0). It is suggested, that as participants had an additional second to use the augmentation while performing the blue task, a rapid motion onset detection was not crucial, thus, maintaining the gaze closer to the center of the visual field was not necessary in contrast to Conditions 3 and 4.

4.5. Conclusion and outlook
To conclude, visual augmentation can support people’s strategy in a challenging dual-task-interleaving scenario with the unequal importance of the tasks. In particular, the individual response-time-based augmentation with an abrupt peripheral motion stimulus onset, and augmentation with constantly present peripheral motion stimulus, showed advantage compared to when augmentation was not present. A visual augmentation based on the individual response time was more beneficial in terms of the mean duration on the less important task before switching to the more important one. However, when a rapid detection of peripheral augmentation was crucial, the natural gaze direction was altered being confined toward the center of the visual field indicating the allocation of attentional resources covertly toward the periphery. A continuously moving stimulus in Condition 1 also required attention, however, given a relatively long presentation time of the augmentation, it did not interfere with the gaze direction.

A possible confounding factor affecting performance is the general learning of the task, which could have impacted the performance level. From the trial-by-trial data evaluation, there was only a small and not significant learning effect, thus, the findings of the current study originate indeed mostly from the impact of augmentation. It is also important to note, that in the current experiment, the sub-tasks of the dual-task setting, even though implemented in a dynamic 3D environment, were relatively simple for the participants and involved mostly a motor component. In this task setting, visual augmentation did not significantly capture participants’ attention as no reduction of performance was observed when augmentation was present. However, when the sub-tasks in a dual-task scenario are more cognitively demanding, capturing attentional resources by the augmentation can potentially lead to a drop in the total performance. Beyond the scope of the current work, future studies should address this question. It is also important to mention, that in this experiment, the parameters of the augmentation were selected based on the known dynamics of the task. In real life, the task environments are not always well understood and in those scenarios application of such an augmentation with predefined parameters would be challenging. Nonetheless, such an augmentation could find its application in well-understood environments, such as performing a routine task at work, where the strategy of the user would not be overruled by an algorithm but rather supported by a subtle augmentation system. Furthermore, while the proposed 3D VR-based paradigm is comparable to 2D screen-based experiments in terms of capturing performance in a task-interleaving scenario, it also permits a more natural experimental setting. In particular, it allows free head movement which enables evaluation of natural head and eye movements. This advantage as well as the possibility to implement dynamic interactive tasks makes virtual reality a powerful tool for studying task interleaving in future studies. The results of this study provide an insight into potential visual augmentation designs aiming to improve user’s performance in a challenging dual-task-interleaving setting.

Acknowledgements
This study was conducted with support of the Integrative Augmented Reality (I-AR) intramural funding of the University of Tuebingen, Germany.

Data availability statement
The data for all experiments are available at https://osf.io/p23s7/ and experiment was not preregistered.

ORCID
Olga Lukashova-Sanz http://orcid.org/0000-0002-4053-3894 Siegfried Wahl http://orcid.org/0000-0003-3437-6711 Katharina Rifai http://orcid.org/0000-0002-7218-8547

References
Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. Psychological Science, 14(5), 427–432. https://doi.org/10.1111/1467-9280.01458
Akoumianakis, D., & Stephanidis, C. (2005). Building consensus in human-computer interaction design: Integrated activity-oriented design environments. International Journal of Human-Computer Interaction, 18(1), 85–103. https://doi.org/10.1207/s15327590jhci1801_5
Booth, T., Sridharan, S., McNamara, A., Grimm, C., & Bailey, R. (2013). Guiding attention in controlled real-world environments. In Proceedings of the ACM Symposium on Applied Perception - SAP '13 (pp. 75). ACM Press.
Brumby, D. P., Salvucci, D. D., & Howes, A. (2009). Focus on driving. In Proceedings of the 27th international conference on Human factors in computing systems - CHI '09 (pp. 1629). ACM Press.
Coughlan, J. M., & Miele, J. (2017). ARAVI: AR as an accessibility tool for people with visual impairments. In International Symposium on Mixed and Augmented Reality: (ISMAR) [proceedings]. IEEE and ACM International Symposium on Mixed and Augmented Reality (pp. 288–292).
Dabbish, L., Mark, G., & González, V. M. (2011). Why do I keep interrupting myself?: Environment, habit and self-interruption. In Conference on human factors in computing systems – Proceedings (pp. 3127–3130).
Dey, A., Billinghamurst, M., Lindeman, R. W., & Swan, J. E. (2018). A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. *Frontiers in Robotics and AI*, 5. Article 37. https://doi.org/10.3389/frobt.2018.00037

Farmer, G. D., Jansen, C. P., Nguyen, A. T., & Brumby, D. P. (2018). Dividing attention between tasks: Testing whether explicit payoff functions elicit optimal dual-task performance. *Cognitive Science*, 42 (3), 820–849. https://doi.org/10.1111/cogs.12513

Favela, J., Tentori, M., & Gonzalez, V. M. (2010). Ecological validity and pervasiveness in the evaluation of ubiquitous computing technologies for health care. *International Journal of Human-Computer Interaction*, 26(5), 414–444. https://doi.org/10.1080/10447311003719896

Howes, A., Lewis, R. L., & Vera, A. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review*, 116(4), 717–751. https://doi.org/10.1037/a0017187

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 99–104. https://doi.org/10.1109/MCSE.2007.55

Imaoka, Y., Flury, A., & De Bruin, E. D. (2020). Assessing saccadic eye movements with head-mounted display virtual reality technology. *Frontiers in Psychiology*, 11, Article 572938. https://doi.org/10.3389/fpsyg.2020.572938

Jansen, C. P., & Brumby, D. P. (2010). Strategic adaptation to performance objectives in a dual-task setting. *Cognitive Science*, 34(8), 1548–1560. https://doi.org/10.1111/j.1556-6709.2010.01124.x

Jansen, C. P., & Brumby, D. P. (2015). Strategic adaptation to task characteristics, incentives, and individual differences in dual-tasking. *PLoS ONE*, 10(7), e0130009. https://doi.org/10.1371/journal.pone.0130009

Jansen, C. P., Brumby, D. P., & Garnett, R. (2012). Natural break points. *Journal of Cognitive Engineering and Decision Making*, 6(1), 5–29. https://doi.org/10.1111/j.1553-5444.2011.01323.x

Jansen, C. P., Everaert, E., Hendriksen, H. M. A., Mensing, G. L., Tichelaar, L. J., & Nunner, H. (2019). The influence of rewards on (sub-)optimal interleaving. *PLoS ONE*, 14(3), e0214027. https://doi.org/10.1371/journal.pone.0214027

Jansen, C. P., Gould, S. J., Li, S. Y., Brumby, D. P., & Cox, A. L. (2015). Integrating knowledge of multitasking and interruptions across different perspectives and research methods. *International Journal of Human Computer Studies*, 79, 1–5. https://doi.org/10.1016/j.ijhcs.2015.03.002

Jansen, C. P., & Gray, W. D. (2012). When, what, and how much to reward in reinforcement learning-based models of cognition. *Cognitive Science*, 36(2), 333–358. https://doi.org/10.1111/j.1551-6709.2011.0222x

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, 4 (4), 138–147. https://doi.org/10.1016/S1364-6613(00)01452-2

Lee, F. J., & Taatgen, N. A. (2002). Multitasking as skill acquisition frank. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 572–577. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315782379-134/multitasking-skill-acquisition-frank-lee-niels-taatgen

Lukashova-Sanz, O., & Wahl, S. (2021). Salience-Aware subtle augmentation improves human visual search performance in VR. *Brain Sciences*, 11(3), 283. https://doi.org/10.3390/brainsci11030283

Mckinney, W. (2010). Data structures for statistical computing in Python. In *Proceedings of the 9th Python in Science Conference T (SCIPY 2010)* (pp. 56–61).

Mone, M. A., & Shalley, C. E. (1995). Effects of Task Complexity and Goal Specificity on Change in Strategy and Performance Over Time. *Human Performance*, 8(4), 243–262. https://doi.org/10.1207/s15327043hup0804_1

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Raja, V., & Calvo, P. (2017). Augmented reality: An ecological blend. *Cognitive Systems Research*, 42, 58–72. https://doi.org/10.1016/j.cogsys.2016.11.009

Unity Technologies. (2019). *Unity*. Retrieved December 1, 2019, from https://hadoop.apache.org

Van Der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2), 22–30. https://doi.org/10.1109/MCSE.2011.37

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R., Jones, E., Kern, R., Larson, E., Carey, C. J., van Mulbregt, P.,... Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17 (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

VIVE S@napi SDK. (2021). Retrieved February 1, 2019, from https://hub.vive.com/en-US/download

Wickens, C. (2021). Attention: Theory, principles, models and applications. *International Journal of Human–Computer Interaction*, 37(5), 403–417. https://doi.org/10.1080/10447318.2021.1874741

Zarraonandia, T., Aedo, I., Díaz, P., & Montero Montes, A. (2014). Augmented presentations: Supporting the communication in presentations by means of augmented reality. *International Journal of Human-Computer Interaction*, 30(10), 829–838. https://doi.org/10.1080/10447318.2014.927283

**About the Authors**

Olga Lukashova-Sanz is a physicist with a M. Sc. in physics with a specialization in solid state physics, and a M. Sc. with a specialization in medical technology. She is a PhD student at the ZEISS Vision Science Lab within the Ophthalmic Research Institute and ZEISS Vision Care Business Group.

Siegfried Wahl is a physicist with a Ph.D. in developmental biology of the visual system. He is Honorary Professor of the University of Tübingen and Principal Investigator of the ZEISS Vision Science Lab within the Ophthalmic Research Institute and ZEISS Vision Care Business Group.

Katharina Rifai is a physicist and did her Ph.D. in the field of Visual Neuroscience. She is a member of the ZEISS Vision Science Lab and Advanced Development department within Carl Zeiss Vision International GmbH.