On the impact of VR assessment on the Quality of Experience of Highly Realistic Digital Humans
A Volumetric Video Case Study

Irene Viola · Shishir Subramanyam · Jie Li · Pablo Cesar

Received: date / Accepted: date

Abstract Fuelled by the increase in popularity of virtual and augmented reality applications, point clouds have emerged as a popular 3D format for acquisition and rendering of digital humans, thanks to their versatility and real-time capabilities. Due to technological constraints and real-time rendering limitations, however, the visual quality of dynamic point cloud contents is seldom evaluated using virtual and augmented reality devices, instead relying on prerecorded videos displayed on conventional 2D screens. In this study, we evaluate how the visual quality of point clouds representing digital humans is affected by compression distortions. In particular, we compare three different viewing conditions based on the degrees of freedom that are granted to the viewer: passive viewing (2DTV), head rotation (3DoF), and rotation and translation (6DoF), to understand how interacting in the virtual space affects the perception of quality. We provide both quantitative and qualitative results of our evaluation involving 78 participants, and we make the data publicly available. To the best of our knowledge, this is the first study evaluating the quality of dynamic point clouds in virtual reality, and comparing it to traditional viewing settings. Results highlight the dependency of visual quality on the content under test, and limitations in the way current data sets are used to evaluate compression solutions. Moreover, influencing factors in quality evaluation in VR, and shortcomings in how point cloud encoding solutions handle visually-lossless compression, are discussed.

Keywords Virtual Reality · Point Clouds · Quality of Experience (QoE) · Visual Quality Assessment · Degrees of Freedom (DoF) · Subjective Quality Evaluation

1 Introduction

Recent technological advances in devices for capturing and rendering immersive media contents, together with the fast processing capabilities of commodity hardware, have fostered the development of new applications for Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (XR). Such applications usher a new way to engage and interact with media contents: whereas in traditional 2D consumption, users are passive receivers and have limited possibilities of manipulating the contents they are visualizing, immersive media allow for more interactivity in deciding which content should be displayed by each user. Commonly, immersive media applications can be classified based on the Degrees of Freedom (DoF) that are available to the end user to explore the virtual world: 3DoF refers to the availability of only head rotation as a tool for interaction, as for example in omnidirectional imagining, whereas 6DoF refers to the ability to operate translational movements as well as rotational movements in the 3D space.

In order to populate immersive VR, AR and XR applications, volumetric contents are needed. In this context, point clouds have emerged as a popular format to capture and represent volumetric reconstructions of real-world objects and people, due to their simplicity and versatility. Geometrical representation in point clouds is obtained through a collection of points with x, y and z coordinates in Euclidean space; in addition, attributes such as colour may be included included at each point location. This enables a simple representation

This work is funded by the European Commission H2020 program, under the grant agreement 762111, VRTogether, http://vrtogether.eu/

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Irene Viola and Jie Li
CWI, Amsterdam, The Netherlands
E-mail: irene@cwi.nl, jie.li@cwi.nl
Shishir Subramanyam and Pablo Cesar
CWI, Amsterdam, and TU Delft, Delft, The Netherlands
E-mail: subraman@cwi.nl, p.s.cesar@cwi.nl
that requires no additional pre-processing, is resilient to noise introduced during capture, and enforces no restrictions on the attributes that can be encoded at each point location. However, one main drawback for the deployment of this type of content is the large amount of data that is required in order to produce a photorealistic representation: uncompressed, a single point cloud frame containing one million points requires roughly 20MB to be transmitted. Compression becomes therefore essential for efficient storage and feasible transmission over bandwidth-limited networks. Thus, significant research and industrial effort has been focused on optimizing encoding and transmission, as demonstrated by the ongoing standardization endeavors by bodies such as JPEG and MPEG.

Given the significant storage and bandwidth requirements for dense dynamic point clouds, decisions need to be taken regarding the delivery (type of encoder, bit-rate) to ensure an acceptable quality of experience, depending on the viewing conditions. In a previous paper, we analysed the impact of different viewing conditions in VR environments, namely, with 3DoF or 6DoF locomotion. With this work, we aim to extend our previous analysis by including results and discussions obtained in a baseline viewing condition using traditional 2D screens (2DTV), which is by large the most commonly used environment for user studies for point cloud quality assessment.

In this paper, we report findings obtained in a user study involving 78 participants assessing 72 stimuli based on eight dynamic point clouds sequences depicting humans. Each point cloud sequence was compressed using two encoding solutions at 4 bit-rates, and evaluated in three viewing conditions (6DoF, 3DoF and 2DTV). The gathered data include rating scores, presence questionnaires, simulator sickness reports, along with average watching time. Contributions of the paper are three-fold:

1. An extensive evaluation of the quality of highly realistic digital humans represented as dynamic point clouds in immersive and traditional TV viewing conditions is provided. Existing protocols did not consider the dynamic nature of the point clouds, focused on one type of data set, did not take into account VR viewing conditions, and did not compare VR findings with 2DTV conditions using dynamic contents;
2. Quantitative subjective results about the perceived quality of the contents, along with qualitative insights on what is important for users in interacting with digital humans in VR, are presented. Such results will help in better configuring the network conditions for the delivery of point clouds for real-time transmission, and have implications over ongoing research and standardisation work regarding the underlying compression technology;
3. The collected raw data, which is comprised of rating scores, presence questionnaires, and simulator sickness reports, is made available to the research community, along with scripts to faithfully recreate the stimuli under examination. This will aid reproducibility, while contributing to ongoing research in the area.

The remainder of the paper is organised as follows. Section 2 summarises the related work in the field of point cloud compression and subjective visual quality assessment. Section 3 details the methodology that was followed to conduct the experiments and analyse the data. In Section 4, we report the quantitative and qualitative results of the subjective visual quality assessment, along with commenting the findings in terms of simulator sickness, presence, and interaction time. Key factors and issues for visual quality assessment of dynamic point clouds are discussed in Section 5. Finally, Section 6 concludes the paper. Additional data regarding the statistical analysis of the results is offered in Appendix 6.

2 Related work

2.1 Quality assessment for point clouds

There is a growing interest on subjective quality assessment of point clouds rendered on 2D displays. Zhang et al. evaluated the quality degradation effect of resolution, shape and color on static point clouds. The results indicate that resolution is almost linearly correlated with the perceived quality, and color has less impact than shape on the perceived quality. Zerman et al. compressed two dynamic human point clouds using a state-of-the-art algorithm, and assessed the effects of this algorithm and input point counts on the perceived quality. Their results showed that no direct correlation was found between human viewers’ quality ratings and input point counts. In a recent study, a protocol to conduct subjective quality evaluations and benchmark objective quality metrics were proposed. The viewers passively assessed the quality of a set of static point clouds, as animations with predefined movement path. In a comprehensive work by Alexiou et al., the entire set of emerging point cloud compression encoders developed in the MPEG committee were evaluated through a series of subjective quality assessment experiments. Nine static models, including both humans and objects, were used in the experiments. The experiments provided insights regarding the performance of the encoders and the types of degradation they introduce. Zerman et al. compared the visual quality of point cloud and mesh contents compressed using state-of-the-art algorithms, concluding that, while meshes are more suitable for high bitrate streaming, point cloud compression appears to be more advantageous at lower bitrates. Perry et al. conducted an experiment in 4 laboratories, comparing the latest standardized compression solutions on static point cloud contents.

1 https://github.com/cwi-dis/2DTV_VR_QoE
On the impact of VR assessment on the Quality of Experience of Highly Realistic Digital Humans

Fig. 1: Point Cloud Digital Humans compressed using two point cloud codecs, V-PCC (left) and MPEG anchor (right), at the 4 selected bit-rates.

Only a limited number of point cloud quality assessment studies have been conducted in immersive environments. Mekuria et al. [34] evaluated the subjective quality of their codec performance in a realistic 3D tele-immersive system, in which users were represented as 3D avatars and/or 3D dynamic point clouds, and could navigate in the virtual space using mouse cursor in a desktop setting. Several aspects of quality, such as level of immersiveness, togetherness, realism, quality of motion, were considered. Alexiou and Ebrahimi [6] proposed the use of AR to subjectively evaluate the quality of colorless point cloud geometry. Zerman et al. [53] presented a behaviour analysis of users interacting with colored volumetric media in a AR application. Tran et al. [48] suggested that, in case of evaluating video quality in an immersive setup, aspects such as cybersickness and presence should not be overlooked. Recently, an evaluation of static point cloud contents was conducted in a VR environment [9].

Our work aims at comparing different viewing paradigms for dynamic point clouds, namely 6DoF, 3DoF, and 2DTV conditions. Such a comparison is largely absent in the literature, as previous work has mainly focused on static point cloud contents and single viewing conditions.

2.2 Point cloud compression

A single point cloud frame is represented by an unordered collection of points sampled from the surface of an object. In a dynamic sequence of point clouds, there are no correspondences of points maintained across frames. Thus, detecting spatial and temporal redundancies is often difficult, making point cloud compression challenging. Octrees have been used extensively as a space partitioning structure to represent point cloud geometry [26, 33]. They are a 3D extension of the 2D quadtree used to encode video and images.

Research into point cloud compression can be broadly divided into two categories: model-based and projection-based. The first uses signal processing or deep learning techniques to compress either the geometrical composition of the point cloud, or its attributes, such as color. Zhang et al. [55] proposed a method to compress point cloud attributes using a Graph Fourier Transform. They assume that an octree has been created and separately coded for geometry prior to coding attributes. De Queiroz and Chou [40] used a Region Adaptive Hierarchical Transform (RAHT) to use the colors of nodes in lower levels of the octree to predict the colors of nodes in the next level. In [11], authors adopt techniques from traditional image and video processing, using 3D block prediction in combination with shape-adaptive DCT and graph transforms.

The second category of point cloud codecs aim at projecting the point cloud information onto a 2D canvas, subsequently using legacy image and video compression solutions to encode them. Intra Frame coding in octrees can be achieved by entropy coding the occupancy codes, and then compress the color attributes by mapping them to a 2D grid and using legacy JPEG image compression, as shown in [34]. In 2017, MPEG started a standardization activity to determine a new standard codec for point clouds. They used the codec created by Mekuria et al. [34] as an anchor to evaluate proposals. To encode dynamic point cloud sequences, MPEG has currently standardized one method for dynamic dense point clouds, namely V-PCC [22], and is in the process of standardizing one method for dynamically acquired, sparse point clouds, namely G-PCC [21].

Recently, deep learning solutions for point cloud compressions have been proposed, to encode either geometry or color information, or a combination thereof. Quach et al. [38] propose the use of an auto-encoder to efficiently compress geometry information, and they subsequently analyse the impact of several parameters on the performance [39]. Similarly, Guarda et al. [18] propose a Convolutional Neural Network (CNN) architecture to encode and decode point cloud contents. They further analyse its performance in [17], and extend the work in [19] by employing explicit and implicit quantization. A deeper architecture is proposed in [50], which uses 3D convolutional layers along with variational auto-encoders to achieve favorable compression efficiency. In [5], Alexiou et al. propose a deep learning architecture to encode both geometry and color attributes, and analyse the
performance of various parameters on the coding efficiency and visual quality.

A complete survey of point cloud compression solutions can be found here \[10\]. In our work, we elected to adopt the MPEG Anchor that was used to evaluate the Call for Proposals for the MPEG standardization efforts in point cloud compression \[43\], and the MPEG standard for dynamic point clouds V-PCC \[22\], as they have both been widely used in quality evaluation campaigns in the literature.

3 Methodology

3.1 Dataset Preparation

A dataset of dynamic point cloud sequences was used from the MPEG repository. All sequences were clipped to five seconds and sampled at 30 frames per second. This included point cloud sequences \[12\] captured using photogrammetry (\textit{Longdress, Loot, Red and black, Soldier}, shown in figure \[2\]) and one sequence of a synthetic character sampled from an animated mesh (\textit{Queen}). Four additional point cloud sequences; \textit{Manfred, Despoina, Sarge} (shown in Figure \[2\]) were added for the evaluation. These sequences were created using motion-captured animated mesh sequences.

Keyframes were selected at 30 frames per second and extracted along with the associated mesh materials. Particular care was put in ensuring the selected sequences have the characters facing the user and speaking in their general direction. Then, 1 million points were randomly sampled, independently per key frame to create a consistent groundtruth dataset. The points were sampled from the mesh surface with a probability proportional to the area of the underlying mesh face. This was done to ensure no direct point correspondences across point cloud frames, to mimic realistic acquisition and maintain consistency with the rest of the dataset. The X, Y, Z coordinates of each point was represented using an unsigned integer, as is required for the current version of the V-PCC software. Texture information was encoded as 8-bit RGB.

The contents were compressed using two widely available codecs: the MPEG V-PCC codec in the Release 7.0 \[43\] (\textit{C1}), as the state-of-the-art solution in point cloud compression, and the MPEG anchor codec \[34\] (\textit{C2}), as a baseline release with real-time capabilities. Bitrate points were selected based on the provided presets for \textit{C1}, to ensure fair use. For the additional point cloud sequences for which no configuration file was available, the one provided by MPEG for the \textit{Queen} sequence was used for all the contents. We selected the rate points 1, 3 and 5 from the provided preset V-PCC configurations and extended it with an additional final rate point using a texture Quantization Parameter (QP) of 8, a geometry QP of 12, and an occupancy precision of 2. We re-label the rate points as R1, R2, R3 and R4, respectively. All sequences are encoded using the C2AI (Category 2 All Intra) config. For the photogrammetry sequences, we use the predefined dedicated configuration files for each sequence, at the same rate points.

\textit{C2} is used in an all-intra configuration to match the bit-rates per sequence and rate point (R1-R4) with a tolerance of 10\%, as defined in the MPEG call for proposals. We use an octree depth from 7 to 10 for the rate points R1 to R4, respectively. The highest possible JPEG quantization parameter values were then chosen per sequence, while meeting the target bit rate set using \textit{C1}. An example of content \textit{Loot} encoded with the two compression solutions at the selected rate points is shown in Figure \[1\].

3.2 Experiment setup

All point cloud sequences were rendered using the Unity game engine, by storing all the points of each frame in a vertex buffer, and then drawing procedural geometry on the GPU. The point clouds were rendered using a quadrilateral at each point location with a fixed offset of 0.08 units (this corresponds to a side length of approximately 2mm) around each point (placed at the centre) for all the sequences, to be consistent. In the case of bitrate R1 generated using the MPEG anchor, we increased the offset value to 0.16 by eye, as the resulting point clouds were too sparse (shown in Figure \[1\]). We maintain a fixed frame rate of 30fps throughout the experiment.

Three viewing conditions were selected for comparison: 6DoF, 3DoF, and 2DTV condition. For the 6DoF and 3DoF viewing conditions, participants were asked to wear an Oculus Rift CV1 HMD to view each of the point cloud sequences. For the 3DoF condition, participants were asked to sit on a swivel chair placed at a fixed location in the room and navigate using head movements alone, whereas for the 6DoF condition, participants were allowed to navigate freely within the room. Each sequence was 5 seconds long, after which the playback looped around. We set the background of the virtual room to mid-grey, to avoid distractions. The Oculus Guardian System was used to display in-application wall and floor markers if the participants got too close to the boundary. We used a workstation with 2 GeForce GTX 1080 Ti in SLI for the GPU and an Intel Core i9 Skylake-X 2.9GHz CPU. For the 2DTV condition, the videos were created offline, using the same rendering as the other two viewing condition, and played back to the users using MPV\[2\] (mpv://mpv.io/) and VLC empirical color calibration device according to the following profile: sRGB Gamut, D65 white point, 120 cd/m2 brightness, and minimum black level of 0.2 cd/m2. The test was performed in a room with controlled lighting and mid-grey

\[2\] https://mpv.io/
On the impact of VR assessment on the Quality of Experience of Highly Realistic Digital Humans

3.3 Subjective methodology

To perform the experiments, the subjective methodology Absolute Category Rating with Hidden References (ACR-HR) was selected, according to ITU-T Recommendations P.910 [24]. Participants were asked to observe the video sequences depicting digital humans, and rate the corresponding visual quality on a scale from 1 to 5 (1-Bad, 2-Poor, 3-Fair, 4-Good, and 5-Excellent).

A series of pilot studies were conducted to determine the positioning of digital humans in the virtual space and the length of each sequence, to ensure the sequences were running smoothly within the limited computer RAM. Due to the huge size of the test material, it was not possible to evaluate all 8 point cloud contents in one single session, as long loading times would have brought fatigue to the participants and corrupted the results. Thus, we decided to split the evaluation into two separate tests: one focused on the evaluation of contents obtained from random sampling of meshes (T1: contents Queen, Manfred, Despoina and Sarge), and one focused on contents acquired through photogrammetry (T2: contents Long dress, Soldier, Red and black, and Loot). From each sequence, a subset of frames comprising 5 seconds was selected.

Before the test took place, 3 training sequences depicting examples of 1-Bad, 5-Excellent and 3-Fair were shown to the users to help them familiarize with the viewing condition and test setup, and to guide their rating. Following ITU-T Recommendation BT.500-13 [23], the training sequences were created using one additional content not shown during the test, to prevent biased results. Each content sequence was encoded using the point cloud compression algorithms under test.

For each test and viewing condition, 36 stimuli were evaluated. For each stimulus, the 5 second sequence was played at least once in full, and kept in loop until the participants gave their score. The order of the displayed stimuli was randomized per participant and per viewing condition, and the same content was never displayed twice in a row to avoid bias. Moreover, the presentation order of viewing conditions was randomized between participants, to prevent any confounding effect. Two dummy samples were added at the beginning of each viewing session to ease participants into the task, and the corresponding scores were subsequently discarded.

For the 6DoF and 3DoF viewing conditions, a total of 27 participants were recruited for T1 (12 males, 15 female, average age: 22.48 years old), whereas 25 participants were recruited for T2 (17 males, 8 females, average age: 28.39 years old). The 2DTV viewing condition was conducted separately, with 26 participants for both T1 and T2 (17 males, 9 females, average age: 34.76 years old). All participants were screened for color vision and visual acuity, using Ishihara and Snellen charts, respectively, according to ITU-T Recommendations P.910 [24].

3.4 Data analysis

Outlier detection was performed separately for each of the test datasets T1 and T2, following ITU-T Recommendations
The recommended threshold values of $r_1 = 0.75$ and $r_2 = 0.8$ were used. One outlier was found in test dataset T1, and the corresponding scores were discarded. No outliers were found in the scores collected for T2. Since the 2DTV viewing condition was tested with a different subject population, outlier detection was conducted separately. No outlier was detected for this viewing condition.

After outlier detection, the Mean Opinion Score (MOS) was computed for each stimulus, independently per viewing condition. The associated 95% Confidence Intervals (CIs) were obtained assuming a Student’s t-distribution. Additionally, the Differential MOS (DMOS) was obtained by applying HR removal, following the procedure described in ITU-T Recommendations P913 [25]. Non-parametric statistical analysis was then used to analyze if there are statistical differences among variables, using the MATLAB Statistics and Machine Learning Toolbox, along with the ARTool package in R [15,28].

## 4 Results

### 4.1 Subjective quality assessment

Figures 3 and 4 shows the results of the subjective quality assessment of the contents comprising test T1 and test T2, for the 6DoF, 3DoF and 2DTV viewing conditions. In particular, the DMOS scores associated with the compressed contents are shown with solid lines, along with relative CIs, whereas the dashed lines represent the respective HR scores with the corresponding confidence intervals.

To assess whether significant differences could be found between the three viewing conditions under test, we ran a non-parametric Kruskal-Wallis test, separately for test T1 and T2. The test was chosen as the gathered data was not found to be normally distributed, according to the Shapiro-Wilk normality test (T1: $W = 0.906, p < .001$; T2: $W = 0.909, p < .001$). We found a significant effect of the viewing condition on the scores for test T1 ($\chi^2 = 37.56, p < .001$). Post-hoc analysis using Mann-Whitney U test with Bonferroni correction ($\alpha = .05/3$) revealed significant differences between 6DoF and 2DTV viewing conditions ($p < .001, r = 0.09$), and between 3DoF and 2DTV viewing conditions ($p < .001, r = 0.14$), but not between 6DoF and 3DoF ($p = 0.101, r = 0.04$). Values seem to indicate an effect of VR viewing condition with respect to traditional TV viewing on the final scores; however, the effect sizes suggest that, if the effect indeed exists, it is small. For test T2, no significant effect was found for viewing condition on the scores ($\chi^2 = 5.19, p = 0.075$).

It can be observed that codec CI has generally a more favorable performance with respect to C2. This is especially evident for the contents acquired through photogrammetry (see Fig. 4), for which the gap among the two codecs is more pronounced. Mann-Whitney U test confirmed statistical significance for the two codecs (T1: $Z = 6.60, p < .001$, T2: $Z = 22.06, p < .001$), albeit with notably different effect sizes between test T1 and T2 ($r = 0.13$ and $r = 0.45$, respectively).

A Kruskal-Wallis test performed on the scores revealed a significant effect of the content on the final scores, for both sets of contents (T1: $\chi^2 = 64.91, p < .001$, T2: $\chi^2 = 35.23, p < .001$). Table 1 shows the results of the post-hoc test conducted using Mann-Whitney U test with Bonferroni correction ($\alpha = .05/6$). Contents Long dress - Loot and Red and black - Soldier all show statistical significance with respect to content Queen ($p < .001, r \geq 0.15$ for all pairs). For contents acquired through photogrammetry, statistical significance was found between contents Loot and Soldier, and Loot and Red and black ($p < .006, r = 0.07$). Results suggest that contents Red and black appear to be given different scores with respect to contents Loot and Soldier. However, the effect sizes suggest that the effect, if existent, is quite small.

We also ran a Kruskal-Wallis test on the scores to assess whether the selected bit-rates were showing statistical significance. Results confirmed that the bit-rates have a significant effect, if existent, is quite small. For test T1 and T2 ($p < .001, r > 0.20$ for all pairs).

### 4.2 Data analysis

#### 4.2.1 T1

In order to further analyze the effect of DoF conditions, contents, codecs and bit-rates, and relative interactions, on the

|                | Z   | p   | r   |
|----------------|-----|-----|-----|
| **T1**         |     |     |     |
| Manfred - Sarge| -1.10| 0.270| 0.03 |
| Manfred - Despoina| -1.59| 0.111| 0.04 |
| Manfred - Queen | 5.57 | <.001| 0.15 |
| Sarge - Despoina| -0.54| 0.588| 0.01 |
| Sarge - Queen   | 6.75 | <.001| 0.18 |
| Despoina - Queen| 7.04 | <.001| 0.19 |

|                | Z   | p   | r   |
|----------------|-----|-----|-----|
| **T2**         |     |     |     |
| Long dress - Loot| -5.19| <.001| 0.14 |
| Long dress - Red and black | -0.05| 0.960| 0.00 |
| Long dress - Soldier | -2.73| 0.006| 0.07 |
| Loot - Red and black | 4.99 | <.001| 0.14 |
| Loot - Soldier   | 2.08 | 0.037| 0.06 |
| Red and black - Soldier | -2.61| 0.009| 0.07 |
On the impact of VR assessment on the Quality of Experience of Highly Realistic Digital Humans

---

![Graphs of DMOS against achieved bit-rate for Manfred, Sarge, Despoina, and Queen.](image)

Fig. 3: DMOS against achieved bit-rate. HR scores are shown using a dashed plot. Each column represents a content in test T1, whereas the rows depict results obtained using the viewing conditions 6DoF, 3DoF and 2DTV, respectively.

Table 2: Pairwise post-hoc test on the bitrates for test T1 and T2, using Wilcoxon signed-rank test with Bonferroni correction.

|       | Z   | p    | r    |
|-------|-----|------|------|
| R1 - R2 | -14.30 | <.001 | 0.41 |
| R1 - R3 | -25.54 | <.001 | 0.73 |
| R1 - R4 | -27.36 | <.001 | 0.78 |
| R2 - R3 | -17.03 | <.001 | 0.49 |
| R2 - R4 | -21.05 | <.001 | 0.60 |
| R3 - R4 | -7.30  | <.001 | 0.21 |

For test T1, analysis of deviance on the full mixed-effects model showed significance for main effects Content ($F = 53.98, df = 3, p < .001$), Codec ($F = 65.19, df = 1, p < .001$) and bit-rate ($F = 595.38, df = 3, p < .001$), but not for DoF ($F = 0.07, df = 2, p = 0.936$). Moreover, significant interaction effects were found for DoF - Content ($F = 11.39, df = 6, p < .001$), Content - Codec ($F = 9.97, df = 3, p < .001$), Content - bit-rate ($F = 6.04, df = 9, p < .001$) and Codec - bit-rate ($F = 8.96, df = 3, p < .001$). No interaction effect beyond the first level was found to be significant. The full results of the analysis of deviance can be found in Table 3.

For test T2, analysis of deviance on the full mixed-effects model showed significance for main effects Content ($F = 53.98, df = 3, p < .001$), Codec ($F = 65.19, df = 1, p < .001$) and bit-rate ($F = 595.38, df = 3, p < .001$), but not for DoF ($F = 0.07, df = 2, p = 0.936$). Moreover, significant interaction effects were found for DoF - Content ($F = 11.39, df = 6, p < .001$), Content - Codec ($F = 9.97, df = 3, p < .001$), Content - bit-rate ($F = 6.04, df = 9, p < .001$) and Codec - bit-rate ($F = 8.96, df = 3, p < .001$). No interaction effect beyond the first level was found to be significant. The full results of the analysis of deviance can be found in Table 3.

For test T1, analysis of deviance on the full mixed-effects model showed significance for main effects Content ($F = 53.98, df = 3, p < .001$), Codec ($F = 65.19, df = 1, p < .001$) and bit-rate ($F = 595.38, df = 3, p < .001$), but not for DoF ($F = 0.07, df = 2, p = 0.936$). Moreover, significant interaction effects were found for DoF - Content ($F = 11.39, df = 6, p < .001$), Content - Codec ($F = 9.97, df = 3, p < .001$), Content - bit-rate ($F = 6.04, df = 9, p < .001$) and Codec - bit-rate ($F = 8.96, df = 3, p < .001$). No interaction effect beyond the first level was found to be significant. The full results of the analysis of deviance can be found in Table 3.

For test T2, analysis of deviance on the full mixed-effects model showed significance for main effects Content ($F = 53.98, df = 3, p < .001$), Codec ($F = 65.19, df = 1, p < .001$) and bit-rate ($F = 595.38, df = 3, p < .001$), but not for DoF ($F = 0.07, df = 2, p = 0.936$). Moreover, significant interaction effects were found for DoF - Content ($F = 11.39, df = 6, p < .001$), Content - Codec ($F = 9.97, df = 3, p < .001$), Content - bit-rate ($F = 6.04, df = 9, p < .001$) and Codec - bit-rate ($F = 8.96, df = 3, p < .001$). No interaction effect beyond the first level was found to be significant. The full results of the analysis of deviance can be found in Table 3.

---

Gathered scores, we fitted a full linear mixed-effects model on the data, accounting for randomness introduced by the participants. Due to the non-normality of our data, the aligned rank transform was applied prior to the fitting [51]. Since the transform is designed for a fully randomized test, it is not suitable for the scores collected during the test, as the HR addition makes the design matrix rank deficient. However, the transform can be applied to the differential scores used to obtain DMOS, as it follows a fully randomized design. Thus, it was decided to perform the analysis on the differential scores. Post-hoc contrast tests were performed using the ART-C tool [15].

For test T1, analysis of deviance on the full mixed-effects model showed significance for main effects Content ($F = 53.98, df = 3, p < .001$), Codec ($F = 65.19, df = 1, p < .001$) and bit-rate ($F = 595.38, df = 3, p < .001$), but not for DoF ($F = 0.07, df = 2, p = 0.936$). Moreover, significant interaction effects were found for DoF - Content ($F = 11.39, df = 6, p < .001$), Content - Codec ($F = 9.97, df = 3, p < .001$), Content - bit-rate ($F = 6.04, df = 9, p < .001$) and Codec - bit-rate ($F = 8.96, df = 3, p < .001$). No interaction effect beyond the first level was found to be significant. The full results of the analysis of deviance can be found in Table 3.

For test T2, analysis of deviance on the full mixed-effects model showed significance for main effects Content ($F = 53.98, df = 3, p < .001$), Codec ($F = 65.19, df = 1, p < .001$) and bit-rate ($F = 595.38, df = 3, p < .001$), but not for DoF ($F = 0.07, df = 2, p = 0.936$). Moreover, significant interaction effects were found for DoF - Content ($F = 11.39, df = 6, p < .001$), Content - Codec ($F = 9.97, df = 3, p < .001$), Content - bit-rate ($F = 6.04, df = 9, p < .001$) and Codec - bit-rate ($F = 8.96, df = 3, p < .001$). No interaction effect beyond the first level was found to be significant. The full results of the analysis of deviance can be found in Table 3.

Post-hoc interaction analysis using ART-C revealed significant differences at 5% level in 3DoF for content pairs Manfred - Queen ($p < .001$), Sarge - Queen ($p < .001$) and Despoina - Queen ($p < .001$); in 6DoF, for content pairs Man-
Fig. 4: DMOS against achieved bit-rate. HR scores are shown using a dashed plot. Each column represents a content in test T2, whereas the rows depict results obtained using the viewing conditions 6DoF, 3DoF and 2DTV, respectively.

Table 3: Analysis of deviance on the full mixed-effects model, for test T1.

|                  | F    | df | p    |
|------------------|------|----|------|
| DoF              | 0.07 | 2  | 0.936|
| Content          | 53.98| 3  | < .001|
| Codec            | 65.19| 1  | < .001|
| Bitrate          | 595.38| 3  | < .001|
| DoF: Content     | 11.39| 6  | < .001|
| DoF: Codec       | 0.58 | 2  | 0.660|
| Content: Codec   | 9.97 | 3  | < .001|
| DoF: Bitrate     | 1.70 | 6  | 0.116|
| Content: Bitrate | 6.04 | 9  | < .001|
| Codec: Bitrate   | 8.96 | 3  | < .001|
| DoF: Content: Codec | 0.71 | 6  | 0.643|
| DoF: Content: Bitrate | 0.98 | 18 | 0.480|
| Content: Codec: Bitrate | 0.73 | 6  | 0.627|
| DoF: Content: Codec: Bitrate | 1.23 | 9  | 0.273|

Fred - Sarge (p < .001) and Despoina - Queen (p < .001); in the 2DTV condition, for content pairs Manfred - Sarge (p = 0.019), Manfred - Queen (p < .001), Sarge - Despoina (p < .001), Sarge - Queen (p < .001), and Despoina - Queen (p = 0.024). Additionally, several content pairs exhibited significant contrasts when different DoF were employed; for a full report of the contrasts, we invite readers to consult the appendix. Most notably, no significant effect was found when the same content was displayed in different DoF mediums; that is, for every content under exam, the pairs 3DoF - 6DoF, 3DoF - 2DTV, and 6DoF - 2DTV were consistently above the 5% threshold of significance.

Regarding the interaction between factors Content and Codec, post-hoc analysis using ART-C showed significant differences at 5% level between the two codecs for all contents (p < .001) except Queen, for which the two codecs were deemed equivalent (p = 0.995). Furthermore, statistically significant differences were found, for C1, in content pairs Manfred - Despoina (p = 0.006), and Despoina - Queen (p < .001), and for C2, in content pairs Manfred - Despoina (p = 0.010), Manfred - Queen (p < .001), Sarge - Queen (p < .001), and Despoina - Queen (p < .001). Several content pairs were found to have significant contrasts when different codecs were employed; for a complete overview, we refer readers to the appendix.
Table 4: Analysis of deviance on the full mixed-effects model, for test T2.

|                  | F   | df | p   |
|------------------|-----|----|-----|
| DoF              | 2.59| 2  | 0.082 |
| Content          | 165.56| 3  | < .001 |
| Codec            | 1059.81| 1  | < .001 |
| Bitrate          | 702.54| 3  | < .001 |
| DoF: Content     | 1.99| 6  | 0.064 |
| DoF: Codec       | 1.08| 2  | 0.340 |
| Content: Codec   | 5.81| 3  | < .001 |
| DoF: Bitrate     | 0.39| 6  | 0.881 |
| Content: Bitrate | 6.30| 9  | < .001 |
| Codec: Bitrate   | 44.89| 3  | < .001 |
| DoF: Content: Codec | 0.80| 6  | 0.569 |
| DoF: Content: Bitrate | 1.10| 18 | 0.341 |
| DoF: Codec: Bitrate | 1.12| 6  | 0.346 |
| Content: Codec: Bitrate | 2.08| 9  | 0.028 |
| DoF: Content: Codec: Bitrate | 0.69| 18 | 0.820 |

Significant differences at 5% level were also found when considering post-hoc interactions between factors Content and Bitrate. In particular, the pair Manfred - Queen was found to have significant differences for rate R2 (p < .001) and R3 (p = 0.018); pair Serge - Queen for rate R1 (p = 0.033), R2 (p < .001), and R3 (p = 0.005); pair Despoina - Queen, for rate R1 (p < .001), R2 (p < .001), and R3 (p < .001). No significant difference was found among contents at rate R4, indicating that, at the highest quality level, the contents were rated similarly. As expected, most of the comparisons between different bitrates, be it with the same or among different contents, yield statistically significant differences; the complete results are available in the appendix. Finally, results of the post-hoc analysis of interactions between factors Codec and Bitrate showed statistically significant differences at 5% level among the codecs for rate R1 (p < .001) and R3 (p = 0.007), but not for rate R2 (p = 0.052 and R4) (p = 0.986). In the first case, the p-value is quite close to significance, whereas in the second, the test confirms our previous observations: at the highest quality level, the difference among codecs seems to be imperceptible. The rest of the combinations between codecs and bitrates lead to significant differences, with the exception of C1-R3 versus C2-R4 (p = 0.054), indicating that the null hypothesis cannot be rejected for C2 at rate R4 when compared to C1 at rate R3; in other words, to achieve similar ratings to codec C1 (MPEG V-PCC), codec C2 (the MPEG anchor) requires higher bandwidth, which is in line with what observed in the previous section. More exhaustive results can be found in the appendix.

4.2.2 T2

Results of analysis of deviance on the full mixed-effects model for test T2 showed significance for main effects Content (F = 165.56, df = 3, p < .001), Codec (F = 1059.81, df = 1, p < .001) and Bitrate (F = 702.54, df = 3, p < .001), but not for DoF (F = 2.59, df = 2, p = 0.0825), similarly to what was seen for test T1. Two-way interactions were found significant at 5% level between Content and Codec (F = 5.81, df = 3, p < .001), Content and Bitrate (F = 6.30, df = 9, p < .001), and Codec and Bitrate (F = 44.89, df = 3, p < .001), as well as the three-way interaction between Content, Codec, and Bitrate (F = 2.08, df = 9, p = 0.028).

Post-hoc interaction analysis using ART-C for the three-way interaction between factors Content, Codec, and Bitrate revealed significant differences at 5% level between the two codecs under exam, when fixing content and bitrate level, for touples involving content Longdress at bitrate R1 (p < .001), R2 (p < .001), and R3 (p < .001), but not for the highest bitrate R4 (p = 0.964). Similarly, for content Soldier, touples had significant interactions at rate R1 (p = 0.010), R2 (p < .001), and R3 (p < .001), but not R4 (p = 0.085). For the other two contents, all touples at same bitrate were significant (Loot, C1, R1 - Loot, C2, R1: p < .001; Loot, C1, R2 - Loot, C2, R2: p < .001; Loot, C1, R3 - Loot, C2, R3: p < .001; Loot, C1, R4 - Loot, C2, R4: p = 0.016; Red and black, C1, R1 - Red and black, C2, R1: p < .001; Red and black, C1, R2 - Red and black, C2, R2: p < .001; Red and black, C1, R3 - Red and black, C2, R3: p < .001; Red and black, C1, R4 - Red and black, C2, R4: p = 0.047). This indicates that, with the exception of Longdress and Soldier at bitrate R4, for all rate points and all contents the two codecs were significantly different at 5% level.

Post-hoc interaction also revealed statistical difference among different contents: considering codec C1 at bitrate R1, significant differences were found in content touples Longdress - Loot (p < .001), Longdress - Soldier (p < .001), Loot - Soldier (p = 0.004), and Red and black - Soldier (p < .001), but not for touples Longdress - Red and black (p = 0.626) and Loot - Red and black (p = 0.091); analogously, at bitrate R2, differences were found for touples Longdress - Loot (p = 0.004), Longdress - Soldier (p < .001), Loot - Soldier (p < .001), and Red and black - Soldier (p < .001), but not for touples Longdress - Red and black (p = 0.198) and Loot - Red and black (p = 1). At bitrate R3, significant differences were found only between touples Longdress - Soldier (p < .001); at bitrate R4, all content touples failed to reject the null hypothesis. This indicates that differences in DMOS scores among different contents are more prominent at lowest bitrates for codec C1, whereas for higher bitrates, contents received similar scores. When considering codec C2, at bitrate R1, significant differences were observed for content touples Longdress - Loot (p < .001), Longdress - Soldier (p < .001), Loot - Red and black (p = 0.047), and Red and black - Soldier (p = 0.004), but not for touples Longdress - Red and black (p = 0.999) and Loot - Soldier (p = 1); similarly, at bitrate R2, differences were deemed significant for content touples Longdress - Loot (p < .001),
Longdress - Soldier (p < .001), Loot - Red and black, (p < .001), and Red and black - Soldier (p < .001), but not for Longdress - Red and black (p = 0.994) and Loot - Soldier (p = 1). This seems to indicate that, a lower bitrates, contents presenting the same gender were rated similarly. At bitrate R3, statistical difference could be observed among touples Longdress - Loot (p < .001), Longdress - Red and black (p = 0.018), Longdress - Soldier (p < .001), Loot - Red and black (p = 0.009), and Red and black - Soldier (p < .001), but not for touple Loot - Soldier (p = 1), whereas for bitrate R4, statistical differences at a significant level were only observed between Longdress - Soldier (p = 0.015). Results indicate that for the highest bitrate, similar trends can be generally observed between different contents for both codecs under consideration, whereas as bitrate decreases, more differences can be spotted in DMOS trends.

Finally, post-hoc interaction analysis revealed significant differences at 5% level, for content Longdress encoded with codec C1, between bitrate R1 with respect with all the other bitrates (R1 - R2: p < .001; R1 - R3: p < .001; R1 - R4: p < .001); however, no significant difference in rating was found when comparing R2, R3, and R4. Remarkably, different trends can be observed for codec C2, for which only bitrates R3 - R4 are to be considered statistically equivalent (p = 0.055), whereas all other touples for content Longdress present significant differences (R1 - R2: p = 0.005; R1 - R3: p < .001; R1 - R4: p < .001; R2 - R3: p < .001; R2 - R4: p < .001). For content Loot, statistical differences are observed, for codec C1, among all bitrates, bar R2 - R3 (p = 0.142) and R3 - R4 (p = 1); the rest falls below the significance level (R1 - R2: p < .001; R1 - R3: p < .001; R1 - R4: p < .001; R2 - R4: p = 0.004). For codec C2, however, a different trend is observed: bitrates R1 - R2 are the only tuple, for Loot, for which no statistical difference is found (p = 0.377), whereas all other cases present statistically significant differences (p < .001 for all tuples). Content Red and black exhibits similar behaviour as Longdress: for codec C1, only differences among bitrate R1 with all the other bitrates are significant (R1 - R2: p < .001; R1 - R3: p < .001; R1 - R4: p < .001), whereas for codec C2, all bitrate tuples are statistically different at 5% significance level (R1 - R2: p = 0.018; R1 - R3: p < .001; R1 - R4: p < .001; R2 - R3: p < .001; R2 - R4: p < .001; R3 - R4: p < .001). For content Soldier, significant differences were found, for codec C1, for all bitrate tuples, bar R3 - R4 (p = 0.648; for all other bitrates, p < .001); conversely, for codec C2, the only bitrates who did not exhibit significant differences were R1 - R2 (p = 0.373; for all other bitrates, p < .001). This seems to indicate that generally, for codec C1 statistical differences in score distributions are usually found between lower bitrates, whereas the highest bitrates have a more uniform behaviour; however, for codec C2, more differences can be found across score distribution for all the bitrates, bar certain cases (such as Loot and Soldier) for which the lowest bitrates are deemed equivalent. Several other combinations of the three factors under exam were deemed statistically significant, more than what we could cover in this pages: we refer interested readers to the appendix for a full coverage of the post-hoc interaction results.

4.3 Additional questionnaires and interaction data

4.3.1 IPQ & SSQ Questionnaires

For T1 and T2, the collected IPQ data under each subscale are all normally distributed as examined by the Shapiro-Wilk test (p > 0.05). A paired sample t-test was applied to check the differences between 3DoF and 6DoF in terms of SP, INV, REAL and G. For T1, there was a significant difference in SP between 3DoF (M=4.13, SD=0.92) and 6DoF (M=5.04, SD=0.67), t(26)=-4.44, p < .001, Cohen’s d = 0.52 and also a significant difference in G between 3DoF (M=4.11, SD=1.28) and 6DoF (M=4.96, SD=1.13), t(26)= -2.60, p < .01, Cohen’s d = 0.64. For T2, SP was also significantly different in 3DoF (M=4.16, SD=1.17) and 6DoF (M=4.83, SD=1.12), t(24)= 3.48, p < .01, Cohen’s d = 0.45 and so was G between 3DoF (M=4.20, SD=1.61) and 6DoF (M=5.08, SD=1.19), t(24)= 3.56, p < .01, Cohen’s d = 0.71. Other factors showed no significant differences between 3DoF and 6DoF in both T1 and T2.

With respect to SSQ, no significant differences (p > 0.05) were found between 3DoF and 6DoF in terms of cybersickness. We further tested whether there were order effects in experiencing cybersickness, where half of the participants started with 6DoF as the first condition and 3DoF as the second, and the remainder the inverse. No significant differences (p > 0.05) were found for any order effects in experiencing cybersickness.
4.3.2 Interaction time

Interaction time was found to be strongly correlated with MOS values in a study conducted on light field image quality assessment [49], as well as in studies conducted with point cloud contents in interactive environments [9]. In particular, it was found that users tended to spend more time interacting with contents at high quality, whereas for low quality scores, less time was spent looking at the contents. In order to see whether similar trends could be observed in our data, we compared the average time spent watching the sequence in 3DoF and 6DoF, separately for each quality score given by the participants. Since no interactivity was present in the 2DTV condition, the data was discarded for the analysis. Results are shown in Fig. 5. A positive trend can be observed between the given score and the average time spent looking at the sequence, with the exception of score 5, which for test T2 shows a negative trend with respect to the time. However, it should be considered that on average, a small percentage of scores equal to 5 were given in test T2 (10% of the total scores), thus, variations may be due to the difference in sample size. It is also worth noting that, on average, participants spent more time looking at the sequences in 6DoF, with respect to the 3DoF case. Indeed, several participants pointed out that the lowest scores were the fastest to be given, whereas for higher quality, it was harder to decide on the rating.

4.3.3 Interviews

The interviews were only conducted for the VR conditions, due to time limitations. We asked the same interview questions for T1 and T2; so, we combined the interview transcripts of 52 participants (T1=27, T2=25). Participants were labelled as T1P1-T1P27 or T2P1-T2P25. The categorized answers are presented as follows:

Factors considered when assessing quality. 57% of the participants mentioned that they assessed the quality based on three criteria: 1) overall outline and pattern distortion on body and on clothes, 2) natural gestures and movements of the digital humans, and 3) visual artifacts such as blockiness, blurriness, and extraneous floating artifacts. As T2P3 commented, “I paid attention to the blurriness of the clothing patterns, the fingers, and whether the gestures or movements were smooth.” 48% of the participants mentioned the quality assessment criteria are content related, who agreed that it is easier to spot artifacts for the content with complex patterns (e.g., Long dress) and dominant colors (e.g., Red and black) than the content with uniformed colors (e.g., Soldier and Sarge). T2P5 said, “The two ladies were easier to assess the visual quality, because their clothes have strong colors. The man playing keyboard or something is quite difficult. His clothes were mainly monotone.” 46% of the participants considered facial expressions as an unignorable factor for quality assessment, which they believe is an important cue for social connectedness. For example, T1P15 mentioned, “The robotic lady was static and had no [facial] expression at all. It was difficult to tell the difference [between different quality levels].” Similarly, T2P8 said, “If I could see her [the lady with red dress] smile and her teeth, I gave the score of fair to good.” For the extraneous floating artifacts (e.g., bubbles flickering outside the digital humans), 23% found it very annoying and lowered the overall quality for the content, but a few participants (8%) thought these artifacts do not influence their quality judgement. T1P4 commented, “The flickering blocks were annoying and distracting at the beginning, but later I got used to them. It felt like watching an old movie.”

Difficulties in assessment. 42% of the participants pointed out the difficulties in assessing the quality, especially for the high quality contents, which are not perfect and still have missing details like blurry faces or wrong fingers. 15% of the participants specifically pointed out that it is difficult to distinguish between quality level 3 to 5. As T2P21 mentioned, “It was difficult to give [score] 5, the best ones were still missing many details like fingers or feet. I was hesitating all the time between [score] 4 and 5.” 17% of the participants commented that it gradually became easier in rating the quality when they adapted to the contents. So, the second viewing condition was easier for them. For example, T1P12 said, “I noticed that the rating got easier as I got familiar to the quality levels, and had seen the best and the worst qualities.”

Comparison between 3DoF and 6DoF. 52% of the participants preferred 6DoF, because it allowed them to move closer to examine the details (e.g., shoes and fingers). They felt more realistic when walking in the virtual space. However, they also commented that 3DoF offered a fixed distance between them and digital humans, enabling a more stable and focused assessment. For example, T2P13 commented, “Walking around [6DoF] allowed me to get really close [to the digital humans] and see more details, like pixel sizes, shoes, fingers. It felt more realistic. Sitting down [3DoF] has a distance [between me and the digital humans]. Sometimes, I found it difficult to assess the quality.” 21% of the participants preferred relaxation and passiveness in 3DoF, because they did not find much differences between 3DoF and 6DoF in terms of quality assessment, but they found 3DoF is less nauseous than 6DoF. As T1P13 mentioned, “I felt more focused, secured and relaxed when sitting down, I got worried to be trapped by the cables. I also felt more dizzy when I walked around [in VR].”
5 Discussion

5.1 Testing in VR

Results of our experiment indicate a very small, if not negligible, effect of viewing condition on the distribution of the scores. In particular, the viewing condition was deemed to have a significant effect on the distribution of the scores for the first set of contents we tested, revealing significant differences between VR testing and the 2DTV counterpart, with small effect sizes; for the second set of contents, no significant difference was spotted. The implications of it are twofold. On one hand, results show that the score distributions follow similar trends in VR with respect to passive video consumption, thus confirming the validity of testing volumetric contents on traditional 2D screens, as it has been done for the majority of tests performed in the literature. More specifically, no significant interaction was found, for both tests, between viewing condition and codec under exam, indicating that differences among compression solutions are not affected by the choice of displaying device and interaction paradigm. However, the interaction between viewing condition and contents was found to be significant for test T1: in particular, results indicate that differences among contents vary depending on the visualization medium. This might be due to a variety of factors: the possibility of interacting with the content, moving closer to inspect details; the presence of a fixed viewing point, which allows for easier comparison; the absence of confounding factors such as simulator-induced sickness or novelty effect. Particular care should be adopted in choosing contents depending on the type of test that needs to be carried, making sure that artifacts are visible at the distance that is selected for passive viewing, for example.

The second implication involves the constituents of quality of experience. The MOS of visual quality is only one of the factors that influences the quality of experience of a given user; other factors, such as presence and immersion, are equally important in determining the enjoyment of a given user using the system. Even though small or no effect of viewing condition on the MOS distribution was found, results of the IPQ revealed a strong effect of viewing condition on spatial presence and general sense of being there. Such factors should be taken into consideration when designing new experiments: if visual quality is the main constituent that needs to be assessed, traditional screens might be substituted (with caution) for VR assessment; however, if other factors need to be evaluated, such choice might have a larger impact.

5.2 Datasets

Despite the rich literature in point cloud acquisition and compression, few point cloud datasets are publicly available. This is especially true when considering point cloud datasets depicting photo-realistic humans. One of the most popular and widely used full-body dataset, created by 8i Labs [12], consists of only 4 individual contents, whereas the HHI Fraunhofer dataset has 1 individual content [13]. In the context of point cloud compression, such scarcity of available data may lead to compression solutions being designed, optimized and tested while considering a considerably narrow range of input data, thus leading to algorithms that are overfitted to the specifics of the acquisition method used to obtain the contents. The consequences of such a scenario are reflected in our results. Whereas for the contents assessed in test T2 a large difference was observed between codec V-PCC and the MPEG anchor, for the contents in test T1 the gap was markedly lower, and indeed the significance of the effect of the codec selection had a smaller effect size for test T1 with respect to test T2, as seen in section 4.1. Test T2 consisted of contents that had been used in multiple quality assessment experiments [17, 37, 44, 47], notably including the performance evaluation of the upcoming MPEG standard [45]. On the other hand, test T1 included contents that have not been used so far in assessment of point cloud compression solutions. The discrepancies in the results of the subjective quality assessment campaign indicate that performance gains may vary considerably when new contents are evaluated. A larger body of contents depicting digital humans, involving several acquisition technologies, is needed in order to properly design, train and evaluate new compression solutions in a robust way.

5.3 Personal preferences and bias

Subjective evaluation experiments are complicated by many aspects of human psychology and viewing conditions, such as participants’ vision ability, translation of quality perception into ranking scores, adaptations and personal preferences for contents. Through carefully following the ITU-T Recommendations P.913 [25], we are able to control some of the aspects. For example, eliminate the scores given by the participants with vision problems; train participants to help them understand the quality levels; randomize the stimuli and viewing conditions to minimize the order effects. However, we noticed that personal preferences towards certain contents are difficult to control. Satgunam et al. [41] found that their participants were divided into two preference groups: prefer sharper content versus smoother content. Similarly, Kortum and Sullivan [30] found that the “desirability” of participants had an impact on video quality responses, with a more desirable video clip being given a higher rating. In our experiments, content Queen is generally given lower ratings with respect to the other contents. In the interviews, 40% of the participants showed their preference towards Soldier, due to his high-resolution facial features, unitoned clothes
and natural movements, whereas 27% expressed dislike towards *Queen*, because of her lifeless look and static gestures. Body motion has been shown to be important in increasing the naturalness of virtual characters, especially when involving complex motions [35]. Moreover, several studies have highlighted the importance of matching the appearance of the avatar to the naturalness of the motion, indicating that “appropriateness” between visual presentation and gesture plays an important role in realism [16][31]. Our findings suggest that realism and naturalness of interaction might have an impact on the perceived visual quality of the contents as well. Quality assessment may need to be adjusted based on content and viewer preferences, and offering training with different contents, as well as account for factors such as realism, naturalness of motion, and uncanny valley effect.

5.4 Technological constraints and limitations

The two codecs used in this experiment introduce different distortions during compression. As the MPEG anchor codec uses the octree data structure to represent geometry, the number of points in the decoded cloud varies exponentially based on the tree depth. Thus, at lower bitrates, the decoded point clouds are quite sparse, and when the point size is increased to make them appear watertight, they have a block-y appearance. However, the low delay encoding and decoding of this codec makes it suitable for real time applications such as social VR. On the other hand, the V-PCC codec leverages existing 2D video codecs to compress both geometry and color, which introduces noise in terms of extraneous objects, and general geometric artifacts such as misaligned seams (see Figure 1). However, the approach yields better results at low bitrates, as demonstrated in our results. The codec is optimized for human perception of 2D video, which might lead to undesired results when applied to 3D objects in VR. The mapping from 3D to 2D is critical to codec performance, which explains why the encoding phase has high complexity. Decoding has a lower delay, as it benefits from hardware acceleration of video decoders on GPUs, making this approach suitable for on-demand streaming.

One of the main shortcoming of both compression solutions lies in their inability to reach visually-lossless quality, as demonstrated by our results. Achieving a visually pleasant result is of paramount importance for the market adoption of the technology; indeed, poor visual quality might lead consumers to tune off from the experience altogether [1]. Visual perception should be taken into account when designing compression solutions, especially at high bitrates, to ensure that in absence of strict bandwidth constraints, excellent quality can be achieved.

5.5 Rendering environment considerations

Previous research on subjective assessment of dynamic point cloud contents has been primarily conducted in desktop setups [2][52], whereas VR/AR technologies have been employed with static content [6][8], or limited amount of dynamic contents [53]. However, placing dynamic contents to be rendered in a VR/AR environment in real time, in order for users to interact, adds several technological constraints. In previous research on subjective assessment of point cloud content such as the MPEG standardisation activity [2], participants were asked to view a video of the point cloud sequence rendered from a predefined camera path. The same approach was used by van der Hooft et al. to subjectively and objectively assess the quality of adaptive streaming for point cloud contents [29]; however, the influence of camera paths on objective quality assessment was shown to be significant in [45]. In order to allow users to interact with the content in a VR/AR environment, the dynamic point clouds need to be rendered in real time. This is, however, extremely resource intensive, and poses many technical constraints. For our test in particular, the point clouds needed to be stored as uncompressed files, to avoid confounding factors with the compression solutions under evaluation. The point cloud files were stored as binary PLY files to allow faster read from disk. Yet, real-time progressive loading of the sequences was not possible, as the loading operation was interfering with the rendering, thus leading to drops in the frame rate. Waiting for each sequence to be loaded, on the other hand, was unattainable, as the read time from disk would mean long waiting times between one sequence and another, thus adding to subjects’ fatigue. To fix the issue, we decided to load all sequences in physical memory before the test. However, this impacted the amount of sequences that could be tested in one session, as well as the length of each. New systems for rendering point clouds in real time, while respecting the constraints introduced by quality assessment scenarios, should be developed in order to foster research in the field.

5.6 Protocols for subjective assessment in VR

Choosing the right methodology to follow in order to collect users’ opinions is a delicate matter, as it can influence the statistical power of the collected score, and in some cases lead to difference in results. Single stimulus methodologies, in particular, lead to larger CIs with respect to double stimulus methodologies, and are more subject to be influenced by individual content preference [25]. An early study comparing single and double stimulus methodologies for the evaluation of colorless point cloud contents indicated that the latter was more consistent in recognizing the level of impairment, as relative differences facilitate the rating task [3]. However, the study pointed out that the single stimulus methodology shows
more discrimination power for compression-like artifacts, albeit at the cost of wider CIs.

Double stimulus methodologies, while commonly used in video quality assessment and widely adopted in 2D-based quality assessment of point cloud contents, are tricky to adopt in VR technology, due to the difficulties in displaying both contents simultaneously in a perceptually satisfying way, while ensuring a fair comparison between the contents under evaluation. When dealing with interactive methodologies, in particular, synchronous display of any modification in viewpoint is usually enforced, to ensure that the two contents are always visible at the same condition. This is clearly challenging to implement in a 6DoF scenario, in which users are free to change their position in the VR space at any given time. Positioning the two contents side by side in the same virtual space would mean that, at any given time, they are seen from two different angles; the same problem would arise when temporal sequencing is employed. A toggle-based method like the one proposed in [36] is not applicable to moving sequences, as different frames would be seen between stimuli.

In our study, we saw that content preference had an impact on the ratings, as several contents were deemed of lower quality, as the scores given to the HR exemplify. Such bias resulted in a reduced rating range for the contents. Results of the interviews also pointed out that naturalness of gestures were an important criteria in assessing the visual quality. Such components would not be normally evaluated in a double stimulus scenario; however, they are important in understanding how human perception reacts to digital humans.

6 Conclusion

In this paper, we extend our previous work by comparing the performance of the point cloud compression standard V-PCC against an octree-based anchor codec (MPEG anchor). Participants were invited to assess the quality of digital humans represented as dynamic point clouds, in 2DTV screen, as well as VR with 3DoF and 6DoF conditions. Results indicate a small effect of viewing condition on the final scores for one of the two sets of contents under test. Moreover, results show that codec V-PCC has a more favorable performance than the MPEG anchor, especially at low bit-rates. For the highest bit-rate, the two codecs are often statistically equivalent. The contents under test appear to have a significant influence on how the scores are distributed; thus, new data sets are needed in order to comprehensively evaluate compression distortions. Moreover, current encoding solutions, while efficient at low bitrates, are unable to provide visually lossless results, even when large volumes of data are available, revealing significant shortcomings in point cloud compression. We also point out that commonly-used double stimulus methodologies for quality evaluation often reduce the rating task to a difference recognition, while insights on the quality of the original contents are missed. The raw data is made available at the following link: https://github.com/cwi-dis/2DTV_VR_QoE.

References

1. OTT: Beyond Entertainment Consumer Survey Report. https://www.conviva.com/research/ott-beyond-entertainment/
2. Call for proposals for point cloud compression iso/iec jtc1/sc29 wg11 n16732. geneva ch (2017)
3. Alexiou, E., Ebrahimi, T.: On the performance of metrics to predict quality in point cloud representations. In: Applications of Digital Image Processing XL, vol. 10396, p. 103961H. International Society for Optics and Photonics (2017)
4. Alexiou, E., Ebrahimi, T.: Impact of visualisation strategy for subjective quality assessment of point clouds. In: 2018 IEEE International Conference on Multimedia & Expo Workshops (ICMIEW), pp. 1–6. IEEE (2018)
5. Alexiou, E., Tung, K., Ebrahimi, T.: Towards neural network approaches for point cloud compression. In: Applications of Digital Image Processing XLIII, vol. 11510, p. 1151008. International Society for Optics and Photonics (2020)
6. Alexiou, E., Upenik, E., Ebrahimi, T.: Towards subjective quality assessment of point cloud imaging in augmented reality. In: 2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), pp. 1–6. IEEE (2017)
7. Alexiou, E., Viola, I., Borges, T.M., Fonseca, T.A., de Queiroz, R.L., Ebrahimi, T.: A comprehensive study of the rate-distortion performance in mpeg point cloud compression. APSIPA Transactions on Signal and Information Processing 8, 27 (2019). DOI 10.1017/ATSIP.2019.20. URL http://infoscience.epfl.ch/record/272124
8. Alexiou, E., Xia, P., Ebrahimi, T.: Towards modelling of visual saliency in point clouds for immersive applications. In: 26th IEEE International Conference on Image Processing (ICIP) (2019)
9. Alexiou, E., Yang, N., Ebrahimi, T.: Pointx: A toolbox for visualisation and subjective evaluation of point clouds in virtual reality. In: 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6. IEEE (2020)
10. Cao, C., Preda, M., Zakharchenko, V., Jing, E.S., Zaharia, T.: Compression of sparse and dense dynamic point clouds–methods and standards. Proceedings of the IEEE (2021)
11. Cohen, R.A., Tian, D., Vetro, A.: Point cloud attribute compression using 3-d intra prediction and shape-adaptive transforms. In: 2016 Data Compression Conference (DCC), pp. 141–150. IEEE (2016)
12. d’Eon, E., Harrison, B., Myers, T., Chou, P.A.: Si Voxelized Full Bodies - A Voxelized Point Cloud Dataset, ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) input document WG11M40059/WG1M74006, Geneva (2017)
13. Ebner, T., Feldmann, I., Schreer, O., Kauff, P., v. Unger, T.: HHI Point cloud dataset of a boxing trainer, ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) input document MPEG2018/m42921, Ljubljana (2018)
14. Ebrahimi, T., Foesel, S., Pereira, F., Schelkens, P.: Jpeg pleno: Toward an efficient representation of virtual reality. Ieee Multimedia 25(4), 14–20 (2016)
15. Elkin, L.A., Kay, M., Higgins, J.J., Wobbrock, J.O.: An aligned rank transform procedure for multifactor contrast tests. arXiv preprint arXiv:2102.11824 (2021)
16. Ferstl, Y., Thomas, S., Guiard, C., Ennis, C., McDonnell, R.: Human or robot? investigating voice, appearance and gesture motion realism of conversational social agents. In: Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents, pp. 76–83 (2021)
On the impact of VR assessment on the Quality of Experience of Highly Realistic Digital Humans

17. Guarda, A.F., Rodrigues, N.M., Pereira, F.: Deep learning-based point cloud coding: A behavior and performance study. In: 2019 8th European Workshop on Visual Information Processing (EUVIP), pp. 34–39. IEEE (2019)
18. Guarda, A.F., Rodrigues, N.M., Pereira, F.: Point cloud coding: Adopting a deep learning-based approach. In: 2019 Picture Coding Symposium (PCS), pp. 1–5. IEEE (2019)
19. Guarda, A.F., Rodrigues, N.M., Pereira, F.: Deep learning-based point cloud geometry coding: Rd control through implicit and explicit quantization. In: 2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp. 1–6. IEEE (2020)
20. van der Hooft, J., Vega, M.T., Timmerer, C., Begen, A.C., De Turck, F., Scharz, R.: Objective and subjective qoe evaluation for adaptive point cloud streaming. In: 2020 twelfth international conference on quality of multimedia experience (QoMEX), pp. 1–6. IEEE (2020)
21. ISO/IEC DIS 23090-5: Geometry-based point cloud compression. International Organization for Standardization
22. ISO/IEC DIS 23090-5: Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC). International Organization for Standardization (2021)
23. ITU-R BT.500-13: Methodology for the subjective assessment of the quality of television pictures. International Telecommunication Union (2012)
24. ITU-T P.910: Subjective video quality assessment methods for multimedia applications. International Telecommunication Union (2008)
25. ITU-T P.913: Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment. International Telecommunication Union (2016)
26. Jackins, C.L., Tanimoto, S.L.: Oct-trees and their use in representing three-dimensional objects. Computer Graphics and Image Processing 14(3), 249–270 (1980)
27. Jang, E.S., Preda, M., Mamou, K., Tourapis, A.M., Kim, J., Graziosi, D.B., Rhyu, S., Budagavi, M.: Video-based point-cloud-compression standard in mpeg: from evidence collection to committee draft [standards in a nutshell]. IEEE Signal Processing Magazine 36(3), 118–123 (2019)
28. Kay, M., Wobbrock, J.: mjskay/artool: Artool 0.10.6 (2019). DOI 10.5281/zenodo.2556415. URL https://doi.org/10.5281/zenodo.2556415
29. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. The international journal of aviation psychology 3(3), 203–220 (1993)
30. Kortum, P., Sullivan, M.: The effect of content desirability on subjective video quality ratings. Human factors 52(1), 105–118 (2010)
31. Kucherenko, T., Jonell, P., Yoon, Y., Wolfert, P., Henter, G.E.: A large, crowdsourced evaluation of gesture generation systems on common data: The genea challenge 2020. In: 26th International Conference on Intelligent User Interfaces, pp. 11–21 (2021)
32. Mamou, K.: PCC test model category 2 v0. ISO/IEC JTC1/SC29/ WG11 N17248 I (2017)
33. Meagher, D.: Geometric modeling using octree encoding. Computer graphics and image processing 19(2), 129–147 (1982)
34. Mekuria, R., Blom, K., Cesar, P.: Design, implementation, and evaluation of a point cloud codec for tele-immersive video. IEEE Transactions on Circuits and Systems for Video Technology 27(4), 828–842 (2017)
35. Neff, M.: Hand gesture synthesis for conversational characters. Handbook of Human Motion pp. 1–12 (2016)
36. Perrin, A.F., Bist, C., Cozot, R., Ebrahimi, T.: Measuring quality of omnidirectional high dynamic range content. In: Applications of Digital Image Processing XL, vol. 10396, p. 1039613. International Society for Optics and Photonics (2017)
37. Perry, S., Cong, H.P., da Silva Cruz, L.A., Prazeres, J., Pereira, M., Pinheiro, A., Dumic, E., Alexiou, E., Ebrahimi, T.: Quality evaluation of static point clouds encoded using mpeg codecs. In: 2020 IEEE International Conference on Image Processing (ICIP), pp. 3428–3432. IEEE (2020)
38. Quach, M., Valenzise, G., Dufaux, F.: Learning convolutional transforms for lossy point cloud geometry compression. In: 2019 IEEE International Conference on Image Processing (ICIP), pp. 4320–4324. IEEE (2019)
39. Quach, M., Valenzise, G., Dufaux, F.: Improved deep point cloud geometry compression. In: 2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP), pp. 1–6. IEEE (2020)
40. Queiroz, R.D., Chou, P.A.: Compression of 3D Point Clouds Using a Region-Adaptive Hierarchical Transform. IEEE Transactions on Image Processing 25 (2016)
41. Satgunam, P.N., Woods, R.L., Bronstad, P.M., Peli, E.: Factors affecting enhanced video quality preferences. IEEE Transactions on Image Processing 22(12), 5146–5157 (2013)
42. Schubert, T.W.: The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. Zietschrift für Medienpsychologie 15(2), 69–71 (2003)
43. Schwarz, S., Preda, M., Baroncini, V., Budagavi, M., Cesar, P., Chou, P.A., Cohen, R.A., Krivokuca, M., Lasserre, S., Li, Z., et al.: Emerging MPEG standards for point cloud compression. IEEE Journal on Emerging and Selected Topics in Circuits and Systems 9(1), 133–148 (2018)
44. da Silva Cruz, L.A., Dumic, E., Alexiou, E., Prazeres, J., Duarte, R., Pereira, M., Pinheiro, A., Ebrahimi, T.: Point cloud quality evaluation: Towards a definition for test conditions. In: 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6. IEEE (2019)
45. Subramanyam, S., Li, J., Viola, I., Cesar, P.: Comparing the quality of highly realistic digital humans in 3dof and 6dof: A volumetric video case study. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 127–136. IEEE (2020)
46. Subramanyam, S., Viola, I., Hanjalic, A., Cesar, P.: User centered adaptive streaming of dynamic point clouds with low complexity tiling. In: Proceedings of the 28th ACM International Conference on Multimedia, pp. 3669–3677 (2020)
47. Torlig, E.M., Alexiou, E., Fonseca, T.A., de Queiroz, R.L., Ebrahimi, T.: A novel methodology for quality assessment of voxelized point clouds. In: Applications of Digital Image Processing XLII, vol. 10752, p. 107520I. International Society for Optics and Photonics (2017)
48. TT Tran, H., Ngoc, N.P., Pham, C.T., Jung, Y.J., Thang, T.C.: A subjective study on user perception aspects in virtual reality. Applied Sciences 9(16), 3384 (2019)
49. Viola, I., Ebrahimi, T.: A new framework for interactive quality assessment with application to light field coding. In: Applications of Digital Image Processing XL, vol. 10396, p. 103961F. International Society for Optics and Photonics (2017)
50. Wang, J., Zhu, H., Liu, H., Ma, Z.: Lossy point cloud geometry compression via end-to-end learning. IEEE Transactions on Circuits and Systems for Video Technology (2021)
51. Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J.: The Aligned Rank Transform for nonparametric factorial analyses using only ANOVA procedures. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp. 143–146. ACM (2011)
52. Zerman, E., Gao, P., Ozcinar, C., Smolic, A.: Subjective and objective quality assessment for volumetric video compression. In: Fast track article for IST International Symposium on Electronic Imaging 2019: Image Quality and System Performance XVI proceedings (2019)
53. Zerman, E., Kulkarni, R., Smolic, A.: User behaviour analysis of volumetric video in augmented reality. In: 2021 13th International Conference on Quality of Multimedia Experience (QoMEX), pp. 129–132. IEEE (2021)

54. Zerman, E., Ozcinar, C., Gao, P., Smolic, A.: Textured mesh vs coloured point cloud: A subjective study for volumetric video compression. In: 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6. IEEE (2020)

55. Zhang, C., Florencio, D., Loop, C.: Point cloud attribute compression with graph transform. Image Processing (ICIP), 2014 IEEE International Conference on (2014)

56. Zhang, J., Huang, W., Zhu, X., Hwang, J.N.: A subjective quality evaluation for 3D point cloud models. In: 2014 International Conference on Audio, Language and Image Processing, pp. 827–831. IEEE (2014)
Appendix A

Table 5: Contrast test between factors DoF and Content

| contrast                        | estimate  | SE     | df    | t.ratio | p.value |
|---------------------------------|-----------|--------|-------|---------|---------|
| 3DoF, Manfred - 3DoF, Sarge     | -9.0337   | 60.7066| 2294.00| -0.149  | 1.0000  |
| 3DoF, Manfred - 3DoF, Despoina | 178.6466  | 60.7066| 2294.00| 2.943   | 0.1268  |
| 3DoF, Manfred - 3DoF, Queen    | -388.5697 | 60.7066| 2294.00| -6.401  | <0.0001 |
| 3DoF, Manfred - 6DoF, Manfred  | -88.2236  | 108.4749| 126.03| -0.813  | 0.9999  |
| 3DoF, Manfred - 6DoF, Sarge    | -188.9303 | 108.4749| 126.03| -1.742  | 0.8452  |
| 3DoF, Manfred - 6DoF, Queen    | 279.4808  | 108.4749| 126.03| 2.576   | 0.3044  |
| 3DoF, Sarge - 3DoF, Despoina   | 187.6803  | 60.7066| 2294.00| 3.092   | 0.0848  |
| 3DoF, Sarge - 3DoF, Queen      | -379.5361 | 60.7066| 2294.00| -6.252  | <0.0001 |
| 3DoF, Sarge - 6DoF, Manfred    | -84.8082  | 108.4749| 126.03| -0.767  | 0.9998  |
| 3DoF, Sarge - 6DoF, Sarge      | 189.7498  | 108.4749| 126.03| 1.732   | 0.8499  |
| 3DoF, Sarge - 6DoF, Despoina   | -567.2163 | 60.7066| 2294.00| -9.344  | <0.0001 |
| 3DoF, Sarge - 6DoF, Queen      | -300.8752 | 108.4749| 126.03| -2.746  | 0.2166  |
| 3DoF, Sarge - 2DTV, Manfred    | -21.7965  | 109.5543| 126.03| -0.199  | 1.0000  |
| 3DoF, Sarge - 2DTV, Sarge      | 11.1032   | 109.5543| 126.03| 0.101   | 1.0000  |
| 3DoF, Sarge - 2DTV, Despoina   | -262.7268 | 109.5543| 126.03| -2.398  | 0.4151  |
| 3DoF, Sarge - 2DTV, Queen      | -479.5218 | 109.5543| 126.03| -4.377  | 0.0015  |
| 3DoF, Queen - 6DoF, Manfred    | 300.3462  | 108.4749| 126.03| 2.769   | 0.2064  |
| 3DoF, Queen - 6DoF, Sarge      | 316.1755  | 108.4749| 126.03| 2.915   | 0.1485  |
| 3DoF, Queen - 6DoF, Despoina   | 668.0505  | 108.4749| 126.03| 6.159   | <0.0001 |
| 3DoF, Queen - 6DoF, Queen      | 199.6394  | 108.4749| 126.03| 1.840   | 0.7924  |
| 3DoF, Queen - 2DTV, Manfred    | 357.7395  | 109.5543| 126.03| 3.265   | 0.0601  |
| 3DoF, Queen - 2DTV, Sarge      | 578.3195  | 109.5543| 126.03| 5.279   | <0.0001 |
| 3DoF, Queen - 2DTV, Despoina   | 304.4895  | 109.5543| 126.03| 2.779   | 0.2018  |
| 3DoF, Queen - 2DTV, Queen      | 87.6945   | 109.5543| 126.03| 0.800   | 0.9997  |
| 6DoF, Manfred - 6DoF, Sarge    | 15.8293   | 60.7066| 2294.00| 0.261   | 1.0000  |
| 6DoF, Manfred - 6DoF, Despoina | 367.7043  | 60.7066| 2294.00| 6.057   | <0.0001 |
| 6DoF, Manfred - 6DoF, Queen    | -100.7067 | 60.7066| 2294.00| -1.659  | 0.8865  |
| 6DoF, Manfred - 2DTV, Manfred  | 57.3934   | 109.5543| 126.03| 0.524   | 1.0000  |
| 6DoF, Manfred - 2DTV, Sarge    | 277.9734  | 109.5543| 126.03| 2.537   | 0.3272  |
| 6DoF, Manfred - 2DTV, Despoina | 4.1434    | 109.5543| 126.03| 0.038   | 1.0000  |
| 6DoF, Manfred - 2DTV, Queen    | -212.6516 | 109.5543| 126.03| -1.941  | 0.7312  |
| 6DoF, Sarge - 6DoF, Despoina   | 351.8750  | 60.7066| 2294.00| 5.796   | <0.0001 |
| 6DoF, Sarge - 6DoF, Queen      | -116.5361 | 60.7066| 2294.00| -1.920  | 0.7970  |
| 6DoF, Sarge - 2DTV, Manfred    | 41.5640   | 109.5543| 126.03| 0.379   | 1.0000  |
### Table 5 – Continued from previous page

| contrast                          | estimate   | SE       | df   | t.ratio | p.value |
|----------------------------------|------------|----------|------|---------|---------|
| 6DoF, Sarge - 2DTV, Sarge        | 262.1440   | 109.5543 | 126.03 | 2.393   | 0.4187  |
| 6DoF, Sarge - 2DTV, Despoina    | -11.6860   | 109.5543 | 126.03 | -0.107  | 1.0000  |
| 6DoF, Sarge - 2DTV, Queen       | -228.4810  | 109.5543 | 126.03 | -2.086  | 0.6338  |
| 6DoF, Despoina - 2DTV, Sarge    | -468.4111  | 60.7066  | 2294.00 | -7.716  | <.0001  |
| 6DoF, Despoina - 2DTV, Manfred  | -310.3110  | 109.5543 | 126.03 | -2.832  | 0.1794  |
| 6DoF, Despoina - 2DTV, Queen    | -363.5610  | 109.5543 | 126.03 | -3.319  | 0.0517  |
| 6DoF, Queen - 2DTV, Manfred     | 158.1001   | 109.5543 | 126.03 | 1.443   | 0.9524  |
| 6DoF, Queen - 2DTV, Sarge       | 378.6801   | 109.5543 | 126.03 | 3.457   | 0.0345  |
| 6DoF, Queen - 2DTV, Despoina    | 104.8501   | 109.5543 | 126.03 | 0.957   | 0.9983  |
| 6DoF, Queen - 2DTV, Queen       | -111.9449  | 109.5543 | 126.03 | -1.022  | 0.9969  |
| 2DTV, Manfred - 2DTV, Sarge     | 220.5800   | 61.9088  | 2294.00 | 3.563   | 0.0193  |
| 2DTV, Manfred - 2DTV, Despoina  | -53.2500   | 61.9088  | 2294.00 | -0.860  | 0.9994  |
| 2DTV, Manfred - 2DTV, Queen     | -273.8300  | 61.9088  | 2294.00 | -4.362  | 0.0008  |
| 2DTV, Sarge - 2DTV, Despoina    | -490.6250  | 61.9088  | 2294.00 | -7.925  | <.0001  |
| 2DTV, Sarge - 2DTV, Queen       | -216.7950  | 61.9088  | 2294.00 | -3.502  | 0.0238  |

Results are averaged over the levels of: Codecs, Bitrates

Degrees-of-freedom method: kenward-roger

P value adjustment: tukey method for comparing a family of 12 estimates

### Table 6: Contrast test between factors DoF and Content

| contrast                  | estimate   | SE       | df   | t.ratio | p.value |
|---------------------------|------------|----------|------|---------|---------|
| Manfred, C1 - Manfred, C2 | 259.8001   | 49.6388  | 2294 | 5.234   | <.0001  |
| Manfred, C1 - Sarge, C1   | 37.6988    | 49.6388  | 2294 | 0.759   | 0.9950  |
| Manfred, C1 - Sarge, C2   | 358.0433   | 49.6388  | 2294 | 7.213   | <.0001  |
| Manfred, C1 - Despoina, C1| 181.0136   | 49.6388  | 2294 | 3.647   | 0.0066  |
| Manfred, C1 - Despoina, C2| 435.7672   | 49.6388  | 2294 | 8.779   | <.0001  |
| Manfred, C1 - Queen, C1   | -103.4447  | 49.6388  | 2294 | -2.084  | 0.4256  |
| Manfred, C1 - Queen, C2   | -141.3435  | 49.6388  | 2294 | -2.847  | 0.0842  |
| Manfred, C2 - Sarge, C1   | -222.1013  | 49.6388  | 2294 | -4.474  | 0.0002  |
| Manfred, C2 - Sarge, C2   | 98.2432    | 49.6388  | 2294 | 1.979   | 0.4963  |
| Manfred, C2 - Despoina, C1| -78.7865   | 49.6388  | 2294 | -1.587  | 0.7582  |
| Manfred, C2 - Despoina, C2| 175.9671   | 49.6388  | 2294 | 3.545   | 0.0095  |
| Manfred, C2 - Queen, C1   | -363.2449  | 49.6388  | 2294 | -7.318  | <.0001  |
| Manfred, C2 - Queen, C2   | -401.1436  | 49.6388  | 2294 | -8.081  | <.0001  |
| Sarge, C1 - Sarge, C2     | 320.3445   | 49.6388  | 2294 | 6.454   | <.0001  |
| Sarge, C1 - Despoina, C1  | 143.3147   | 49.6388  | 2294 | 2.887   | 0.0756  |
| Sarge, C1 - Despoina, C2  | 398.0683   | 49.6388  | 2294 | 8.019   | <.0001  |
| Sarge, C1 - Queen, C1     | -141.1436  | 49.6388  | 2294 | -2.843  | 0.0851  |
| Sarge, C1 - Queen, C2     | -179.0423  | 49.6388  | 2294 | -3.607  | 0.0076  |
| Sarge, C2 - Despoina, C1  | -177.0297  | 49.6388  | 2294 | -3.566  | 0.0088  |
| Sarge, C2 - Despoina, C2  | 77.7238    | 49.6388  | 2294 | 1.566   | 0.7708  |
| Sarge, C2 - Queen, C1     | -461.4881  | 49.6388  | 2294 | -9.297  | <.0001  |
| Sarge, C2 - Queen, C2     | -499.3868  | 49.6388  | 2294 | -10.060 | <.0001  |
| Despoina, C1 - Despoina, C2| 254.7536   | 49.6388  | 2294 | 5.132   | <.0001  |
| Despoina, C1 - Queen, C1  | -284.4583  | 49.6388  | 2294 | -5.731  | <.0001  |
| Despoina, C1 - Queen, C2  | -322.3571  | 49.6388  | 2294 | -6.494  | <.0001  |
| Despoina, C2 - Queen, C1  | -539.2119  | 49.6388  | 2294 | -10.863 | <.0001  |
Table 6 – Continued from previous page

| contrast                      | estimate  | SE     | df    | t.ratio  | p.value |
|-------------------------------|-----------|--------|-------|----------|---------|
| Despoina, C2 - Queen, C2     | -577.1106 | 49.6388| 2294  | -11.626  | <.0001  |
| Queen, C1 - Queen, C2        | -37.8987  | 49.6388| 2294  | -0.763   | 0.9949  |

Results are averaged over the levels of: DoF, Bitrates

Degrees-of-freedom method: kenward-roger

P value adjustment: tukey method for comparing a family of 8 estimates

Table 7: Contrast test between factors Content and Bitrate

| contrast                      | estimate  | SE     | df    | t.ratio  | p.value |
|-------------------------------|-----------|--------|-------|----------|---------|
| Manfred, R1 - Manfred, R2    | -261.0203 | 56.8360| 2294  | -4.593   | 0.0005  |
| Manfred, R1 - Manfred, R3    | -813.6400 | 56.8360| 2294  | -14.316  | <.0001  |
| Manfred, R1 - Manfred, R4    | -1080.8721| 56.8360| 2294  | -19.017  | <.0001  |
| Manfred, R1 - Sarge, R1      | 90.9803   | 56.8360| 2294  | 1.601    | 0.9684  |
| Manfred, R1 - Sarge, R2      | -221.5797 | 56.8360| 2294  | -3.899   | 0.0097  |
| Manfred, R1 - Sarge, R3      | -794.1628 | 56.8360| 2294  | -13.973  | <.0001  |
| Manfred, R1 - Sarge, R4      | -996.3062 | 56.8360| 2294  | -17.529  | <.0001  |
| Manfred, R1 - Queen, R1      | -111.2328 | 56.8360| 2294  | -1.957   | 0.8485  |
| Manfred, R1 - Queen, R2      | -697.9362 | 56.8360| 2294  | -12.280  | <.0001  |
| Manfred, R1 - Queen, R3      | -1025.7954| 56.8360| 2294  | -18.048  | <.0001  |
| Manfred, R1 - Queen, R4      | -1068.4754| 56.8360| 2294  | -18.799  | <.0001  |
| Manfred, R2 - Manfred, R3    | 352.0005  | 56.8360| 2294  | 6.193    | <.0001  |
| Manfred, R2 - Sarge, R1      | 904.6203  | 56.8360| 2294  | 15.916   | <.0001  |
| Manfred, R2 - Sarge, R2      | 592.2641  | 56.8360| 2294  | 10.417   | <.0001  |
| Manfred, R2 - Sarge, R3      | 19.4772   | 56.8360| 2294  | 0.343    | 1.0000  |
| Manfred, R2 - Sarge, R4      | -182.6662 | 56.8360| 2294  | -3.214   | 0.0958  |
| Manfred, R2 - Despoina, R1   | 156.6036  | 56.8360| 2294  | 2.755    | 0.2961  |
| Manfred, R2 - Despoina, R2   | -85.3759  | 56.8360| 2294  | -1.502   | 0.9824  |
| Manfred, R2 - Despoina, R3   | -649.9931 | 56.8360| 2294  | -11.436  | <.0001  |
| Manfred, R2 - Despoina, R4   | -1005.8467| 56.8360| 2294  | -17.697  | <.0001  |

On the impact of VR assessment on the Quality of Experience of Highly Realistic Digital Humans
| contrast          | estimate | SE     | df  | t.ratio | p.value   |
|------------------|----------|--------|-----|---------|-----------|
| Manfred, R4 - Sarge, R1 | 1171.8523 | 56.8360 | 2294 | 20.618  | <.0001    |
| Manfred, R4 - Sarge, R2 | 859.2923  | 56.8360 | 2294 | 15.119  | <.0001    |
| Manfred, R4 - Sarge, R3 | 286.7092  | 56.8360 | 2294 | 5.044   | 0.0001    |
| Manfred, R4 - Sarge, R4 | 84.5659   | 56.8360 | 2294 | 1.488   | 0.9839    |
| Manfred, R4 - Sarge, R1 | 969.6392  | 56.8360 | 2294 | 17.060  | <.0001    |
| Manfred, R4 - Sarge, R2 | 382.9359  | 56.8360 | 2294 | 6.738   | <.0001    |
| Manfred, R4 - Sarge, R3 | 1237.4756 | 56.8360 | 2294 | 21.773  | <.0001    |
| Manfred, R4 - Sarge, R4 | 995.4962  | 56.8360 | 2294 | 17.515  | <.0001    |
| Manfred, R4 - Despoina, R1 | 346.3131  | 56.8360 | 2294 | 6.093   | <.0001    |
| Manfred, R4 - Despoina, R2 | 72.1692   | 56.8360 | 2294 | 1.270   | 0.9968    |
| Manfred, R4 - Despoina, R3 | 910.9303  | 56.8360 | 2294 | 16.027  | <.0001    |
| Sarge, R1 - Sarge, R2 | -312.5600 | 56.8360 | 2294 | -5.499  | <.0001    |
| Sarge, R1 - Sarge, R3 | -885.1431 | 56.8360 | 2294 | -15.574 | <.0001    |
| Sarge, R1 - Sarge, R4 | -1087.2864| 56.8360 | 2294 | -19.130 | <.0001    |
| Sarge, R1 - Despoina, R1 | 65.6233   | 56.8360 | 2294 | 1.155   | 0.9989    |
| Sarge, R1 - Despoina, R2 | -176.3562 | 56.8360 | 2294 | -3.103  | 0.1300    |
| Sarge, R1 - Despoina, R3 | -740.9733 | 56.8360 | 2294 | -13.037 | <.0001    |
| Sarge, R1 - Despoina, R4 | -1096.8269| 56.8360 | 2294 | -19.298 | <.0001    |
| Sarge, R2 - Sarge, R1 | 110.3469  | 56.8360 | 2294 | 1.941   | 0.8564    |
| Sarge, R2 - Sarge, R2 | -476.3564 | 56.8360 | 2294 | -8.381  | <.0001    |
| Sarge, R2 - Sarge, R3 | -804.2156 | 56.8360 | 2294 | -14.150 | <.0001    |
| Sarge, R2 - Sarge, R4 | -846.8956 | 56.8360 | 2294 | -14.901 | <.0001    |
| Sarge, R2 - Despoina, R1 | 378.1833  | 56.8360 | 2294 | 6.654   | <.0001    |
| Sarge, R2 - Despoina, R2 | 136.2038  | 56.8360 | 2294 | 2.396   | 0.5505    |
| Sarge, R2 - Despoina, R3 | -428.4133 | 56.8360 | 2294 | -7.538  | <.0001    |
| Sarge, R2 - Despoina, R4 | -784.2669 | 56.8360 | 2294 | -13.799 | <.0001    |
| Sarge, R3 - Sarge, R1 | 114.1697  | 56.8360 | 2294 | 2.537   | 0.4447    |
| Sarge, R3 - Despoina, R1 | 950.7664  | 56.8360 | 2294 | 16.728  | <.0001    |
| Sarge, R3 - Despoina, R2 | 708.7869  | 56.8360 | 2294 | 12.471  | <.0001    |
| Sarge, R3 - Despoina, R3 | 144.1697  | 56.8360 | 2294 | -3.724  | 0.0185    |
| Sarge, R3 - Queen, R1 | 682.9300  | 56.8360 | 2294 | 12.016  | <.0001    |
| Sarge, R3 - Queen, R2 | 96.2267   | 56.8360 | 2294 | 1.693   | 0.9489    |
| Sarge, R3 - Queen, R3 | -231.6326 | 56.8360 | 2294 | -4.075  | 0.0048    |
| Sarge, R3 - Queen, R4 | -274.3126 | 56.8360 | 2294 | -4.826  | 0.0002    |
| Sarge, R4 - Despoina, R1 | 1152.9097 | 56.8360 | 2294 | 20.285  | <.0001    |
| Sarge, R4 - Despoina, R2 | 910.9303  | 56.8360 | 2294 | 16.027  | <.0001    |
| Sarge, R4 - Despoina, R3 | 346.3131  | 56.8360 | 2294 | 6.093   | <.0001    |
| Sarge, R4 - Despoina, R4 | -9.5405   | 56.8360 | 2294 | -0.168  | 1.0000    |
| Sarge, R4 - Queen, R1 | 885.0733  | 56.8360 | 2294 | 15.572  | <.0001    |
| Sarge, R4 - Queen, R2 | 298.3700  | 56.8360 | 2294 | 5.250   | <.0001    |
| Sarge, R4 - Queen, R3 | -29.4892  | 56.8360 | 2294 | -0.519  | 1.0000    |
| Sarge, R4 - Queen, R4 | -72.1692  | 56.8360 | 2294 | -1.270  | 0.9968    |
| Despoina, R1 - Despoina, R2 | -241.9795 | 56.8360 | 2294 | -4.258  | 0.0023    |
Table 7 – Continued from previous page

| contrast                          | estimate | SE   | df   | t.ratio | p.value  |
|----------------------------------|----------|------|------|---------|----------|
| Despoina, R1 - Despoina, R3     | -806.5967| 56.8360 | 2294 | -14.192 | <.0001   |
| Despoina, R1 - Despoina, R4     | -1162.4503| 56.8360 | 2294 | -20.453 | <.0001   |
| Despoina, R1 - Queen, R1        | -267.8364 | 56.8360 | 2294 | -4.712  | 0.0003   |
| Despoina, R1 - Queen, R2        | -854.5397 | 56.8360 | 2294 | -15.035 | <.0001   |
| Despoina, R1 - Queen, R3        | -1182.3990| 56.8360 | 2294 | -20.804 | <.0001   |
| Despoina, R2 - Despoina, R3     | -920.4708 | 56.8360 | 2294 | -16.195 | <.0001   |
| Despoina, R2 - Queen, R1        | -25.8569  | 56.8360 | 2294 | -0.455  | 1.0000   |
| Despoina, R2 - Queen, R2        | -612.5603 | 56.8360 | 2294 | -10.778 | <.0001   |
| Despoina, R2 - Queen, R3        | -940.4195 | 56.8360 | 2294 | -16.546 | <.0001   |
| Despoina, R2 - Queen, R4        | -983.0995 | 56.8360 | 2294 | -17.297 | <.0001   |
| Despoina, R3 - Despoina, R4     | 538.7603  | 56.8360 | 2294 | 9.479   | <.0001   |
| Despoina, R3 - Queen, R1        | -47.9431  | 56.8360 | 2294 | -0.844  | 1.0000   |
| Despoina, R3 - Queen, R2        | -375.8023 | 56.8360 | 2294 | -6.612  | <.0001   |
| Despoina, R3 - Queen, R4        | -418.4823 | 56.8360 | 2294 | -7.363  | <.0001   |
| Despoina, R4 - Queen, R1        | 894.6138  | 56.8360 | 2294 | 15.740  | <.0001   |
| Despoina, R4 - Queen, R2        | 307.9105  | 56.8360 | 2294 | 5.418   | <.0001   |
| Despoina, R4 - Queen, R3        | -19.9487  | 56.8360 | 2294 | -0.351  | 1.0000   |
| Despoina, R4 - Queen, R4        | -62.6287  | 56.8360 | 2294 | -1.102  | 0.9994   |
| Queen, R1 - Queen, R2           | -586.7033 | 56.8360 | 2294 | -10.323 | <.0001   |
| Queen, R1 - Queen, R3           | -914.5626 | 56.8360 | 2294 | -16.091 | <.0001   |
| Queen, R1 - Queen, R4           | -957.2426 | 56.8360 | 2294 | -16.842 | <.0001   |
| Queen, R2 - Queen, R3           | -327.8592 | 56.8360 | 2294 | -5.769  | <.0001   |
| Queen, R2 - Queen, R4           | -370.5392 | 56.8360 | 2294 | -6.519  | <.0001   |
| Queen, R3 - Queen, R4           | -42.6800  | 56.8360 | 2294 | -0.751  | 1.0000   |

Results are averaged over the levels of: DoF, Codecs

Degrees-of-freedom method: kenward-roger

P value adjustment: tukey method for comparing a family of 16 estimates

Table 8: Contrast test between factors Content, Codec and Bitrate

| contrast                          | estimate | SE   | df   | t.ratio | p.value  |
|----------------------------------|----------|------|------|---------|----------|
| Longdress, C1, R1 - Longdress, C1, R2 | -535.4682 | 70.4123 | 2263 | -7.605  | <.0001   |
| Longdress, C1, R1 - Longdress, C1, R3 | -619.6446 | 70.4123 | 2263 | -8.800  | <.0001   |
| Longdress, C1, R1 - Longdress, C1, R4 | -657.2574 | 70.4123 | 2263 | -9.335  | <.0001   |
| Longdress, C1, R1 - Longdress, C2, R1 | 636.6415  | 70.4123 | 2263 | 9.042   | <.0001   |
| Longdress, C1, R1 - Longdress, C2, R2 | 328.2036  | 70.4123 | 2263 | 4.661   | 0.0015   |
| Longdress, C1, R1 - Longdress, C2, R3 | -242.6631 | 70.4123 | 2263 | -3.446  | 0.1438   |
| Longdress, C1, R1 - Longdress, C2, R4 | -507.2836 | 70.4123 | 2263 | -7.204  | <.0001   |
| Longdress, C1, R1 - Loot, C1, R1  | 446.7518  | 70.4123 | 2263 | 6.345   | <.0001   |
| Longdress, C1, R1 - Loot, C1, R2  | -224.4031 | 70.4123 | 2263 | -3.187  | 0.2794   |
| Longdress, C1, R1 - Loot, C1, R3  | -467.3159 | 70.4123 | 2263 | -6.637  | <.0001   |
| Longdress, C1, R1 - Loot, C1, R4  | -535.5154 | 70.4123 | 2263 | -7.605  | <.0001   |
| Longdress, C1, R1 - Loot, C2, R1  | 1014.3595 | 70.4123 | 2263 | 14.406  | <.0001   |
| Longdress, C1, R1 - Loot, C2, R2  | 799.7723  | 70.4123 | 2263 | 11.358  | <.0001   |
| Longdress, C1, R1 - Red and black, C1, R1 | 343.0297 | 70.4123 | 2263 | 4.872   | 0.0005   |
| Longdress, C1, R1 - Red and black, C2, R2 | -247.0190 | 70.4123 | 2263 | -3.508  | 0.1204   |
| Longdress, C1, R1 - Red and black, C1, R2 | -301.1995 | 70.4123 | 2263 | -4.278  | 0.0078   |
| contrast                                      | estimate | SE  | df  | t.ratio | p.value |
|----------------------------------------------|----------|-----|-----|---------|---------|
| Longdress, C1, R1 - Red and black, C1, R3   | -537.738 | 70.412 | 2263 | -7.637  | <.0001  |
| Longdress, C1, R1 - Red and black, C1, R4   | -565.375 | 70.412 | 2263 | -8.030  | <.0001  |
| Longdress, C1, R1 - Red and black, C2, R1   | 746.592  | 70.412 | 2263 | 10.603  | <.0001  |
| Longdress, C1, R1 - Red and black, C2, R2   | 459.903  | 70.412 | 2263 | 6.532   | <.0001  |
| Longdress, C1, R1 - Red and black, C2, R3   | 43.685   | 70.412 | 2263 | 0.620   | 1.0000  |
| Longdress, C1, R1 - Soldier, C1, R1         | 760.319  | 70.412 | 2263 | 10.798  | <.0001  |
| Longdress, C1, R1 - Soldier, C1, R2         | 150.912  | 70.412 | 2263 | 2.143   | 0.9611  |
| Longdress, C1, R1 - Soldier, C1, R3         | -281.403 | 70.412 | 2263 | -3.996  | <.0001  |
| Longdress, C1, R1 - Soldier, C1, R4         | 472.649  | 70.412 | 2263 | 6.713   | <.0001  |
| Longdress, C1, R2 - Longdress, C1, R3       | -84.176  | 70.412 | 2263 | -1.195  | 1.0000  |
| Longdress, C1, R2 - Longdress, C1, R4       | -121.807 | 70.412 | 2263 | -1.730  | <.0001  |
| Longdress, C1, R2 - Longdress, C2, R1       | 1172.109 | 70.412 | 2263 | 16.646  | <.0001  |
| Longdress, C1, R2 - Longdress, C2, R2       | 863.671  | 70.412 | 2263 | 12.266  | <.0001  |
| Longdress, C1, R2 - Longdress, C2, R3       | 292.805  | 70.412 | 2263 | 4.158   | 0.0126  |
| Longdress, C1, R2 - Loot, C1, R1            | 982.220  | 70.412 | 2263 | 13.950  | <.0001  |
| Longdress, C1, R2 - Loot, C1, R2            | 311.065  | 70.412 | 2263 | 4.418   | 0.0043  |
| Longdress, C1, R2 - Loot, C1, R3            | 68.152   | 70.412 | 2263 | 0.968   | 1.0000  |
| Longdress, C1, R2 - Loot, C1, R4            | -0.047   | 70.412 | 2263 | -0.001  | 1.0000  |
| Longdress, C1, R2 - Loot, C2, R1            | 1549.828 | 70.412 | 2263 | 22.011  | <.0001  |
| Longdress, C1, R2 - Loot, C2, R2            | 1335.240 | 70.412 | 2263 | 18.963  | <.0001  |
| Longdress, C1, R2 - Loot, C2, R3            | 878.497  | 70.412 | 2263 | 4.418   | 0.0043  |
| Longdress, C1, R2 - Loot, C2, R4            | 288.449  | 70.412 | 2263 | 4.097   | 0.0161  |
| Longdress, C1, R2 - Red and black, C1, R1   | 728.598  | 70.412 | 2263 | 10.348  | <.0001  |
| Longdress, C1, R2 - Red and black, C1, R2   | 1172.109 | 70.412 | 2263 | 16.646  | <.0001  |
| Longdress, C1, R2 - Red and black, C1, R3   | -2.270   | 70.412 | 2263 | -0.032  | 1.0000  |
| Longdress, C1, R2 - Red and black, C1, R4   | 234.268  | 70.412 | 2263 | 3.327   | 0.1984  |
| Longdress, C1, R2 - Red and black, C2, R1   | 995.372  | 70.412 | 2263 | 14.136  | <.0001  |
| Longdress, C1, R2 - Red and black, C2, R2   | 579.153  | 70.412 | 2263 | 8.225   | <.0001  |
| Longdress, C1, R2 - Red and black, C2, R3   | 237.975  | 70.412 | 2263 | 3.380   | 0.1727  |
| Longdress, C1, R2 - Red and black, C2, R4   | 1295.786 | 70.412 | 2263 | 18.403  | <.0001  |
| Longdress, C1, R2 - Soldier, C1, R1         | 686.380  | 70.412 | 2263 | 9.748   | <.0001  |
| Longdress, C1, R2 - Soldier, C1, R2         | 254.066  | 70.412 | 2263 | 3.608   | 0.0890  |
| Longdress, C1, R2 - Soldier, C1, R3         | 62.819   | 70.412 | 2263 | 0.892   | 1.0000  |
| Longdress, C1, R2 - Soldier, C2, R1         | 1592.915 | 70.412 | 2263 | 22.623  | <.0001  |
| Longdress, C1, R2 - Soldier, C2, R2         | 1377.954 | 70.412 | 2263 | 19.570  | <.0001  |
| Longdress, C1, R2 - Soldier, C2, R3         | 939.434  | 70.412 | 2263 | 13.342  | <.0001  |
| Longdress, C1, R2 - Soldier, C2, R4         | 318.059  | 70.412 | 2263 | 4.517   | 0.0028  |
| Longdress, C1, R3 - Longdress, C1, R4       | -37.630  | 70.412 | 2263 | -0.534  | 1.0000  |
| Longdress, C1, R3 - Longdress, C2, R1       | 1256.286 | 70.412 | 2263 | 17.842  | <.0001  |
| Longdress, C1, R3 - Longdress, C2, R2       | 947.848  | 70.412 | 2263 | 13.461  | <.0001  |
| Longdress, C1, R3 - Longdress, C2, R3       | 376.981  | 70.412 | 2263 | 5.354   | <.0001  |
| Longdress, C1, R3 - Longdress, C2, R4       | 112.361  | 70.412 | 2263 | 1.596   | 0.9996  |
| Longdress, C1, R3 - Loot, C1, R1            | 1066.396 | 70.412 | 2263 | 15.145  | <.0001  |
| Longdress, C1, R3 - Loot, C1, R2            | 395.241  | 70.412 | 2263 | 5.613   | <.0001  |
Table 8 – Continued from previous page

| contrast                        | estimate  | SE      | df  | t.ratio | p.value |
|---------------------------------|-----------|---------|-----|---------|---------|
| Longdress, C1, R3 - Loot, C1, R3| 152.3287  | 70.4123 | 2263| 2.163   | 0.9564  |
| Longdress, C1, R3 - Loot, C1, R4| 84.1292   | 70.4123 | 2263| 1.195   | 1.0000  |
| Longdress, C1, R3 - Loot, C2, R1| 1634.0041 | 70.4123 | 2263| 23.206  | <.0001  |
| Longdress, C1, R3 - Loot, C2, R2| 1419.4169 | 70.4123 | 2263| 20.159  | <.0001  |
| Longdress, C1, R3 - Red and black, C1, R3| 812.7749 | 70.4123 | 2263| 11.543  | <.0001  |
| Longdress, C1, R3 - Red and black, C2, R1| 1366.2369 | 70.4123 | 2263| 19.403  | <.0001  |
| Longdress, C1, R3 - Red and black, C2, R2| 1079.5482 | 70.4123 | 2263| 15.332  | <.0001  |
| Longdress, C1, R3 - Red and black, C2, R3| 663.3292 | 70.4123 | 2263| 9.421   | <.0001  |
| Longdress, C1, R3 - Red and black, C2, R4| 322.1518 | 70.4123 | 2263| 4.575   | 0.0021  |
| Longdress, C1, R4 - Longdress, C2, R1| 1293.9169 | 70.4123 | 2263| 18.376  | <.0001  |
| Longdress, C1, R4 - Longdress, C2, R2| 985.4790  | 70.4123 | 2263| 13.996  | <.0001  |
| Longdress, C1, R4 - Longdress, C2, R3| 414.6123  | 70.4123 | 2263| 5.888   | <.0001  |
| Longdress, C1, R4 - Longdress, C2, R4| 149.9918  | 70.4123 | 2263| 2.130   | 0.9639  |
| Longdress, C1, R4 - Loot, C1, R1| 1104.0272 | 70.4123 | 2263| 15.679  | <.0001  |
| Longdress, C1, R4 - Loot, C1, R2| 432.8723  | 70.4123 | 2263| 6.148   | <.0001  |
| Longdress, C1, R4 - Loot, C1, R3| 189.9595  | 70.4123 | 2263| 2.698   | 0.6625  |
| Longdress, C1, R4 - Loot, C1, R4| 121.7600  | 70.4123 | 2263| 1.729   | 0.9984  |
| Longdress, C1, R4 - Loot, C2, R1| 1671.6349 | 70.4123 | 2263| 23.741  | <.0001  |
| Longdress, C1, R4 - Loot, C2, R2| 1457.0477 | 70.4123 | 2263| 20.693  | <.0001  |
| Longdress, C1, R4 - Loot, C2, R3| 1000.3051 | 70.4123 | 2263| 14.206  | <.0001  |
| Longdress, C1, R4 - Loot, C2, R4| 410.2564  | 70.4123 | 2263| 5.826   | <.0001  |
| Longdress, C1, R4 - Red and black, C1, R1| 850.4056 | 70.4123 | 2263| 12.078  | <.0001  |
| Longdress, C1, R4 - Red and black, C1, R2| 356.0759 | 70.4123 | 2263| 5.057   | 0.0002  |
| Longdress, C1, R4 - Red and black, C1, R3| 119.5369 | 70.4123 | 2263| 1.698   | 0.9989  |
| Longdress, C1, R4 - Red and black, C1, R4| 91.8969  | 70.4123 | 2263| 1.305   | 1.0000  |
| Longdress, C1, R4 - Red and black, C2, R1| 1403.8677 | 70.4123 | 2263| 19.938  | <.0001  |
| Longdress, C1, R4 - Red and black, C2, R2| 1117.1790 | 70.4123 | 2263| 15.866  | <.0001  |
| Longdress, C1, R4 - Red and black, C2, R3| 700.9600  | 70.4123 | 2263| 9.955   | <.0001  |
| Longdress, C1, R4 - Red and black, C2, R4| 359.7826  | 70.4123 | 2263| 5.110   | 0.0002  |
| Longdress, C1, R4 - Soldier, C1, R1| 1417.5938 | 70.4123 | 2263| 20.133  | <.0001  |
| Longdress, C1, R4 - Soldier, C1, R2| 808.1877  | 70.4123 | 2263| 11.478  | <.0001  |
| Longdress, C1, R4 - Soldier, C1, R3| 375.8728  | 70.4123 | 2263| 5.338   | <.0001  |
| Longdress, C1, R4 - Soldier, C1, R4| 184.6267  | 70.4123 | 2263| 2.622   | 0.7219  |
| Longdress, C1, R4 - Soldier, C2, R1| 1714.7226 | 70.4123 | 2263| 24.353  | <.0001  |
| Longdress, C1, R4 - Soldier, C2, R2| 1499.7615 | 70.4123 | 2263| 21.300  | <.0001  |
| Longdress, C1, R4 - Soldier, C2, R3| 1061.2421 | 70.4123 | 2263| 15.072  | <.0001  |
| Longdress, C1, R4 - Soldier, C2, R4| 439.8667  | 70.4123 | 2263| 6.247   | <.0001  |
| Longdress, C2, R1 - Longdress, C2, R2| -308.4379 | 70.4123 | 2263| -4.380  | 0.0051  |
Table 8 – Continued from previous page

| contrast                          | estimate | SE     | df    | t.ratio | p.value |
|----------------------------------|----------|--------|-------|---------|---------|
| Longdress, C2, R1 - Longdress, C2, R3 | -879.3046 | 70.4123 | 2263  | -12.488 | <.0001  |
| Longdress, C2, R1 - Longdress, C2, R4 | -1143.9251 | 70.4123 | 2263  | -16.246 | <.0001  |
| Longdress, C2, R1 - Loot, C1, R1  | -189.8897 | 70.4123 | 2263  | -2.697  | 0.6633  |
| Longdress, C2, R1 - Loot, C1, R2  | -861.0446 | 70.4123 | 2263  | -12.229 | <.0001  |
| Longdress, C2, R1 - Loot, C1, R3  | -1103.9574 | 70.4123 | 2263  | -15.678 | <.0001  |
| Longdress, C2, R1 - Loot, C2, R1  | 377.7179 | 70.4123 | 2263  | 5.364   | <.0001  |
| Longdress, C2, R1 - Loot, C2, R2  | 163.1308 | 70.4123 | 2263  | 2.317   | 0.9058  |
| Longdress, C2, R1 - Loot, C2, R3  | -293.6118 | 70.4123 | 2263  | -4.170  | 0.0120  |
| Longdress, C2, R1 - Loot, C2, R4  | -1172.1569 | 70.4123 | 2263  | -16.647 | <.0001  |
| Longdress, C2, R1 - Red and black, C1, R1 | -443.5113 | 70.4123 | 2263  | -6.299  | <.0001  |
| Longdress, C2, R1 - Red and black, C1, R2 | -937.8410 | 70.4123 | 2263  | -13.319 | <.0001  |
| Longdress, C2, R1 - Red and black, C1, R3 | -1174.3800 | 70.4123 | 2263  | -16.679 | <.0001  |
| Longdress, C2, R1 - Red and black, C1, R4 | -1202.0200 | 70.4123 | 2263  | -17.071 | <.0001  |
| Longdress, C2, R1 - Red and black, C2, R1 | 109.9508 | 70.4123 | 2263  | 1.562   | 0.9998  |
| Longdress, C2, R1 - Red and black, C2, R2 | -176.7379 | 70.4123 | 2263  | -2.510  | 0.8014  |
| Longdress, C2, R1 - Red and black, C2, R3 | -94.3444 | 70.4123 | 2263  | -1.3267 | <.0001  |
| Longdress, C2, R1 - Red and black, C2, R4 | 123.6769 | 70.4123 | 2263  | 1.756   | 0.9979  |
| Longdress, C2, R2 - Longdress, C2, R3 | -570.8667 | 70.4123 | 2263  | -8.107  | <.0001  |
| Longdress, C2, R2 - Longdress, C2, R4 | -835.4872 | 70.4123 | 2263  | -11.866 | <.0001  |
| Longdress, C2, R2 - Loot, C1, R1  | 118.5482 | 70.4123 | 2263  | 1.684   | 0.9990  |
| Longdress, C2, R2 - Loot, C1, R2  | -552.6067 | 70.4123 | 2263  | -7.848  | <.0001  |
| Longdress, C2, R2 - Loot, C1, R3  | -795.5195 | 70.4123 | 2263  | -11.298 | <.0001  |
| Longdress, C2, R2 - Loot, C1, R4  | -863.7190 | 70.4123 | 2263  | -12.267 | <.0001  |
| Longdress, C2, R2 - Loot, C2, R1  | 686.1559 | 70.4123 | 2263  | 9.745   | <.0001  |
| Longdress, C2, R2 - Loot, C2, R2  | 471.5687 | 70.4123 | 2263  | 6.976   | <.0001  |
| Longdress, C2, R2 - Loot, C2, R3  | 14.8262 | 70.4123 | 2263  | 0.211   | 1.0000  |
| Longdress, C2, R2 - Loot, C2, R4  | -575.2226 | 70.4123 | 2263  | -8.169  | <.0001  |
| Longdress, C2, R2 - Red and black, C1, R1 | -135.0733 | 70.4123 | 2263  | -1.918  | 0.9915  |
| Longdress, C2, R2 - Red and black, C1, R2 | -629.4031 | 70.4123 | 2263  | -8.939  | <.0001  |
| Longdress, C2, R2 - Red and black, C1, R3 | -865.9421 | 70.4123 | 2263  | -12.298 | <.0001  |
| Longdress, C2, R2 - Red and black, C1, R4 | -893.5821 | 70.4123 | 2263  | -12.691 | <.0001  |
| Longdress, C2, R2 - Red and black, C2, R1 | 418.3887 | 70.4123 | 2263  | 5.942   | <.0001  |
| Longdress, C2, R2 - Red and black, C2, R2 | 131.7000 | 70.4123 | 2263  | 1.870   | 0.9942  |
| Longdress, C2, R2 - Red and black, C2, R3 | -284.5190 | 70.4123 | 2263  | -4.041  | 0.0199  |
| Longdress, C2, R2 - Red and black, C2, R4 | -625.6964 | 70.4123 | 2263  | -8.886  | <.0001  |
| Longdress, C2, R2 - Soldier, C1, R1 | 432.1149 | 70.4123 | 2263  | 6.137   | <.0001  |
| Longdress, C2, R2 - Soldier, C1, R2 | -177.2913 | 70.4123 | 2263  | -2.518  | 0.7962  |
| Longdress, C2, R2 - Soldier, C1, R3 | -609.6062 | 70.4123 | 2263  | -8.658  | <.0001  |
| Longdress, C2, R2 - Soldier, C1, R4 | -800.8523 | 70.4123 | 2263  | -11.374 | <.0001  |
| Longdress, C2, R2 - Soldier, C2, R1 | 729.2436 | 70.4123 | 2263  | 10.357  | <.0001  |
| Longdress, C2, R2 - Soldier, C2, R2 | 514.2826 | 70.4123 | 2263  | 7.304   | <.0001  |
| Longdress, C2, R2 - Soldier, C2, R3 | 75.7631 | 70.4123 | 2263  | 1.076   | 1.0000  |
Table 8 – Continued from previous page

| contrast | estimate | SE     | df | t.ratio | p.value |
|----------|----------|--------|----|---------|---------|
| Longdress, C2, R2 - Soldier, C2, R4 | -545.6123 | 70.4123 | 2263 | -7.749  | <.0001  |
| Longdress, C2, R3 - Longdress, C2, R4 | -264.6205 | 70.4123 | 2263 | -3.758  | 0.0548  |
| Longdress, C2, R3 - Loot, C1, R1 | 689.4149  | 70.4123 | 2263 | 9.791   | <.0001  |
| Longdress, C2, R3 - Loot, C1, R2 | 18.2600   | 70.4123 | 2263 | 0.259   | 1.0000  |
| Longdress, C2, R3 - Loot, C1, R3 | -224.6528 | 70.4123 | 2263 | -3.191  | 0.2771  |
| Longdress, C2, R3 - Loot, C2, R1 | 1257.0226 | 70.4123 | 2263 | 17.852  | <.0001  |
| Longdress, C2, R3 - Loot, C2, R2 | 1042.4354 | 70.4123 | 2263 | 14.805  | <.0001  |
| Longdress, C2, R3 - Loot, C2, R3 | 585.6928  | 70.4123 | 2263 | 8.318   | <.0001  |
| Longdress, C2, R3 - Loot, C2, R4 | -4.3559   | 70.4123 | 2263 | -0.062  | 1.0000  |
| Longdress, C2, R3 - Red and black, C1, R1 | 435.7933 | 70.4123 | 2263 | 6.189   | <.0001  |
| Longdress, C2, R3 - Red and black, C1, R2 | -58.5364 | 70.4123 | 2263 | -0.831  | 1.0000  |
| Longdress, C2, R3 - Red and black, C1, R3 | -295.0754 | 70.4123 | 2263 | -4.191  | 0.0111  |
| Longdress, C2, R3 - Red and black, C1, R4 | -322.7154 | 70.4123 | 2263 | -4.583  | 0.0021  |
| Longdress, C2, R3 - Red and black, C2, R1 | 989.2554  | 70.4123 | 2263 | 14.049  | <.0001  |
| Longdress, C2, R3 - Red and black, C2, R2 | 702.5667  | 70.4123 | 2263 | 9.978   | <.0001  |
| Longdress, C2, R3 - Red and black, C2, R3 | 286.3477  | 70.4123 | 2263 | 4.067   | 0.0180  |
| Longdress, C2, R3 - Red and black, C2, R4 | -54.8297  | 70.4123 | 2263 | -0.779  | 1.0000  |
| Longdress, C2, R4 - Loot, C1, R1 | -23.8487  | 70.4123 | 2263 | -0.339  | 1.0000  |
| Longdress, C2, R4 - Loot, C1, R2 | -13.0526  | 70.4123 | 2263 | -0.186  | 1.0000  |
| Longdress, C2, R4 - Loot, C1, R3 | 37.6854   | 70.4123 | 2263 | 0.534   | 1.0000  |
| Longdress, C2, R4 - Loot, C1, R4 | -14.9260  | 70.4123 | 2263 | -0.208  | 1.0000  |
| Longdress, C2, R4 - Loot, C2, R1 | -253.8799 | 70.4123 | 2263 | -3.501  | 0.0003  |
| Longdress, C2, R4 - Loot, C2, R2 | 1024.4532 | 70.4123 | 2263 | 14.501  | <.0001  |
| Longdress, C2, R4 - Loot, C2, R3 | -433.8833 | 70.4123 | 2263 | -6.189  | <.0001  |
| Longdress, C2, R4 - Loot, C2, R4 | -64.0948  | 70.4123 | 2263 | -0.911  | 1.0000  |
| Longdress, C2, R4 - Red and black, C1, R1 | 700.4138 | 70.4123 | 2263 | 9.947   | <.0001  |
| Longdress, C2, R4 - Red and black, C1, R2 | 206.0841 | 70.4123 | 2263 | 2.927   | 0.4726  |
| Longdress, C2, R4 - Red and black, C1, R3 | -30.4549  | 70.4123 | 2263 | -0.433  | 1.0000  |
| Longdress, C2, R4 - Red and black, C1, R4 | -58.0949  | 70.4123 | 2263 | -0.825  | 1.0000  |
| Longdress, C2, R4 - Red and black, C2, R1 | 1253.8759 | 70.4123 | 2263 | 17.808  | <.0001  |
| Longdress, C2, R4 - Red and black, C2, R2 | 967.1872  | 70.4123 | 2263 | 13.736  | <.0001  |
| Longdress, C2, R4 - Red and black, C2, R3 | 550.9682  | 70.4123 | 2263 | 7.825   | <.0001  |
| Longdress, C2, R4 - Red and black, C2, R4 | 209.7908  | 70.4123 | 2263 | 2.979   | 0.4300  |
| Longdress, C2, R4 - Soldier, C1, R1 | 1267.6021 | 70.4123 | 2263 | 18.003  | <.0001  |
| Longdress, C2, R4 - Soldier, C1, R2 | 658.1959  | 70.4123 | 2263 | 9.348   | <.0001  |
| Longdress, C2, R4 - Soldier, C1, R3 | 225.8810  | 70.4123 | 2263 | 3.208   | 0.2661  |
| Longdress, C2, R4 - Soldier, C1, R4 | 34.6349   | 70.4123 | 2263 | 0.492   | 1.0000  |
| Longdress, C2, R4 - Soldier, C2, R1 | 1564.7308 | 70.4123 | 2263 | 22.222  | <.0001  |
| Longdress, C2, R4 - Soldier, C2, R2 | 1349.7697 | 70.4123 | 2263 | 19.170  | <.0001  |
| Longdress, C2, R4 - Soldier, C2, R3 | 911.2503  | 70.4123 | 2263 | 12.942  | <.0001  |
| Longdress, C2, R4 - Soldier, C2, R4 | 289.8749  | 70.4123 | 2263 | 4.117   | 0.0149  |
| Loot, C1, R1 - Loot, C1, R2 | -671.1549 | 70.4123 | 2263 | -9.532  | <.0001  |
| contrast                      | estimate | SE    | df     | t.ratio | p.value     |
|-------------------------------|----------|-------|--------|---------|-------------|
| Loot, C1, R1 - Loot, C1, R3  | -914.0677| 70.4123| 2263   | -12.982 | <.0001      |
| Loot, C1, R1 - Loot, C1, R4  | -982.2672| 70.4123| 2263   | -13.950 | <.0001      |
| Loot, C1, R1 - Loot, C2, R1  | 567.6077 | 70.4123| 2263   | 8.061   | <.0001      |
| Loot, C1, R1 - Loot, C2, R2  | 353.0205 | 70.4123| 2263   | 5.014   | 0.0003      |
| Loot, C1, R1 - Loot, C2, R3  | -103.7221| 70.4123| 2263   | -1.473  | 0.9999      |
| Loot, C1, R1 - Loot, C2, R4  | -693.7708| 70.4123| 2263   | -9.853  | <.0001      |
| Loot, C1, R1 - Red and black, C1, R1 | -253.6215| 70.4123| 2263   | -3.602  | 0.0908      |
| Loot, C1, R1 - Red and black, C1, R2 | -747.9513| 70.4123| 2263   | -10.622 | <.0001      |
| Loot, C1, R1 - Red and black, C1, R3 | -984.4903| 70.4123| 2263   | -13.982 | <.0001      |
| Loot, C1, R1 - Red and black, C1, R4 | -1012.1303| 70.4123| 2263   | -14.374 | <.0001      |
| Loot, C1, R1 - Red and black, C2, R1 | 299.8405 | 70.4123| 2263   | 4.258   | 0.0084      |
| Loot, C1, R1 - Red and black, C2, R2 | 13.1518 | 70.4123| 2263   | 0.187   | 1.0000      |
| Loot, C1, R1 - Red and black, C2, R3 | -403.0672| 70.4123| 2263   | -5.724  | <.0001      |
| Loot, C1, R1 - Red and black, C2, R4 | -744.2446| 70.4123| 2263   | -10.570 | <.0001      |
| Loot, C1, R2 - Loot, C1, R3  | 313.5667 | 70.4123| 2263   | 4.453   | 0.0037      |
| Loot, C1, R2 - Loot, C1, R4  | -295.8395| 70.4123| 2263   | -4.202  | 0.0106      |
| Loot, C1, R2 - Loot, C2, R1  | 1238.7626| 70.4123| 2263   | 17.593  | <.0001      |
| Loot, C1, R2 - Loot, C2, R2  | 1024.1754| 70.4123| 2263   | 14.545  | <.0001      |
| Loot, C1, R2 - Loot, C2, R3  | 567.4328 | 70.4123| 2263   | 8.059   | <.0001      |
| Loot, C1, R2 - Loot, C2, R4  | -22.6159 | 70.4123| 2263   | -0.321  | 1.0000      |
| Loot, C1, R2 - Red and black, C1, R1 | 417.5333 | 70.4123| 2263   | 5.930   | <.0001      |
| Loot, C1, R2 - Red and black, C1, R2 | -76.7964 | 70.4123| 2263   | -1.091  | 1.0000      |
| Loot, C1, R2 - Red and black, C1, R3 | -313.3354| 70.4123| 2263   | -4.450  | 0.0037      |
| Loot, C1, R2 - Red and black, C1, R4 | -340.9754| 70.4123| 2263   | -4.843  | 0.0006      |
| Loot, C1, R2 - Red and black, C2, R1 | 970.9954 | 70.4123| 2263   | 13.790  | <.0001      |
| Loot, C1, R2 - Red and black, C2, R2 | 684.3067 | 70.4123| 2263   | 9.719   | <.0001      |
| Loot, C1, R2 - Red and black, C2, R3 | 268.0877 | 70.4123| 2263   | 3.807   | 0.0464      |
| Loot, C1, R2 - Red and black, C2, R4 | -73.0897 | 70.4123| 2263   | -1.038  | 1.0000      |
| Loot, C1, R2 - Soldier, C1, R1 | 984.7215 | 70.4123| 2263   | 13.985  | <.0001      |
| Loot, C1, R2 - Soldier, C1, R2 | 375.3154 | 70.4123| 2263   | 5.330   | 0.0001      |
| Loot, C1, R2 - Soldier, C1, R3 | -56.9995 | 70.4123| 2263   | -0.810  | 1.0000      |
| Loot, C1, R2 - Soldier, C1, R4 | -248.2456| 70.4123| 2263   | -3.526  | 0.1144      |
| Loot, C1, R2 - Soldier, C2, R1 | 1281.8503| 70.4123| 2263   | 18.205  | <.0001      |
| Loot, C1, R2 - Soldier, C2, R2 | 1066.8892| 70.4123| 2263   | 15.152  | <.0001      |
| Loot, C1, R2 - Soldier, C2, R3 | 628.3697 | 70.4123| 2263   | 8.924   | <.0001      |
| Loot, C1, R2 - Soldier, C2, R4 | 6.9944  | 70.4123| 2263   | 0.099   | 1.0000      |
| Loot, C1, R3 - Loot, C1, R4  | -68.1995 | 70.4123| 2263   | -0.969  | 1.0000      |
| Loot, C1, R3 - Loot, C2, R1  | 1481.6754| 70.4123| 2263   | 21.043  | <.0001      |
| Loot, C1, R3 - Loot, C2, R2  | 1267.0882| 70.4123| 2263   | 17.995  | <.0001      |
| Loot, C1, R3 - Loot, C2, R3  | 810.3456 | 70.4123| 2263   | 11.509  | <.0001      |
| Loot, C1, R3 - Loot, C2, R4  | 220.2969 | 70.4123| 2263   | 3.129   | 0.3184      |
| Loot, C1, R3 - Red and black, C1, R1 | 660.4462 | 70.4123| 2263   | 9.380   | <.0001      |
| Loot, C1, R3 - Red and black, C1, R2 | 166.1164 | 70.4123| 2263   | 2.359   | 0.8868      |
| contrast | estimate | SE  | df | t.ratio | p.value |
|----------|----------|-----|----|---------|---------|
| Loot, C1, R3 - Red and black, C1, R3 | -70.4226 | 70.4123 | 2263 | -1.000 | 1.0000 |
| Loot, C1, R3 - Red and black, C1, R4 | -98.0626 | 70.4123 | 2263 | -1.393 | 1.0000 |
| Loot, C1, R3 - Red and black, C2, R1 | 1213.9082 | 70.4123 | 2263 | 17.240 | <.0001 |
| Loot, C1, R3 - Red and black, C2, R2 | 927.2195 | 70.4123 | 2263 | 13.168 | <.0001 |
| Loot, C1, R3 - Red and black, C2, R3 | 1227.6344 | 70.4123 | 2263 | 17.435 | <.0001 |
| Loot, C1, R3 - Red and black, C2, R4 | 618.2282 | 70.4123 | 2263 | 8.780 | <.0001 |
| Loot, C1, R3 - Soldier, C1, R1 | 185.9133 | 70.4123 | 2263 | 2.640 | 0.7079 |
| Loot, C1, R3 - Soldier, C1, R2 | -5.3328 | 70.4123 | 2263 | -0.076 | 1.0000 |
| Loot, C1, R3 - Soldier, C1, R3 | 1227.6344 | 70.4123 | 2263 | 17.435 | <.0001 |
| Loot, C1, R3 - Soldier, C1, R4 | 927.2195 | 70.4123 | 2263 | 13.168 | <.0001 |
| Loot, C1, R4 - Loot, C2, R1 | 1549.8749 | 70.4123 | 2263 | 22.011 | <.0001 |
| Loot, C1, R4 - Loot, C2, R2 | 1335.2877 | 70.4123 | 2263 | 18.964 | <.0001 |
| Loot, C1, R4 - Loot, C2, R3 | 871.2826 | 70.4123 | 2263 | 12.374 | <.0001 |
| Loot, C1, R4 - Loot, C2, R4 | 249.9072 | 70.4123 | 2263 | 3.549 | 0.1066 |
| Loot, C2, R1 - Loot, C2, R2 | -214.5872 | 70.4123 | 2263 | -3.048 | 0.3771 |
| Loot, C2, R1 - Loot, C2, R3 | -671.3297 | 70.4123 | 2263 | -9.534 | <.0001 |
| Loot, C2, R1 - Loot, C2, R4 | -1261.3785 | 70.4123 | 2263 | -17.914 | <.0001 |
| Loot, C2, R1 - Red and black, C1, R1 | -821.2292 | 70.4123 | 2263 | -11.663 | <.0001 |
| Loot, C2, R1 - Red and black, C1, R2 | -1315.5590 | 70.4123 | 2263 | -18.684 | <.0001 |
| Loot, C2, R1 - Red and black, C1, R3 | -1552.0979 | 70.4123 | 2263 | -22.043 | <.0001 |
| Loot, C2, R1 - Red and black, C1, R4 | -1592.9626 | 70.4123 | 2263 | -22.623 | <.0001 |
| Loot, C2, R1 - Red and black, C2, R1 | 43.0877 | 70.4123 | 2263 | 0.612 | 1.0000 |
| Loot, C2, R1 - Red and black, C2, R2 | -171.8733 | 70.4123 | 2263 | -2.441 | 0.8439 |
Table 8 – Continued from previous page

| contrast                        | estimate | SE  | df | t.ratio | p.value |
|---------------------------------|----------|-----|----|---------|---------|
| Loot, C2, R1 - Soldier, C2, R3 | -610.3928| 70.4123| 2263 | -8.669  | <.0001  |
| Loot, C2, R1 - Soldier, C2, R4 | -1231.7682| 70.4123| 2263 | -17.494 | <.0001  |
| Loot, C2, R2 - Loot, C2, R3    | -456.7426| 70.4123| 2263 | -6.487  | <.0001  |
| Loot, C2, R2 - Loot, C2, R4    | -1046.7913| 70.4123| 2263 | -14.867 | <.0001  |
| Loot, C2, R2 - Red and black, C1, R1 | -606.6421| 70.4123| 2263 | -8.616  | <.0001  |
| Loot, C2, R2 - Red and black, C1, R2 | -1100.9718| 70.4123| 2263 | -15.636 | <.0001  |
| Loot, C2, R2 - Red and black, C1, R3 | -1337.5108| 70.4123| 2263 | -18.995 | <.0001  |
| Loot, C2, R2 - Red and black, C1, R4 | -1365.1508| 70.4123| 2263 | -19.388 | <.0001  |
| Loot, C2, R2 - Red and black, C2, R1 | -53.1800 | 70.4123| 2263 | -0.755  | 1.0000  |
| Loot, C2, R2 - Red and black, C2, R2 | -339.8687| 70.4123| 2263 | -4.827  | 0.0007  |
| Loot, C2, R2 - Red and black, C2, R3 | -756.0877| 70.4123| 2263 | -10.738 | <.0001  |
| Loot, C2, R2 - Red and black, C2, R4 | -1097.2651| 70.4123| 2263 | -15.583 | <.0001  |
| Loot, C2, R2 - Red and black, C2, R1 | -468.7600| 70.4123| 2263 | -6.072  | <.0001  |
| Loot, C2, R2 - Red and black, C2, R2 | -1659.7482| 70.4123| 2263 | -23.785 | <.0001  |
| Loot, C2, R2 - Red and black, C2, R3 | -2882.9785| 70.4123| 2263 | -41.687 | <.0001  |
| Loot, C2, R2 - Red and black, C2, R4 | -2882.9785| 70.4123| 2263 | -41.687 | <.0001  |
Table 8 – Continued from previous page

| contrast                  | estimate | SE    | df   | t.ratio | p.value |
|---------------------------|----------|-------|------|---------|---------|
| Loot, C2, R4 - Soldier, C2, R3 | 650.9856 | 70.4123 | 2263 | 9.245   | <.0001  |
| Loot, C2, R4 - Soldier, C2, R4 | 29.6103  | 70.4123 | 2263 | 0.421   | 1.0000  |
| Red and black, C1, R1 - Red and black, C1, R2 | -494.3297 | 70.4123 | 2263 | -7.021  | <.0001  |
| Red and black, C1, R1 - Red and black, C1, R3 | -730.8687 | 70.4123 | 2263 | -10.772 | <.0001  |
| Red and black, C1, R1 - Red and black, C1, R4 | -758.5087 | 70.4123 | 2263 | -10.380 | <.0001  |
| Red and black, C1, R1 - Red and black, C2, R1 | 553.4621  | 70.4123 | 2263 | 7.860   | <.0001  |
| Red and black, C1, R1 - Red and black, C2, R2 | 266.7733  | 70.4123 | 2263 | 3.789   | 0.0494  |
| Red and black, C1, R1 - Red and black, C2, R3 | -149.4456 | 70.4123 | 2263 | -2.122  | 0.9655  |
| Red and black, C1, R1 - Red and black, C2, R4 | -490.6231 | 70.4123 | 2263 | -6.968  | <.0001  |
| Red and black, C1, R2 - Red and black, C2, R1 | 1047.7918 | 70.4123 | 2263 | 14.881  | <.0001  |
| Red and black, C1, R2 - Red and black, C2, R2 | 761.1031  | 70.4123 | 2263 | 10.809  | <.0001  |
| Red and black, C1, R2 - Red and black, C2, R3 | 344.8841  | 70.4123 | 2263 | 4.898   | 0.0005  |
| Red and black, C1, R2 - Red and black, C2, R4 | 3.7067    | 70.4123 | 2263 | 0.053   | 1.0000  |
| Red and black, C1, R3 - Red and black, C2, R1 | 1284.3308 | 70.4123 | 2263 | 18.240  | <.0001  |
| Red and black, C1, R3 - Red and black, C2, R2 | 997.6421  | 70.4123 | 2263 | 14.169  | <.0001  |
| Red and black, C1, R3 - Red and black, C2, R3 | 581.4231  | 70.4123 | 2263 | 8.257   | <.0001  |
| Red and black, C1, R3 - Red and black, C2, R4 | 240.2456  | 70.4123 | 2263 | 3.412   | 0.1582  |
| Red and black, C1, R4 - Red and black, C2, R1 | 1325.6969 | 70.4123 | 2263 | 18.435  | <.0001  |
| Red and black, C1, R4 - Red and black, C2, R2 | 716.2908  | 70.4123 | 2263 | 10.173  | <.0001  |
| Red and black, C1, R4 - Red and black, C2, R3 | 1325.6969 | 70.4123 | 2263 | 18.435  | <.0001  |
| contrast                  | estimate   | SE       | df  | t.ratio | p.value |
|---------------------------|------------|----------|-----|---------|---------|
| Red and black, C1, R4 - Soldier, C1, R4 | 92.7297 | 70.4123 | 2263 | 1.317   | 1.0000  |
| Red and black, C1, R4 - Soldier, C2, R2 | 1407.8646 | 70.4123 | 2263 | 19.995  | <.0001  |
| Red and black, C1, R4 - Soldier, C2, R3 | 969.3451 | 70.4123 | 2263 | 13.767  | <.0001  |
| Red and black, C2, R1 - Red and black, C2, R2 | -286.6887 | 70.4123 | 2263 | -4.072  | 0.0004  |
| Red and black, C2, R1 - Red and black, C2, R3 | -702.9077 | 70.4123 | 2263 | -9.983  | <.0001  |
| Red and black, C2, R1 - Red and black, C2, R4 | -1044.0851 | 70.4123 | 2263 | -14.828 | <.0001  |
| Red and black, C2, R1 - Soldier, C1, R1 | 13.7262 | 70.4123 | 2263 | 0.195   | 1.0000  |
| Red and black, C2, R1 - Soldier, C1, R2 | -595.6800 | 70.4123 | 2263 | -8.460  | <.0001  |
| Red and black, C2, R1 - Soldier, C1, R3 | -1027.9949 | 70.4123 | 2263 | -14.600 | <.0001  |
| Red and black, C2, R1 - Soldier, C1, R4 | -1219.2410 | 70.4123 | 2263 | -17.316 | <.0001  |
| Red and black, C2, R1 - Soldier, C2, R1 | 310.8549 | 70.4123 | 2263 | 4.415   | 0.0044  |
| Red and black, C2, R1 - Soldier, C2, R2 | 95.8938  | 70.4123 | 2263 | 1.362   | 1.0000  |
| Red and black, C2, R1 - Soldier, C2, R3 | -342.6256 | 70.4123 | 2263 | -4.866  | 0.0006  |
| Red and black, C2, R1 - Soldier, C2, R4 | -341.7174 | 70.4123 | 2263 | -4.845  | <.0001  |
| Red and black, C2, R2 - Red and black, C2, R3 | -341.7174 | 70.4123 | 2263 | -4.845  | <.0001  |
| Red and black, C2, R2 - Red and black, C2, R4 | -107.2277 | 70.4123 | 2263 | -1.523  | 0.9999  |
| Red and black, C2, R2 - Soldier, C1, R1 | 716.6338 | 70.4123 | 2263 | 10.178  | <.0001  |
| Red and black, C2, R2 - Soldier, C1, R2 | 107.2277 | 70.4123 | 2263 | 1.523   | 0.9999  |
| Red and black, C2, R2 - Soldier, C1, R3 | -325.0872 | 70.4123 | 2263 | -4.617  | 0.0018  |
| Red and black, C2, R2 - Soldier, C1, R4 | -516.3333 | 70.4123 | 2263 | -7.333  | <.0001  |
| Red and black, C2, R2 - Soldier, C2, R1 | 1013.7626 | 70.4123 | 2263 | 14.398  | <.0001  |
| Red and black, C2, R2 - Soldier, C2, R2 | 798.8015 | 70.4123 | 2263 | 11.345  | <.0001  |
| Red and black, C2, R2 - Soldier, C2, R3 | 360.2821 | 70.4123 | 2263 | 5.117   | 0.0002  |
| Red and black, C2, R2 - Soldier, C2, R4 | -261.0933 | 70.4123 | 2263 | -3.708  | 0.0647  |
| Red and black, C2, R4 - Soldier, C1, R1 | 1057.8113 | 70.4123 | 2263 | 15.023  | <.0001  |
| Red and black, C2, R4 - Soldier, C1, R2 | 448.4051 | 70.4123 | 2263 | 6.368   | <.0001  |
| Red and black, C2, R4 - Soldier, C1, R3 | 16.0903 | 70.4123 | 2263 | 0.229   | 1.0000  |
| Red and black, C2, R4 - Soldier, C1, R4 | -175.1559 | 70.4123 | 2263 | -2.488  | 0.8158  |
| Red and black, C2, R4 - Soldier, C2, R1 | 1354.9400 | 70.4123 | 2263 | 19.243  | <.0001  |
| Red and black, C2, R4 - Soldier, C2, R2 | 1139.7970 | 70.4123 | 2263 | 16.190  | <.0001  |
| Red and black, C2, R4 - Soldier, C2, R3 | 701.4595 | 70.4123 | 2263 | 9.962   | <.0001  |
| Red and black, C2, R4 - Soldier, C2, R4 | 80.0841 | 70.4123 | 2263 | 1.137   | 1.0000  |
| Soldier, C1, R1 - Soldier, C1, R2 | -609.4062 | 70.4123 | 2263 | -8.655  | <.0001  |
| Soldier, C1, R1 - Soldier, C1, R3 | -1041.7210 | 70.4123 | 2263 | -14.795 | <.0001  |
| Soldier, C1, R1 - Soldier, C1, R4 | -1232.9672 | 70.4123 | 2263 | -17.511 | <.0001  |
| Soldier, C1, R1 - Soldier, C2, R1 | 297.1287 | 70.4123 | 2263 | 4.220   | 0.0099  |
| Soldier, C1, R1 - Soldier, C2, R2 | 82.1677 | 70.4123 | 2263 | 1.167   | 1.0000  |
| Soldier, C1, R1 - Soldier, C2, R3 | -356.3518 | 70.4123 | 2263 | -5.061  | 0.0002  |
| Soldier, C1, R1 - Soldier, C2, R4 | -977.7272 | 70.4123 | 2263 | -13.886 | <.0001  |
| Soldier, C1, R2 - Soldier, C1, R3 | -432.3149 | 70.4123 | 2263 | -6.140  | <.0001  |
Table 8 – Continued from previous page

| contrast | estimate  | SE    | df   | t.ratio | p.value |
|----------|-----------|-------|------|---------|---------|
| Soldier, C1, R2 - Soldier, C1, R4 | -623.5610 | 70.4123 | 2263 | -8.856 | <.0001 |
| Soldier, C1, R2 - Soldier, C2, R1 | 906.5349  | 70.4123 | 2263 | 12.875  | <.0001 |
| Soldier, C1, R2 - Soldier, C2, R2 | 691.5738  | 70.4123 | 2263 | 9.822   | <.0001 |
| Soldier, C1, R2 - Soldier, C2, R3 | 253.0544  | 70.4123 | 2263 | 3.594   | 0.0931  |
| Soldier, C1, R3 - Soldier, C1, R4 | 1338.8497 | 70.4123 | 2263 | 19.014  | <.0001  |
| Soldier, C1, R3 - Soldier, C2, R1 | 1123.8887 | 70.4123 | 2263 | 15.962  | <.0001  |
| Soldier, C1, R3 - Soldier, C2, R2 | 685.3692  | 70.4123 | 2263 | 9.734   | <.0001  |
| Soldier, C1, R3 - Soldier, C2, R3 | 63.9938   | 70.4123 | 2263 | 1.0000  | <.0001  |
| Soldier, C1, R3 - Soldier, C2, R4 | -1530.0959| 70.4123 | 2263 | 21.731  | <.0001  |
| Soldier, C2, R1 - Soldier, C2, R2 | 1315.1349 | 70.4123 | 2263 | 18.678  | <.0001  |
| Soldier, C2, R1 - Soldier, C2, R3 | 876.6154  | 70.4123 | 2263 | 12.450  | <.0001  |
| Soldier, C2, R1 - Soldier, C2, R4 | 255.2400  | 70.4123 | 2263 | 0.0845  | <.0001  |
| Soldier, C2, R2 - Soldier, C2, R3 | -214.9610 | 70.4123 | 2263 | -3.053  | 0.3731  |
| Soldier, C2, R2 - Soldier, C2, R4 | -1274.8559| 70.4123 | 2263 | -18.106 | <.0001  |
| Soldier, C2, R3 - Soldier, C2, R4 | -438.5195 | 70.4123 | 2263 | -6.228  | <.0001  |
| Soldier, C2, R1 - Soldier, C2, R2 | -653.4805 | 70.4123 | 2263 | -9.281  | <.0001  |
| Soldier, C2, R1 - Soldier, C2, R3 | -1059.8949| 70.4123 | 2263 | -15.053 | <.0001  |
| Soldier, C2, R1 - Soldier, C2, R4 | -621.3754 | 70.4123 | 2263 | -8.825  | <.0001  |

Results are averaged over the levels of: DoF

Degrees-of-freedom method: kenward-roger

P value adjustment: tukey method for comparing a family of 32 estimates

Table 9: Contrast test between factors DoF and Content

| contrast          | estimate  | SE    | df   | t.ratio | p.value |
|-------------------|-----------|-------|------|---------|---------|
| C1, R1 - C1, R2   | -263.7783 | 40.7258 | 2294 | -6.477  | <.0001  |
| C1, R1 - C1, R3   | -803.3192 | 40.7258 | 2294 | -19.725 | <.0001  |
| C1, R1 - C1, R4   | -962.2388 | 40.7258 | 2294 | -23.627 | <.0001  |
| C1, R1 - C2, R1   | 281.7562  | 40.7258 | 2294 | 6.918   | <.0001  |
| C1, R1 - C2, R2   | -140.8103 | 40.7258 | 2294 | -3.458  | 0.0129  |
| C1, R1 - C2, R3   | -654.9223 | 40.7258 | 2294 | -16.081 | <.0001  |
| C1, R1 - C2, R4   | -925.6729 | 40.7258 | 2294 | -22.729 | <.0001  |
| C1, R2 - C1, R3   | -539.5409 | 40.7258 | 2294 | -13.248 | <.0001  |
| C1, R2 - C1, R4   | -698.4605 | 40.7258 | 2294 | -17.150 | <.0001  |
| C1, R2 - C2, R1   | 545.5345  | 40.7258 | 2294 | 13.395  | <.0001  |
| C1, R2 - C2, R2   | 122.9681  | 40.7258 | 2294 | 3.019   | 0.0521  |
| C1, R2 - C2, R3   | -391.1440 | 40.7258 | 2294 | -9.604  | <.0001  |
| C1, R2 - C2, R4   | -661.8946 | 40.7258 | 2294 | -16.252 | <.0001  |
| C1, R3 - C1, R4   | -158.9196 | 40.7258 | 2294 | -3.902  | 0.0067  |
| C1, R3 - C2, R1   | 1085.0754 | 40.7258 | 2294 | 26.643  | <.0001  |
| C1, R3 - C2, R2   | 662.5090  | 40.7258 | 2294 | 16.268  | <.0001  |
| C1, R3 - C2, R3   | 148.3969  | 40.7258 | 2294 | 3.644   | 0.0067  |
| C1, R3 - C2, R4   | -122.3537 | 40.7258 | 2294 | -3.004  | 0.0545  |
| C1, R4 - C2, R1   | 1243.9950 | 40.7258 | 2294 | 30.546  | <.0001  |
| C1, R4 - C2, R2   | 821.4286  | 40.7258 | 2294 | 20.170  | <.0001  |
| C1, R4 - C2, R3   | 307.3165  | 40.7258 | 2294 | 7.546   | <.0001  |
| C1, R4 - C2, R4   | 36.5659   | 40.7258 | 2294 | 0.898   | 0.9863  |
Table 9 – Continued from previous page

| contrast         | estimate  | SE    | df  | t.ratio | p.value |
|------------------|-----------|-------|-----|---------|---------|
| C2, R1 - C2, R2  | -422.5664 | 40.7258 | 2294 | -10.376 | <.0001  |
| C2, R1 - C2, R3  | -936.6785 | 40.7258 | 2294 | -23.000 | <.0001  |
| C2, R1 - C2, R4  | -1207.4291| 40.7258 | 2294 | -29.648 | <.0001  |
| C2, R2 - C2, R3  | -514.1121 | 40.7258 | 2294 | -12.624 | <.0001  |
| C2, R2 - C2, R4  | -784.8627 | 40.7258 | 2294 | -19.272 | <.0001  |
| C2, R3 - C2, R4  | -270.7506 | 40.7258 | 2294 | -6.648  | <.0001  |

Results are averaged over the levels of: DoF, Contents

Degrees-of-freedom method: kenward-roger

P value adjustment: tukey method for comparing a family of 8 estimates