sciendo

GESTALT THEORY, DOI 10.2478/gth-2022-0006
© 2022 (ISSN 2519-5808); Vol. 44, No. 1-2, 61–78

Original Contributions - Originalbeiträge

Nicole Ruta, Alistair Burleigh & Robert Pepperell

Space and Scale in Medieval Painting Reflects Imagination and Perception

1. Introduction

Before the discovery of linear perspective in Italy in the early 15th century, European artists varied the size of figures in their compositions, often in a way that inverts the laws of optics and linear perspective (White, 1972). For example, in the *Transfiguration* by Duccio di Buoninsegna (c. 1308-11), the central figure of Christ is significantly larger than that of the other characters in the scene, even though he is more distant in space than the figures in the foreground (Figure 1A). By digitally modifying the original painting, as shown in Figure 1B, we can see how the same scene would appear if it followed more closely the laws of optics.

The construction of medieval pictorial space is sometimes referred to as 'inverse' or 'reverse perspective', a term originally coined by the art historian Oskar Wulff in 1907, and the meaning of these pictorial tendencies has since provoked extensive debates in art history (Bunim, 1942; White, 1972, Antonova, 2010, Meiss, 1946), aesthetics (Arnheim, 1972), vision science (Todorovic, 2008) and in contemporary art (Hockney, 2001). The debates have centred around the purpose and meaning of this method of depicting visual space, with two major theses being proposed:

- the standard 'hierarchical scaling' thesis, which claims that artists manipulated the relative sizes of characters following conventions about the status of the figures (Doehlemann, 1910; Bunim, 1942; White, 1972);
- the 'inner view' thesis, according to which artists resized characters in their compositions to direct visual attention 'inside' the paintings towards the most prominent character in the scene and to emotionally engage the viewer by assuming the character's point of view (Wulff, 1907).

1.1. Hierarchical scaling and the inner view

It was Karl Doehlemann (1910, in Antonova, 2010) who first proposed the 'hierarchical scaling' thesis, which he felt reflected the absence of an agreed logic of pictorial space among the community of medieval artists of the kind that was later to be adopted following the invention of linear perspective. According to

Open Access. © 2022 Ruta et al., published by Sciendo. BY-NC-ND This work is licensed under the Creative Commons Attribution NonCommercial-NoDerivatives 4.0 License.



Fig. 1. **(A)** Transfiguration, by Duccio di Buoninsegna (1308-11). The painting is a typical example of the use of figure size variation in the Trecento Sienese and Florentine artistic tradition, where a figure in the background – here figure of Christ – is significantly enlarged compared to characters in the foreground. **(B)** The digitally modified version follows more closely the optical rules of linear perspective: the figure of Christ is reduced to match the other two figures at his sides, and all three background figures are reduced in size so that they are smaller than the figures in the foreground.

Doehlemann, it was the arbitrary nature of medieval pictorial space that allowed figures to be represented in non-consistent ways. This thesis was largely accepted by subsequent scholars (Bunim, 1942; White, 1972), and it became the most prominent explanation of character size variations in medieval art. The gestalt psychologist Arnheim (1972) addressed the topic of figure size variation in medieval painting in the context of what he called the 'illusionistic doctrine' that suffused much art theory. According to this doctrine, the task of painters was to create compelling illusions of optical space using the same principles of projective geometry that underpin photography, i.e., linear perspective. The inadequacy of this doctrine is made evident for Arnheim by the fact that so many artists eschewed it, and in fact blatantly contradicted it, not only in medieval art but also in modernist art too. He cites the examples where artists have significantly enlarged a central figure and diminished surrounding figures in both the European and Asian traditions. Along with other scholars, Arnheim broadly accepts the hierarchical scaling thesis as the best explanation for figure size variations that contradict optical logic. This he does on the grounds that such spatial arrangements most effectively convey the artist's intended symbolic meaning.

In contrast to this view, Wulff (1907) argued that the sizing and arrangement of figures in medieval pictorial space was often an attempt to integrate an omnipotent 'God's Eye' view with the frontal view presented by the picture surface. This allowed the viewer to empathise with this view and subjectively project themselves

into the pictorial space and see figures within it accordingly. Expounding this idea led Wulff to consider two possible reasons for the figure size variations: the 'inner view' and 'scenography' theses. The 'inner view' thesis is that the viewer is mentally drawn inside the pictorial space to assume the point of view of the largest dominating figure. Consequently, objects are depicted as decreasing in size from the point of view of this figure, to facilitate the viewer looking at the scene from his/her perspective. On the other hand, the 'scenography' thesis explained reverse perspective in light of the tradition of Greek scenography, a form of spatial representation that takes into account the physical distance between the painting and the viewer and applies a set of corrections to compensate for optical distortions, which include the practice of enlarging figures that are further away to compensate for the distance. As pointed out in Antonova's review (2010), the two theses originally provided by Wulff are not exempt from the risk of contradiction. The two explanations are indeed mutually exclusive because the 'inner view' thesis implies that the viewer is *inside* the pictorial space, while the 'scenography' thesis assumes that the viewer is looking at the painting from *outside* the pictorial space. Either way, Arnheim was highly critical of Wulff's proposal on the grounds that it contradicted the logic of subjective projection.

To summarise, Wulff's 'inner view' thesis (1907) suggests, contrary to the hierarchical scaling thesis, that medieval artists deliberately manipulated the relative size of objects and figures in pictorial space to direct the viewer's attention towards the most prominent character in the scene and to emotionally engage the viewer to better empathise with this character, as both the *Maestà* by Duccio and the *Baptism of Christ* by Monaco (Figure 2) capture well.

1.2. Limits of previous theses and the role of visual imagination

One of the most problematic implications of the 'hierarchical scaling' thesis is the practice of interpreting medieval art with reference to a set of rules defined at least



Fig. 2. On the left: probably Lorenzo Monaco, Baptism of Christ (1387-8). On the right: Duccio di Buoninsegna, Maestà, central panel (1308-11).

a century after these artworks were created. Terms such as focal point or vanishing lines belong to linear perspective and are not historically relevant for medieval art (Antonova & Kemp, 2005). Moreover, it seems to perpetuate the problematic notion of the superiority of linear perspective, an idea dating back to the Renaissance period itself (Vasari, 1568). Moreover, the 'hierarchical scaling' thesis does not explain paintings such as Lorenzo Monaco's Baptism of Christ (1387-8, Figure 2) where the figure of Jesus, presumably the most important according to the Christian iconography, is represented as smaller than the figure of St. John, the Baptist. The standard 'hierarchical scaling' interpretation reduces a very complex form of art to an oversimplified explanation of a subset of its characteristics. We suggest that the 'inner view' thesis is a valuable and worth investigating alternative to the standard 'hierarchical scaling' one. To our knowledge, Wulff's 'inner view' alternative hypothesis has never been empirically tested. Here, we advance an original 'visual imagination' thesis (Pepperell & Ruta, 2017; Ruta, Burleigh, & Pepperell, 2017) that draws on Wulff's empathetic analysis. Our purpose is to study the relationship between the size of figures in medieval pictorial space and the way we perceive, experience and therefore imagine visual space.

According to Antonova and Kemp (2005), during the Middle Ages, the purpose of religious images was to engage an audience of religious worshippers in liturgical settings, within which temporal viewing processes of scanning are orchestrated by the painter's compositions in a way that serves precisely the spiritual and aesthetic needs of those to whom the images were directed. In other words, these paintings were a tool to educate the faithful: they required active engagement and the ability to 'read' the represented story (Lubbock, 2006). We therefore suggest that both due to the nature of the episodes depicted in these paintings and to fulfil those complicated requirements, it was more likely that medieval artists created these images from their memories and imagination, rather than from direct observation.

Imagined visual space is itself intimately influenced by the way we phenomenologically experience object size while looking at specific locations in visual space. We know that a fundamental property of our visual system consists in the difference between central and peripheral vision (Strasburger, Rentschler, & Jüttner, 2011). In a previous study, we showed that people overestimated the size of equally sized stimuli when presented in central compared to peripheral visual field (Baldwin, Burleigh, Pepperell, & Ruta, 2016). We hypothesise that this fundamental property of visual space was reflected in medieval paintings, showing a correspondence between how artists imagined looking at objects or figures in visual space and how they decided to represent them inside the complex religious images they constructed on their pictorial surfaces. Therefore, the use of relative figure scaling in the Trecento Sienese and Florentine artistic tradition can be interpreted as an

intentional and deliberate choice made by artists to better represent imagining looking at objects in visual space. In line with the 'inner view' thesis, we suggest that medieval artists' intentions were intrinsically different from those of Renaissance artists, with the former relying more on imagination than the latter.

The main aim of this study is to test the implications of both the alternative 'inner view' thesis (Wulff, 1907) and our original 'visual imagination' thesis (Pepperell & Ruta, 2017; Ruta et al., 2017). In relation to these two theoretical positions, we identified the two following experimental hypotheses:

A. According to the 'visual imagination' thesis, imagining looking at a target in visual space is best represented when its size is enlarged compared to other objects inside the pictorial space.
B. According to the 'inner view' thesis, participants will select the bigger and prominent character in the scene as the one attracting most their attention and with whom they most identify but not as the most important one according to Christian iconography.

To test these hypotheses, we selected 10 medieval paintings representative of the Trecento Sienese and Florentine tradition that showed key characteristics of the use of the relative size of visual cue (which we refer here to as R-Svc) and artistically altered them to create a linear perspective-like version (full list of paintings is available on Open Science Framework, see 'Methodological approach' section below). As discussed in more detail in the *Methodological approach* section below, the characters' sizes in the transformed version were artistically manipulated to be 'optically correct' according to the rules of linear perspective. In this way, we were able to investigate if altering the R-Svc:

- affected preference for the painting (Experiment 1 – Task 1);
- had a significant impact on participants' choices regarding painting's composition, when asked to identify with the artist's intentions (Experiment 1 – Tasks 2 and 2a) and
- influenced explicit measures of attention, importance and identification (Experiment 2 – first block).

1.3. Methodological approach: artistically transforming the R-Svc in medieval paintings

We will refer to the character depicted as bigger and prominent in the original version of the paintings as *ChX*. We selected 10 paintings representative of the Trecento Sienese and Florentine art tradition that showed key characteristics of the use of the R-Svc. The position of ChX in pictorial space was counterbalanced across the paintings (centre, left and right), and ChX corresponded to the figure

of Jesus in half of the images. The original image of the painting was artistically manipulated to create a linear perspective-like version, which showed characteristics similar to linear perspective in terms of relative character size. The artistic manipulation was conducted by Prof. R. Pepperell using Adobe Photoshop CC2017. The rationale was to match the size of figures according to the two following aesthetic rules:

1. the size of ChX was reduced: (a) to match the size of the closest figure placed on the same level in the pictorial space or (b) if a reference character was absent, as it would appear in a linear perspective image;
2. if the size of characters in the foreground was represented in the original version as smaller compared to the characters in the background, then their sizes were increased, as they would appear in a linear perspective image.

An anonymised version of the data and code for running the experiments are provided in the Open Science Framework repository (osf.io/wuha6/), together with the entire set of paintings used in this study ('Medieval Paintings Stimuli' folder), examples and rationale for creation of the transformed version of the paintings (Supplementary Materials: osf.io/k3dxt/).

1.4. Experiment 1: preferences for painting composition

Experiment 1 investigated whether our character size manipulation affected the following: (1) preferences for a particular version of the paintings and (2) choices for the painting composition when the participants were explicitly asked to imagine they were an 'artist interested in representing the scene, while looking at one character in particular'. The first point aimed to validate our set of stimuli to ensure that the participants were not positively biased towards the original version. We did not expect our character size manipulation to have a significant influence on the aesthetic value of the paintings because, apart from size, other aesthetic properties of the paintings did not vary between the two versions (Task 1). The second point aimed to test the visual imagination thesis: we expected that most of the participants would choose the original composition rather the transformed one (Task 2a) and would be significantly more likely to click on ChX rather than the other characters in the scene, especially if they had previously chosen the original composition (Task 2b). Due to the semantic content of the paintings selected for this study, we also controlled if preference for the original vs transformed composition was modulated by individuals' religious beliefs: as religious participants might be both more familiar with religious iconography and more likely to attribute to the RSvc, a symbolic meaning, we wanted to test if they also liked more the original compared to the transformed version of the paintings.

2. Materials and Methods

2.1. Participants

Seventy-five participants, 55 identifying as female, gave informed consent to take part in the study and were recruited via the Prolific platform (prolific.ac). The average age was 32 years (SD = 7.9; range = 20–49 years), and 60% of our participants identified themselves as non-religious (Christian = 29.3%, Buddhist = 4%, wished not to answer = 4%). The sample was balanced in terms of art interest (self-rated using a five-point Likert scale, with items corresponding to the following: 'Not at all' = 1.3%; 'Not really' = 16%; 'Neutral' = 26.6%; 'Somewhat' = 38.6%; 'Very interested' = 17.3%). All participants had to speak English as their first language to complete the study (nationality of our sample population was as follows: UK = 69.3%; USA/Canada = 29.3%; India = 1.3%). As additional criteria, we only selected participants that reported that they were using a screen >15 inches (38.9% of the participants had 15-inch screens, 25% had 17-inch screens and 36.1% had screens >17 inches; mobile phone and small tablet users were excluded) and were sitting at a desk with their computer at a comfortable distance in front of them. The average amount paid to participants in our online sample for taking part in the research was £3.75/h. The experiment was approved by the Ethics Committee of Cardiff Metropolitan University and was conducted in accordance with the Declaration of Helsinki (Williams, 2008).

2.2. Procedure

The experiment was divided into two blocks and consisted of two forced-choice tasks:

1. In the first block (Task 1), participants were asked to choose which version of the painting (original vs. transformed) they preferred.
2. In the second block (Task 2a), they were asked to choose which was the best composition if imagining being 'an artist interested in representing the scene, while looking at one character in particular'. Once selected, only the chosen composition remained on-screen, and participants were asked to click on the character that they imagined looking at (Task 2b). The two versions of the paintings were then presented on the screen at the same time, on the left-hand and the right-hand side of the screen, with the order randomised between trials.

3. Analysis and Results

3.1. Task 1: Which is your preferred version of the painting?

To investigate whether the original was preferred significantly more than the transformed version, for each participant, we calculated *preference ratio* – the total

frequencies for the original version (O_{freq}) divided by those for the transformed version (T_{freq}) – as indicated in the following equation:

$$Preference\ Ratio = (O\ freq)/(T\ freq)$$

If the *preference ratio* is equal to 1, it corresponded to chance level, meaning that the preference frequencies for the two versions are the same. If it is >1, the proportion of the frequencies reported by the original version is higher than that of the transformed version; if it is <1, the original version has lower preference frequencies than the transformed version.

Two one-sample Wilcoxon tests were performed, dividing participants into two groups according to their religious affiliation (religious vs. non-religious). The results indicated that the median of the *preference ratio* (Task 1) was not significantly different from 1 for both the religious, $Z = 2.2$, $p > 0.01$, and the non-religious groups, $Z = -1.3$, $p > 0.01$ (adjusted p values for repeated one-sample Wilcoxon tests, according to Bonferroni correction), with a moderate effect size in the case of the religious group ($r = 0.47$) and a low effect in the case of the non-religious group ($r = 0.2$). The median *preference ratio* value was 1.5 for the religious group and 1 for the non-religious. Overall, it is suggested that the proportion of participants who preferred the original over the transformed version was the same for both religious and non-religious participants, as illustrated in Figure 3.

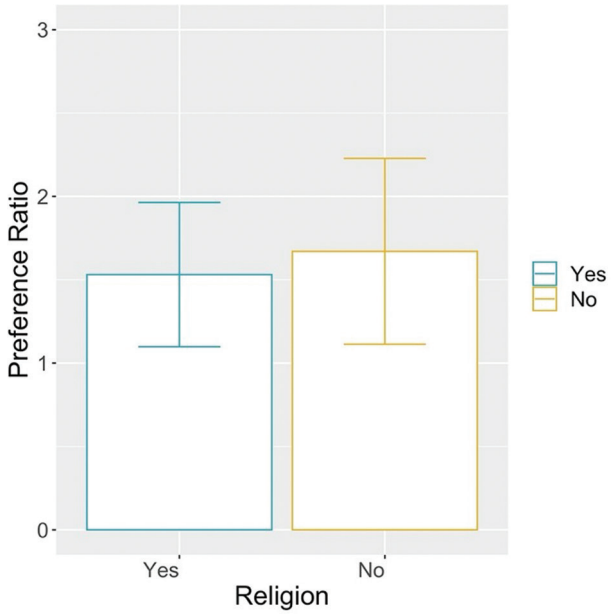


Fig. 3. Bar plot showing mean preference ratio values (Task 1), according to religious affiliation: 'Yes' is in blue; 'No' is in yellow. The cut-off value was defined as 1. Error bars represent confidence intervals.

3.2. Task 2a and Task 2b: If you were the artist, which composition would you choose?

A *composition ratio* was calculated to represent the proportions in the composition choices, described by the total frequencies for the original version (*O freq*) divided by the transformed version (*T freq*), using the same equation used for Task 1. Two one-sample Wilcoxon tests were performed for the religious and non-religious groups. The results indicated that the median *composition ratio* (Task 2a) deviated significantly from 1 in both the religious, $Z = -2.85$, $p < 0.01$, and the non-religious group, $Z = -5.63$, $p < 0.01$ (adjusted p values for repeated one-sample Wilcoxon tests, according to Bonferroni correction), with a strong effect size in both groups: $r = 0.64$ for religious and $r = 0.76$ for non-religious participants. The median *composition ratio* value for the religious group was 2.33, while for the non-religious one, it was 4. The results from Task 2a showed that, in this second task, participants were significantly more likely to choose the original composition compared to the transformed one, with participants who defined themselves as non-religious reporting a higher *composition ratio* value compared to the religious participants, as illustrated in Figure 4.

Additionally, we wanted to test (1) whether the participants clicked on ChX significantly more often than the other characters in the paintings and (2) whether the proportion of clicks was different according to the version previously chosen

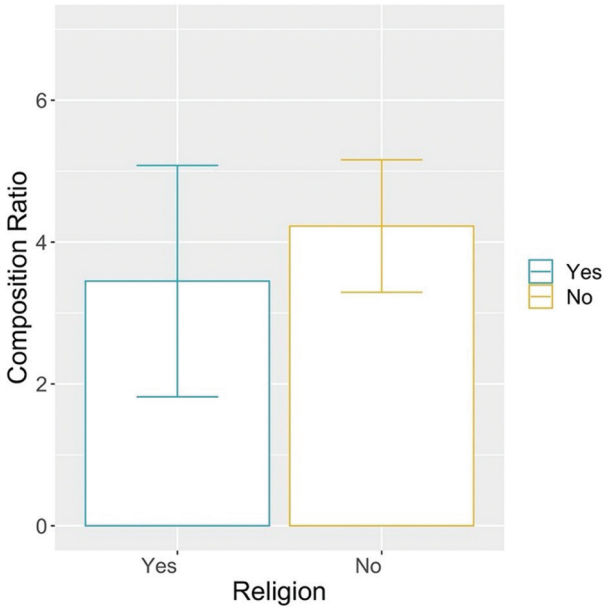


Fig. 4. Bar plot showing mean composition ratio values for composition preference (Task 2a), according to religious affiliation: 'Yes' is in blue; 'No' is in yellow. The cut-off value was defined as 1. Error bars represent confidence intervals.

in Task 2a as having a better composition. A *selections ratio* was calculated for both versions, representing the proportion of the number of times the participants clicked on ChX (ChX_{freq}) divided by the times they clicked on other characters in the paintings ($Other\ Ch_{freq}$), as indicated in the equation below:

$$Selections\ Ratio = \frac{ChX\ freq}{Other\ Ch\ freq}$$

We used Rstudio (RStudio Team, 2015) and lme (Pinheiro et al., 2017) to perform a linear mixed effects analysis to investigate the relationships between *selections ratio* and composition choice (original vs. transformed) and religious affiliation (Yes vs. No). As fixed effects, we entered composition choice, participants' religious affiliation and the interaction between the two into the model. As the random effect, we had the intercept for participants' identity. There was a significant main effect of composition choice, $\chi^2(6) = 74.98$, $p < 0.0001$, meaning that the participants clicked on ChX significantly more than the other characters in the original version (mean *selections ratio* = 5) compared to in the transformed composition (mean *selections ratio* = 0.77), regardless of participants' religious affiliation. The *selections ratio* for religious and non-religious participants were, in general, the same, at $\chi^2(7) = 0.00005$, $p = 0.996$. No significant interaction between composition choice and religious affiliation was found, $\chi^2(8) = 0.99$, $p = 0.319$, as illustrated in Figure 5.

Overall, the results showed that participants did not prefer the original version significantly more than the transformed one (Task 1). However, when asked to

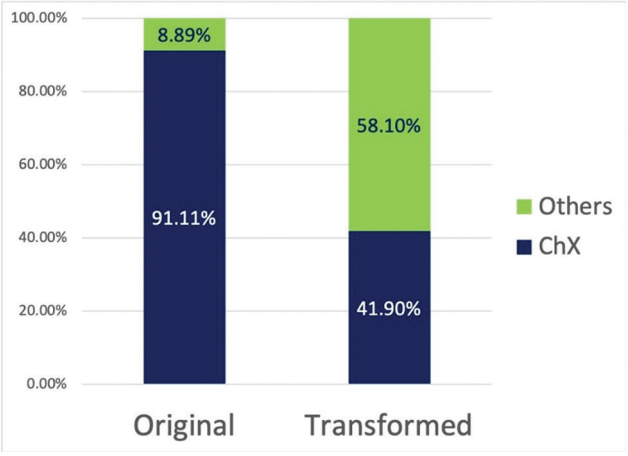


Fig. 5. The compound bar chart above shows the percentage of number times ChX (in blue) and other characters (in green) were selected as the character participants imagined looking at in Task 2b, according to the painting version previously selected in Task 2a: original, on the left-hand side, and transformed, on the right-hand side.

imagine themselves as the artist interested in depicting the scene represented in the artworks, they chose the original composition significantly more often than the transformed version (Task 2a) and reported imagining looking at ChX significantly more than other characters in the scene, if they had previously chosen the original over the transformed painting composition (Task 2b).

3.3. Experiment 2: dynamic evaluation of paintings

The aim of Experiment 2 was to investigate the psychological implications of the 'inner view' thesis and consisted of three blocks: (1) explicit measures of attention, importance and identification (first block); (2) implicit measures of visual attention (e.g., eye movements; first block); (3) ratings for preference, visual comfort and perceived sense of presence associated with each painting (second block) and (4) descriptions of the paintings provided by the participants (third block). For reason of space and conciseness, in this paper, we will only report and discuss results from the first block. For an account of the full study, we refer to Ruta (2019). In line with the 'inner view' thesis, we predicted that participants would report paying attention and identifying ChX more in the original compared to the transformed version of the paintings. Moreover, we collected ratings for importance to test the effectiveness of our manipulation, as in half of our selected stimuli ChX did not correspond to the most important character in the scene according to Christian iconography.

4. Materials and Methods

4.1. Participants

Fifty-eight participants, 24 identifying as female, gave informed consent to take part in the study; 36 of them were recruited using the Prolific platform (prolific.ac) and twenty22 from the Cardiff Metropolitan University student population to record their eye movements during the study. The average age of the participants was 29 years (SD = 5.7; range = 18–44 years). To take part in the online study, the participants had to match the same criteria in terms of native language (nationality of our sample population: UK citizens = 75%; United States citizens = 19.4%; Canadians = 5.5%), desktop size (25.8% of the participants had 15-inch screens, 15.5% had 17-inch screens and 58.6% had screens >17 inches; mobile phone and small tablet users were excluded) and position, as previously reported in Experiment 1. Half of the participants were assigned to the group that saw the original version of the painting (N = 31; females = 8) and the other half to the transformed composition (N = 31; females = 16).

The average amount paid to participants in our online sample for taking part in the research was £5/h, which was paid via Prolific or as a shopping voucher. The experiment was approved by the Ethics Committee of Cardiff Metropolitan

University and was conducted in accordance with the Declaration of Helsinki (2008). All the participants were naïve to the purpose of the study and received a debriefing after completing the experiment.

4.2. Procedure

As discussed before, we will report only results from the first block of Experiment 2. Participants completed an online survey developed using Qualtrics software (qualtrics.com) and were divided in two groups: one saw the original version of the paintings and the other that saw the transformed one. For each one, participants were asked to click directly on:

a) 'the part of the painting that attracted more of their attention' (attention task);
b) 'the part of the painting that they thought was the most important' (importance task) and
c) 'the character they mostly identified with' (identification task).

5. Analysis and Results

5.1. First block: Direct click on the image for attention, importance and identification

Following a similar approach to the one adopted in Experiment 1, for each participant, we calculated a *selections ratio* for attention, importance and identification tasks. In these tasks, the *selections ratio* represented the proportion of the number of times participants clicked on ChX (ChX_{freq}) divided by the times they did not, selecting any other parts of the paintings ($Other\ Ch_{freq}$). Based on the premises of the 'inner view' thesis, we did not expect to find a significant effect for the importance task, but we expected to find a significant difference between the two painting versions for the attention and identification tasks, with participants clicking more on ChX than on other parts of the scene. Consequently, we hypothesised that our character size manipulation would significantly influence which part of the painting participants explicitly reported attracted most their attention (attention task) and which character they felt they could most empathise with (identification task). However, we did not predict that our manipulation would have a significant effect for the importance task, as the semantic content in both two versions of the paintings was unchanged.

On average, the participants that saw the original version of the paintings reported higher *selections ratio* values compared to the ones that saw the transformed version for all three tasks: attention (original mean = 0.964; transformed mean = 0.676), identification (original mean = 1.19; transformed mean = 0.676) and importance (original mean = 1.15; transformed mean = 0.907). Three unpaired

two-samples Wilcoxon tests with Bonferroni correction for repeated tests indicated that this difference was statistically significant only for the attention task ($W = 602$, $p = 0.04$) but not for the identification ($W = 502$, $p = 0.328$) and importance ($W = 500$, $p = 0.392$) ones. Therefore, for the attention task, the proportion of clicks frequencies associated with ChX (*Selections Ratio*) reported by those who viewed the original version is significantly higher than that of those who viewed the transformed version.

6. Discussion

Medieval art has sometimes been regarded both by art history and visual science as 'primitive' (Carli, 1965), having little artistic value (Vasari, 1568; Millard, 1946) since it lacked the apparent mathematical logic and rigor that was supposedly later imparted to it by the adoption of linear perspective. The present work addressed the question of whether the figure size manipulations seen in many works of the period are the result of a far more sophisticated artistic imagination than is often recognised. In fact, the artistic manipulation of figure size for psychological effect was not unique to the Trecento Sienese and Florentine artistic tradition but is a recurring aspect of many other artworks across art history (Pepperell & Haertel, 2014; Mather, 2015; Crawford, 2018). To investigate this phenomenon empirically, we isolated and manipulated the R-Svc independently from the other complex set of properties that characterise medieval pictorial space, often called 'reverse perspective'. There have been many explanations put forward to interpret reverse perspective, but only two of them have directly addressed size manipulation: the predominant view in art history corresponding to the 'hierarchical scaling' thesis, which explains the use of the R-Svc as reflecting the hierarchical importance of certain figures compared to others, and the alternative 'inner view' thesis, which interprets size manipulation as a way of directing the viewers' visual attention and emotionally engaging them to empathise more with the point of view of ChX. In addition to these two theoretical positions, we advanced an original explanation: the visual imagination thesis. In line with the 'inner view' thesis, we suggested that the R-Svc could be explained by the link between imagined visual space and visual perception (Pepperell & Ruta, 2017; Ruta et al., 2017).

First, the results from Task 1 in Experiment 1 showed that our manipulation did not influence the perceived aesthetic value of the paintings themselves, reporting no significant difference in preference and liking ratings between the original and transformed versions of the paintings. These findings confirmed that the participants were not positively biased towards the original version, ensuring that the results from the other psychological variables we investigated were not biased by aesthetic preferences and could be interpreted as a direct effect of our size

manipulation. Moreover, they provide supporting evidence for the independence of the R-Svc from pictorial space.

Second, we investigated our original 'visual imagination' thesis, which considers artists' intentions and explains the use of the R-Svc through the link between key features of the phenomenology of visual perception (Ruta, Burleigh, & Pepperell, 2016) and imagined visual space. The results from Experiment 1 seem to suggest that the character size manipulation in the original composition was a better representation of imagined visual space compared to the transformed version. More precisely, when directly asked to identify themselves with the artist's intentions of wanting to represent the scene by while looking at one character in particular, the participants were significantly more likely to choose the original composition compared to the transformed one (Task 2a in Experiment 1) and to select imagining looking at ChX significantly more than the other characters in the scene, if they had previously chosen the original version (Task 2b in Experiment 1). These findings were in line with our predictions (Hypothesis A), clearly showing that the way Trecento Sienese and Florentine artists used the RSvc in their paintings (Original) was interpreted by the majority of participants as a better visual metaphor to represent 'directly looking at one character in particular'. A limitation of our research is that we cannot exclude the possibility that participants chose a version over the other as they perceived it as being more unusual and therefore more 'artistic'. Future research might aim to replicate our results by using different media (e.g., photography) or a simplified non-artistic version of the paintings reproducing key characteristics of medieval pictorial space (for exploratory study see Ruta, 2019). Another possibility would be comparing different task instructions, framing it only as a perceptual task.

According the 'inner view' thesis, we expected the participants' attention to be directed more towards ChX compared to other characters in the scene (Hypothesis B). In the first block of Experiment 2, we explicitly asked participants to click on the part of the painting (1) that attracted their *attention* more; (2) that they considered the most *important* and (3) that they *identified* with the most. The results from this block only partially confirmed our predictions: our size manipulation had a significant impact only on the explicit judgements concerning attention but not identification as we predicted. We suggest that for the identification task there might have been other confounding factors influencing participants' responses that are not directly related to characters' size, such as visual saliency, colour, contrast or other more socially relevant cues, like garments or facial features. Unfortunately, due to the nature of the data, it was not possible to carry out a more detailed analysis to compare the proportion of clicks associated with ChX against those associated with every other element chosen by the participants. Future research will aim to consider the identity of the other elements selected by

the participants and compare the frequency of clicks associated with each element in the painting to ChX.

It is relevant to mention that for Experiment 2 we also collected eye movements during a free viewing task (Experiment 2b), as an implicit measure of visual attention. Although discussing the results of this task goes beyond the scope of this paper, it is relevant to report that the results were in line with the 'inner view' thesis predictions, showing that the participants spent significantly more time fixating on ChX compared to other characters in the scene and that overall this proportion was significantly higher in the original version compared to the transformed one (Ruta, 2019). This seems to suggest that the R-Svc played a critical role in implicitly diverting the viewers' attention towards ChX. Taken together, these results provide strong evidence for the fact that the way Trecento Sienese and Florentine artists manipulated the relative sizes of characters in their original compositions significantly influenced – and still influences – viewers' visual attention on a more implicit and perhaps emotional level, in line with the hypothesis of the inner view thesis.

Further work in this area might seek to take advantage insights from Gestalt psychology. Although Rudolf Arnheim was firm in his views on the explanation of the figure size effects discussed here, elsewhere in his writings, we find suggestions that might offer other perhaps more deep psychological explanations. In his earlier monograph *Art and Visual Perception* (1956), Arnheim lays great stress on the role of forces and tensions in the organisation of pictorial space and in particular in how we interpret the relationship between figure and ground. Speaking of the dynamic nature of perception in relation to sculpture, he describes the way a figure can act not simply as a bounded area but as a dynamic object that is 'actively outgoing' as it looms forcefully towards the viewer in a convex space. This same looming effect of convex space is evident in some medieval and later painting where the artists give spatial prominence to a larger central figure and diminish the surrounding objects and figures, as with many of the examples shown here. It would be interesting to investigate whether there is a psychological connection between the empathetic thesis studied here and this dynamic perception thesis proposed by Arnheim.

7. Conclusions

Our findings tested the predominant 'hierarchical scaling' thesis (Doehlemann, 1910; Bunim, 1942; White, 1972), reporting supporting evidence in favour of the alternative 'inner view' thesis proposed by the art historian Wulff (1907). We showed that altering the R-Svc of the original compositions affected the way viewers engaged with the artworks, especially in terms of visual attention. We confirmed the critical role of figure scaling (R-Svc) in medieval art as a way

of attracting the viewer's attention towards the bigger character in the scene (ChX).

Drawing upon an emerging pattern from art history (Pepperell & Haertel, 2014; Ruta et al., August 2016) and evidence from scientific literature (Baldwin et al., 2016), the present study suggested an innovative way of integrating knowledge from art history and vision science, investigating the way participants interpreted the artist's intentions in medieval art, which has never been done before. Our results seemed to suggest that enlarged figure size (R-Svc) in medieval art was more likely to be painted from the artists' imagination and interpreted by viewers as a better way of representing 'imagining looking at' certain objects at specific spatial locations, rather than an optically correct way of representing the relative sizes of the same characters, in line with the 'visual imagination' thesis.

Finally, we reported the results of the first of a set of experiments investigating the 'visual imagination' thesis showing that when simply asked to select the part of the painting that attracted most their attention, participants were more likely to select ChX in the original condition compared to the transformed condition.

In line with Wulff's 'inner view' thesis, we conclude that the 'visual imagination' thesis has important implications explaining not only for the use of the R-Svc in medieval art but also more broadly for future studies aimed at testing the correspondence between the key characteristics of the phenomenology of human visual perception and imagined visual space through its artistic representations.

7.1. Open questions and directions for future research

The present study opens a new set of questions that future research should take into account to better explain the correspondence between visual imagination and visual perception: will enabling participants to freely manipulate pictorial space on a wider range of variables result in a better measure of imagined visual space? Does the use of the R-Svc result in a more efficient and natural perception of the scene portrayed (e.g., affecting reaction times for object or the identification of figures)? Do artistically manipulated images presenting the same R-Svc of Trecento Sienese and Florentine paintings have a positive effective connotation? And if so, does this affective connotation have an influence on visual attention and how these images are scanned by viewers? Future research should aim to integrate the insights gained from this study and develop a wider range of stimuli and practical tools to further investigate the visual imagination thesis.

Abstract
Prior to the discovery of linear perspective in the fifteenth century, European artists based their compositions more on imagination than the direct observation of nature. Medieval paintings, therefore, can be thought of as 'mental projections' of space rather than optical

projections, and were sometimes regarded as 'primitive' by historians as they lacked the spatial consistency of later works based on the rules of linear perspective.

There are noticeable differences in the way objects are depicted in paintings of the different periods. For example, human figures in pre-linear perspective works often vary greatly in size in ways that are not consistent with the laws of optics. Some art historians have attributed this to 'hierarchical scaling' in which larger figures have greater narrative significance. But there are examples of paintings that contradict this explanation.

In this paper we will consider an alternative to the hierarchical scaling hypothesis: that medieval artists used relative size to elicit empathy and to reflect the perceptual structure of imagination. This hypothesis was first proposed by the art historian Oskar Wulff, but has largely been dismissed since. We argue that artists of this period, far from being naïve, used sophisticated techniques for directing the attention of the viewer to a particular figure in a painting and encouraging them to 'see' the depicted space from that figure's point of view. We offer some experimental evidence in support of this hypothesis and suggest that the way artists have depicted space in paintings has an important bearing on how we imagine and perceive visual space.

Keywords: visual space, linear perspective, Medieval paintings.

References

Antonova, C. (2010). On the problem of "reverse perspective": definitions east and west. *Leonardo, 43*(5), 464–469.

Antonova, C., & Kemp, M. (2005). "Reverse Perspective": Historical Fallacies and an alternative view. In M. Emmer, *The Visual Mind II* (pp. 399–431). Cambridge, Massachusetts, United States: MIT Press.

Arnheim, R. (1956). *Art and Visual Perception,* London, England: Faber.

Arnheim, R. (1972). Inverted perspective in art: Display and expression. *Leonardo, 5*. 125–135.

Baldwin, J., Burleigh, A., Pepperell, R., & Ruta, N. (2016). The perceived size and shape of objects in peripheral vision. *i-Perception, 7*(4). doi: 10.1177/2041669516661900

Bunim, M. S. (1942). *Space in medieval painting and the forerunners of perspective*. New York: Columbia University Press.

Carli, E. (1965). *Italian Primitives: Panel Painting of the Twelfth and Thirteenth Centuries.* New York, NY: H. N. Abrams.

Crawford, C. (August-September 2018). Understanding metaphor in art: Distinguishing literal giants from metaphorical challenges. In: *International Association of Empirical Aesthetics Congress.* Toronto, Canada.

Doehlemann, K. (1910). *Zur Frage der sog. 'umgekehrte Perspektive.* Berlin, Germany: Repertorium fur Kunstwissenschaft.

Hockney, D. (2001). *Secret Knowledge: Rediscovering the Lost Techniques of the Old Masters.* London, England: Thames and Hudson.

Lubbock, J. (2006). *Storytelling in Christian art from Giotto to Donatello*. New Haven & London: Yale University Press.

Mather, G. (22–23 August 2015). The depiction of visual space in Canaletto's Venetian vedute. In *Proceedings of the Visual Science of Art Conference.* Liverpool, UK.

Meiss, M. (1946). Italian Primitives at Konopiště. *The Art Bulletin*, *28*(1), 1–16.

Millard, M. (1946). Italian primitives at Konopiště. *The Art Bulletin, 28*(1), 1–16.

Pepperell, R., & Haertel, M. (2014). Do artists use linear perspective to depict visual space? *Perception, 43*(5), 395–416.

Pepperell, R., & Ruta, N. (August 2017). Image and imagination: How figure scale in medieval painting reflects visual perception. In *The Visual of Science of Art Conference (VSAC),* Berlin.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer, R. (2017). Package 'nlme'. *Linear and nonlinear mixed effects models*, version, *3*(1).

RStudio Team. (2015). RStudio: Integrated Development Environment for R. Boston, MA. Retrieved from: http://www.rstudio.com/

Ruta, N. (2019) Visual perception in far peripheral visual space and its artistic representations. [Unpublished doctoral dissertation]. Cardiff Metropolitan University. https://ethos.bl.uk/OrderDetails.do?did=1&uin=uk.bl.ethos.802055

Ruta, N., Burleigh, A., & Pepperell, R. (August 2016). Why do artists depict hands smaller than their actual size? In: *The Visual Science of Art Conference VSAC.* Barcelona, Spain.

Ruta, N., Burleigh, A., & Pepperell, R. (December 2017). Why did Medieval painters manipulate figure size? A perceptual psychology approach to art history. In: *Applied Vison Association (AVA) Christmas Meeting.* London.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*(13), doi: 10.1167/11.5.13.

Todorovic, D. (2008). Is pictorial perception robust? The effect of the observer vantage point on the perceived depth structure of linear-perspective images. *Perception, 37*, 106–125.

Vasari, G. (1568). *The Lives of the Most Excellent Painters, Sculptors, and Architects.* Italy: Giunti.

White, J. (1972). *The Birth and Rebirth of Pictorial Space.* London, England: Faber and Faber.

Williams, J. R. (2008). The Declaration of Helsinki and public health. *Bulletin of the World Health Organization*, *86*, 650–652.

Wulff, O. (1907). Die umgekehrte Perspektive und die Niedersicht. Eine Raumanschauungsform der altbyzantischen Kunst und ihre Fortbildung in der Reinassance". In H. Weizsäcker, *Kunstwissenschaftliche Kunstwissenschaftliche Beiträge August Schmarsow gewidmet zum fünfzigsten Semester seiner akademischen Lehrtätigkeit* (pp. 3–42). Leipzig, Germany: K.W. Hiersemann.

Nicole Ruta works as postdoctoral research fellow in the Laboratory of Experimental Psychology at KU Leuven University. She is dedicated to multidisciplinary and interdisciplinary research, working with experts from different disciplines including fine art, computer engineering, graphic coding, neurobiology, and theology. Her PhD thesis studied visual perception in far peripheral visual space and its artistic representations. Her research interests involve aesthetic experience, art history, space perception, affective connotations of visual properties and their applications in architecture and design.
E-mail: nicole.ruta@kuleuven.be,
Tel: +32 16 32 25 72
Orcid: 0000-0001-5300-2913

Robert Pepperell PhD is Professor of Fine Art at Cardiff School of Art and leader of the Fovolab at Cardiff Metropolitan University. He investigates the nature of the conscious mind and visual perception through painting and drawing, scientific experimentation, and philosophical inquiry. As an artist he has exhibited at Ars Electronica, the Barbican Gallery, Glasgow Gallery of Modern Art, the ICA, and the Millennium Dome. As an academic he has published several books, including The Posthuman Condition (1995 and 2003) and The Postdigital Membrane with Michael Punt, 2000), as well as articles, reviews, and papers in the fields of art history, philosophy of mind, artificial intelligence, neuroscience, and perceptual psychology.
Orcid: 0000-0002-4728-4122

Alistair Burleigh Born in Bristol in 1980, Alistair studied Fine Art and Contemporary Multimedia at The University of Wales, Newport. In the early 2000s he developed ground breaking 360-degree digital display techniques in the field of projection mapping, and between 2005 and 2012 led the creative digital mapping studio Wrap3. He has produced complex immersive digital environments for a wide range of functions and clients on a global basis, including full-scale commercial projects for Chanel, the James Bond production 'Quantum of Solace', the United Arab Emirates pavilion at World Expo 2010 Shanghai, the Arcadia Retail Group, GQ, On:Off London Fashion Week, and many art exhibitions and fairs such as The Kinetica Art Fair in London. Alistair is now a member of Robert Pepperell's research group at Cardiff School of Art & Design and is currently Technical Director of Fovotec Ltd.
Orcid: 0000-0001-6226-4763