Computers & Education: X Reality 2 (2023) 100015

Contents lists available at ScienceDirect

Computers & Education: X Reality

journal homepage: www.journals.elsevier.com/computers-and-education-x-reality

Shape distinctness and segmentation benefit learning from realistic visualizations, while dimensionality and perspective play a minor role

Alexander Skulmowski

Digital Education, Institute for Informatics and Digital Education, Karlsruhe, Karlsruhe University of Education, Germany

ARTICLE INFO

Keywords:
Visual design
Learning
Visualization
Cognitive load
Realism

ABSTRACT

Research on the design of visualizations has revealed that realistic visualizations can be beneficial for retention and transfer performance despite their tendency towards causing cognitive load. However, it still can be hard to predict whether a more detailed visualization will be more effective for learning than a more abstract one. The purpose of the reported studies is to investigate whether an emphasis on the specific benefits of realism can enhance learning from realistic visualizations. In three experiments, the dimensionality (i.e., whether flat cross-sections or shapes conveying depth and space are presented), perspective, shape distinctness, and segmentation by color are investigated. In Experiment 1, a visualization of a flat section of the kidney elicited a greater cognitive load if this section was presented as a realistic rendering rather than a schematic drawing, while a more dimensional model of the kidney appears to be harder to cognitively process if it is presented as a drawing. Experiment 2 examined whether a perspective emphasizing depth is more compatible with a realistic rather than a schematic visualization, but this effect failed to reach significance. Lastly, Experiment 3 demonstrated that shape distinctness and segmentation using colors enhance retention performance. The main result of this series of studies is that even minor changes in the presentation of realistic visualizations can have an impact on cognitive load. In conclusion, in order to optimize learning, realistic visualizations should make use of depth, emphasize semantic information by increasing the distinctness of shapes, and use segmentation using colors or different surface materials.

1. Introduction

For decades, differences in the cognitive processing of detailed, realistic and abstract, schematic visualizations have been an important research area (Belenky & Schalk, 2014; Goldstone & Son, 2005), with a growing interest in the last years (Skulmowski et al., 2022). While there are some positions arguing against the use of realism (e.g., Smallman & St. John, 2005), many authors acknowledge that realism and abstraction can have specific advantages for certain objectives. For instance, it has been claimed that detailed visualizations can be memorized more easily due to their clearer representation of real-world objects (Goldstone & Son, 2005), while abstract visualizations are often described to be easier to transfer to different content domains (Goldstone & Son, 2005; Menendez et al., 2020). However, these intuitively plausible guidelines are not always supported by the empirical evidence (e.g., Scheiter et al., 2009; Skulmowski, 2022b), thus raising the question if more specific properties of visualizations interact with realism that have so far been overlooked. Investigations into the effects of visual properties of static visualizations can inform the design of educational X reality (XR) applications and are thus a valuable starting point for research in this field, in particular for desktop virtual reality applications.

The three experiments presented in this paper investigate the impact of visual design factors of realistic visualizations on learning. The research question empirically tested in Experiment 1 is whether the dimensionality (or depth) of models presented in visualizations affects how well learners are able to process schematic and realistic renderings. Similarly, Experiment 2 was conducted to assess the influence of perspective on learning with schematic and realistic visualizations. In these two experiments, it was assumed that more dimensionality or a perspective emphasizing dimensionality, respectively, promote learning with realistic visualizations. The research questions assessed in Experiment 3 concern the role of distinct shapes and segmentation using color coding to improve the cognitive processing of realistic visualizations. Before taking a closer look at the design properties mentioned in this overview, our concept of realism will require a more precise definition. After an introduction of the aspect of realism, more specific design aspects of (realistic) visualizations and their effects on learning will be discussed.

E-mail address: alexander.skulmowski@ph-karlsruhe.de.

https://doi.org/10.1016/j.cexr.2023.100015
Received 9 December 2022; Received in revised form 6 March 2023; Accepted 16 March 2023

2. Literature review

2.1. Realism as a form of perceptual richness

On the technical level, realistic visualizations contain several components. The objects in a virtual scene are usually produced using geometry consisting of a number of polygons that are arranged in a particular shape. Shaders are then applied to the model(s) that determine how shiny, transparent, and bumpy the surface appears. The virtual scene is populated with lights and a "camera" that determine the look of the rendered output. These three components—geometry, shading, and rendering—constitute computer-generated visualizations in the GSR model (Skulmowski et al., 2022) and can be varied to achieve different levels of realism.

From a perceptual standpoint, a higher level of realism usually requires that more details are perceived and processed, resulting in a higher load on the perceptual system (Lavie et al., 2009). Such an abundance of visual information has been referred to as *perceptual richness* (Menendez et al., 2020, 2022), and, during a learning task, may impose demands on learners, for example by requiring them to mentally separate the relevant from the irrelevant information (e.g., Berney et al., 2015). Realistic visualizations featuring many details will need to be mentally segmented into more primitive three-dimensional primitives (as postulated within *geon* theory, Biederman, 1985, 1987; see Skulmowski et al., 2022; Skulmowski, 2022a). This process involving varying levels of perceptual load is thought to result in concrete mental representations that exert a certain cognitive load depending on the number of elements to be kept in mind at the same time (Skulmowski et al., 2022). The relationship between realism and cognitive load is discussed in the following section.

2.2. Cognitive load and realism

Research on realism is often based on cognitive load theory (Sweller et al., 1998, 2019) which assumes a limited working memory capacity that can be filled with relevant (*intrinsic*) and irrelevant (*extraneous*) cognitive load. In this model, learning can be enhanced by minimizing extraneous cognitive load so that mental capacities can be used to engage with intrinsic cognitive load. Therefore, the question arises whether realism needs to be considered to be a form of extraneous cognitive load (see Scheiter et al., 2009).

Realism has been shown to induce extraneous cognitive load in multiple studies (e.g., Skulmowski, 2022a; 2023b). However, research has shown that realism may be necessary for some learning tasks (for overviews, see Nebel et al., 2020; Skulmowski et al., 2022). Beyond task requirements, recent research has investigated the role of more specific properties of realism and their effects on learning. There are intrinsic properties of visualizations that may affect if and how much learners can benefit from realism. There has been a discussion regarding the cognitive costs and benefits that different types of presentation have, leading some authors to speak of cost-benefit models in which some cognitive load may be tolerated in case this obstacle provides other advantages (Skulmowski & Xu, 2022). Thus, it may be necessary to more closely analyze under which specific circumstances realism can benefit learners. In the following section, the design factors of realism investigated in the present studies are discussed.

2.3. Dimensionality and perspective as means to utilize realism

One of the major differences that realistic visualizations have compared to more simplified (or schematic) diagrams is that the former can be used to visually convey the three-dimensional shape of an object in a more detailed manner. More specifically, realistic visualizations typically offer a sense of *dimensionality*. Krüger et al. (2022) define dimensionality as a result of providing depth cues (they cite Wu & Shah, 2004) and as a help in creating a spatially correct mental model (Krüger et al. cite Chen et al., 2015). Accordingly, Krüger et al. (2022) found a positive effect on learning a spatial structure in 3D rather than as a two-dimensional presentation. However, it needs to be noted that the 3D version in this study could be rotated as an element in augmented reality. Other studies in which simpler structures or abstract, verbal information were used did not result in an advantage of dimensionality (He et al., 2022; Keller et al., 2006), suggesting that this quality will mainly benefit spatially complex knowledge. Thus, detailed anatomical models were selected as the learning materials for the present studies.

Since it has been found that learning about the process of mitosis does not benefit from more realistic visualizations (Scheiter et al., 2009), it has been argued that the effects of realism will be stronger in tasks in which learning about the shapes of objects is the main objective (Skulmowski et al., 2022). However, while learning about a process using detailed microscopic photographs resembling cross-sections of cells did not promote learning, learning about the shapes of objects presented as dimensional entities rather than cross-sections consistently benefits from realistic details, at least for specific variables (e.g., Skulmowski, 2022a, 2022b; Skulmowski & Rey, 2021). Therefore, it might be possible that realism primarily has a positive effect if the dimensionality (i.e., whether an object is presented as a cross-section or in its complete form) of a visualization is high. Given the previous findings, emphasizing the dimensionality of a visualization should enhance performance when learning with realistic visualizations.

Beyond the dimensionality of a visualization, the perspective presents another potential possibility to enhance learning from realistic visualizations. Huk et al. (2010, Exp. 1) presented a study in which the two factors format (2D vs. 3D) and cues (with vs. without) were investigated concerning their effects on learning a biological process with an animation. While they did not find an effect of the format on retention, the 3D animation significantly raised comprehension scores compared to the 2D version and increased participants' positive attitude towards the animation. The animation used in the experiment conveyed the processes that a protein is involved in. It is important to acknowledge that the 2D version of the animation not only featured a different perspective compared to the 3D version (an orthographic side view vs. a perspective view), but also a different visual style (solid-filled line diagrams vs. 3D-rendered objects featuring different shaders). Thus, we do not know whether the results can be directly traced back to the perspective or the visual style. In fact, there may be a complex relationship between these two factors. As a result, perspective is investigated as another design factor that may enhance learning with realistic visualizations.

2.4. Shape distinctness and segmentation as semantic features

While the dimensionality and perspective used in visualizations may be considered as more subtle design factors, Experiment 3 investigates how two design aspects that are able to communicate greater semantic information can help learners to optimize working with realistic visualizations. Distantly related to the previously discussed aspect of dimensionality is the *shape distinctness* of the geometry. A 3D object can be modeled in a way that the underlying geometry is rather amorphous, with less emphasis on characteristic shapes and without clear indications of component boundaries. The effect of this aspect on learning were assessed by Skulmowski (2022, Exp. 2 & 3). In the two studies, fictional bone structures were to be learned. There were no main effects of shape distinctness in these studies. However, there was an interaction effect indicating that learners' cognitive load can be lowered using a high shape distinctness when using realistic visualizations, while the opposite effect was found for schematic visualizations (Skulmowski, 2022b). Thus, the effects of shape distinctness should be investigated further in order to determine whether it can have a positive influence on learning with realistic visualizations.

Lastly, an important aspect of realism stems from the ways in which different virtual materials and color cues can affect learning. Potential uses for color cues include the highlighting of important information, the segmentation of visual content, and the coding of different information in order to establish connections between them (Pett & Wilson, 1996). In the context of realism research, Skulmowski and Rey (2018) conducted a study using a realistic computer-generated model of a fictional bone in which color coding increased retention scores if the shading component included a high number of (arbitrary) details in the form of small bumps and ridges on the surface. However, color coding lowered retention performance for a smooth version of the model without such details, leading to an interaction effect between color coding and the level of detail (Skulmowski & Rey, 2018). An explanation for this result came from color coding lowering cognitive load in that study. Using a similar method, Skulmowski (2022c) found that color coding is particularly effective for transfer performance if visualizations contain color coding both in the learning and testing phases, but that color coding can also be detrimental if color cues are only presented during the learning phase. In these last two studies just summarized, color cues were primarily used to facilitate the visual segmentation of an otherwise complex object without clear boundaries between the different components.

3. The present studies

The first two experiments were conducted to investigate whether an alignment between the level of realism and the visual design factors of dimensionality and perspective leads to a higher retention score and a lower cognitive load. This alignment hypothesis is based on the idea that the cognitive costs of realistic details are only worth risking if a specific advantage is to be expected (see Skulmowski & Xu, 2022). In Experiment 1, realism should be particularly helpful when learning using visualizations featuring objects that make use of realistic shading and rendering through their dimensionality. Experiment 2 was conducted to assess a similar relationship between perspective and realism.

In addition, the third experiment examined whether two components of realistic visualizations, namely geometry and shading, will lead to advantages during learning. The dimension of geometry was manipulated by varying the shape distinctness and the effects of shading were tested by defining the boundaries of object components using color coding. As a result the first two experiments deal with more subtle visual design choices, while the third experiment is focused on the effects of including more semantic cues using the geometry and shading of the depicted object.

4. Experiment 1

Experiment 1 investigated whether the dimensionality of objects depicted in visualizations has an effect on cognitive load and retention performance. It was hypothesized that an alignment that builds on the specific strengths of realistic and schematic visualizations should lead to the lowest cognitive load and highest retention scores. For this particular study, an interaction effect between the factors realism and dimensionality was assumed in which dimensionality enhances learning with a realistic, but not with a schematic visualization. On the flipside, a flat cross-section may be easier to process and memorize with a schematic visualization style.

4.1. Method

4.1.1. Participants and design

The sample size was calculated based on previously found effects (η_p^2) of realism that ranged between 0.09 (Skulmowski, 2022b) and 0.14 (Skulmowski & Rey, 2021). The simultaneous manipulation of realism and perspective in the study by Huk et al. (2010, Exp. 1) resulted in smaller effects, with $\eta_p^2 = 0.09$ for comprehension. Thus, a conservative estimate of $\eta_p^2 = 0.09$ was chosen as the basis for the sample calculation, which resulted in a target sample size of 82 using G*Power (Version 3.1.9.2; Faul et al., 2009) with a power of 0.80. The design features the between-subjects factors of realism (schematic vs. realistic) and dimensionality (flat vs. dimensional).

The participants (72 female, 10 male) were aged between 18 and 30 years and were German native speakers enrolled in teacher training courses at a university of education in Germany. They had little or no prior knowledge of the anatomy of the kidney. All between-subject groups contained 21 participants except for the group in which participants viewed the schematic flat cross-section ($n = 19$ participants). Group assignment was performed using block randomization. Additional participants who did not fulfill the participation criteria (age of 18–30, German as native language, and little or no knowledge of the topic) or who indicated that they had been strongly distracted by noise or technical difficulties were not included in the analyses.

The three experiments presented in this paper did not require an ethics approval due to the local legislation. All relevant national and institutional guidelines that apply to this type of study were followed.

4.1.2. Materials

During the learning phase, participants learned with one of the four versions of the kidney anatomy visualization presented in the two top rows of Fig. 1. The top row shows the schematic version of the visualization and the middle row contains the realistic version. The visualizations in the left column display the flat cross-section of the kidney, while the right column shows the dimensional version. In the logic of the GSR model, the factor of dimensionality primarily affects the geometry dimension, as the higher level of dimensionality is achieved by the use of models that emphasize the 3D structure of the kidney. The factor of realism was varied by manipulating the shading and rendering dimensions. While the realistic visualizations use photo-realistic lighting models and detailed shaders, the schematic visualization features cartoon-like shading and was rendered with contour lines. In the bottom row, the two test visualizations can be seen. These were created in a more realistic, but less detailed style intended to offer an “in-between” level of realism that does not bias responses (see Skulmowski, 2022a). The kidney was chosen as the learning material because it can be presented as a cross-section as well as in a more dimensional style. The two test pages featured one of the two visualizations at the top and drop-down menus containing all the items that were learned. For each correctly assigned label, one point was awarded with a maximum of 11 points per test, resulting in a maximum total score of 22 for both tests combined. Wrong responses did not result in any form of penalty. A script was used for scoring in all three studies. The kidney visualizations are based on the materials used for Experiment 2 of Skulmowski (2022a). All stimuli featured in this paper were created using Blender (https://www.blender.org).

Cognitive load was measured using three question items that were developed by Klepsch et al. (2017) to capture extraneous load (for an overview on the selection of appropriate cognitive load in perceptually demanding tasks, see Skulmowski, 2023a). The questions, which ask participants concerning their difficulty in completing a learning task, were presented in a minimally adapted version devised by Skulmowski and Rey (2020) and focus on the visualizations instead of the entire learning task. The extraneous load items had a reliability of McDonald’s (1999) $\omega = 0.78$. The retention test had a reliability of $\omega = 0.85$.

4.1.3. Procedure

The procedure followed that of previous studies (e.g., Skulmowski & Rey, 2021) and the experiment was conducted in a 10-seat laboratory. Participants provided informed consent and entered personal details to check the participation criteria (aged between 18 and 30 years, native German speakers, little or no knowledge regarding the topic). Next, participants received the instructions for the learning task, stating that their objective would be to learn the names, shapes, and locations of the parts of the kidney within 35 s. The following page presented the learning phase including one of the visualizations from Fig. 1 and a countdown timer. Participants were forwarded to a page featuring the three cognitive load question items, after which a 60 s sorting filler task

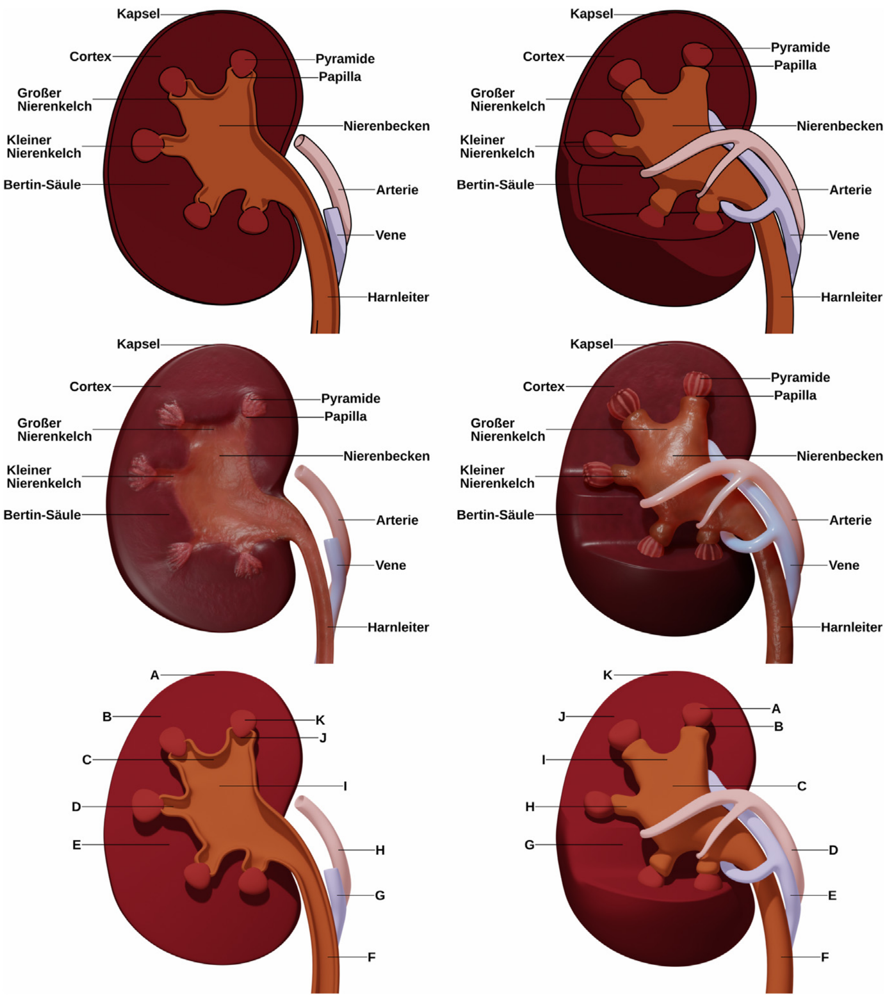

Fig. 1. The learning materials used in Experiment 1 based on Fig. 4 in Skulmowski (2022a, p. 10320) and additional reference materials. The top row shows the kidney in a schematic style, while the middle row presents the realistic versions. The left column displays the flat cross-section, and the dimensional model is shown on the right column. The bottom row contains the two retention tests. Adapted from "Is There an Optimum of Realism in Computer-Generated Instructional Visualizations?," by A. Skulmowski, 2022, *Education and Information Technologies*, *27*, p. 10320 (https://doi.org/10.1007/s10639-022-11043-2). © 2022 Skulmowski (licensed under the Creative Commons Attribution License, http://creativecommons.org/licenses/by/4.0).

was presented. Then, the two retention tests (see Fig. 1, bottom row) were shown on individual pages featuring one of the two test visualizations and drop-down menus. The tests did not impose a time limit on participants. The order of the two test pages was counterbalanced. After the tests, participants were asked to indicate their gender and course of study. They also responded to the two quality control questions assessing whether there were strong distractions or technical difficulties (taken from Skulmowski & Rey, 2020). On the last page, participants were thanked and provided with additional information concerning the experiment. All experiments used SoSci Survey (Leiner, 2021) for the data collection.

4.2. Results

The analyses for all three experiments in this paper were planned as 2 × 2 analyses of variance (ANOVAs) and were conducted using R (Version 4.0.5; R Core Team, 2021) and JASP (Version 0.16.1; JASP Team, 2022). Plots were generated using ggplot2 (Version 3.3.5; Wickham, 2016). The assumptions of variance homogeneity and the normality of the model residuals were tested with Levene's test and the Shapiro-Wilk test. In case any of the two tests reached significance, a nonparametric ANOVA using aligned rank transformations (Wobbrock et al., 2011) was used instead.

4.2.1. Cognitive load

Using a nonparametric ANOVA, the hypothesized interaction effect between realism and dimensionality was confirmed, $F(1, 78) = 5.35$, $p = .023$, $\eta_p^2 = 0.06$. As expected, the cognitive load ratings were highest for the two misaligned conditions that did not make use of the advantages of realism or a simplified schematic style, respectively (see Fig. 2A).

4.2.2. Retention

A nonparametric ANOVA did not result in the hypothesized interaction effect ($p = .369$). However, the descriptive data follow the expected pattern of higher retention scores for the two aligned groups (see Fig. 2B). It needs to be noted that the effect, even if it had reached significance, would likely have been very small ($\eta_p^2 = 0.01$).

5. Experiment 2

Experiment 1 resulted in a significant interaction between the factors realism and dimensionality in which a higher level of realism was beneficial for the more dimensional visualization in terms of cognitive load, while the schematic style resulted in a lower level of cognitive load if combined with the flat cross-section. The retention scores show an inverted version of this pattern, but this effect did not reach significance. Despite this tendency in favor of the alignment hypothesis (i.e., that the more dimensional presentation is favorable for realistic renderings and vice versa), the differences between the visualizations may not have been large enough. Although the visualizations strongly differed on a superficial level concerning the look of the depicted objects, they did not differ on the structural level, as the overall configuration of the parts of the object remained essentially the same across the four visualizations compared in the experiment.

Another aspect related to dimensionality is the perspective used to

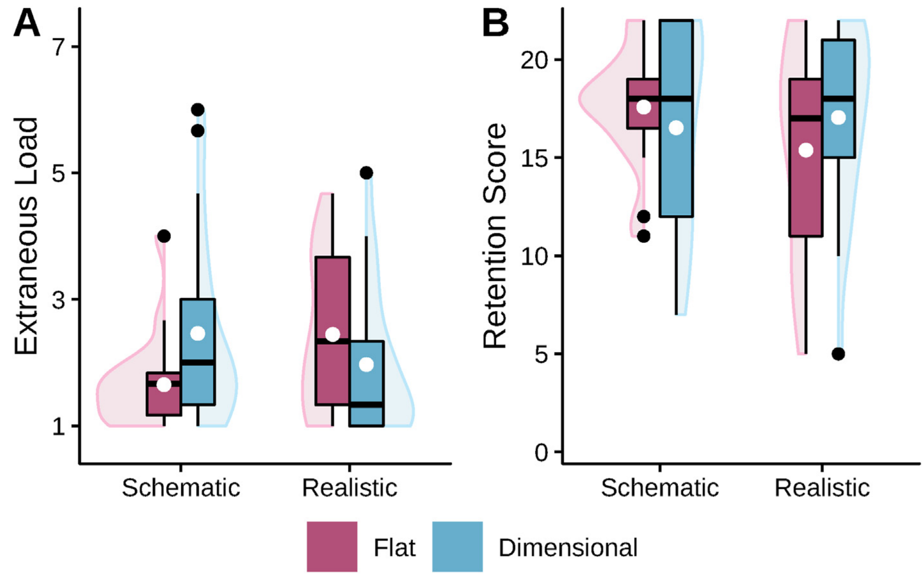

Fig. 2. Boxplots with violin elements of the data from Experiment 1. Fig. 2A: The averaged extraneous cognitive load data. Fig. 2B: The retention scores. White dots indicate the means of groups.

present an object in a visualization. In Huk et al.'s (2010) experiment, the format of visualization was investigated by comparing the perspective and the level of realism at the same time. However, it may be possible that these two aspects exert a divergent impact on visualizations, similar to the tendency found in Experiment 1. While a visualization utilizing an oblique perspective may be easier to process and memorize in a realistic rather than a schematic style as the former rendering mode provides depth cues in the form of shadows, a flat side view in which depth cues are not of concern might be faster to grasp and to learn when presented in a schematic style. Thus, an interaction effect between the factors realism and perspective was hypothesized.

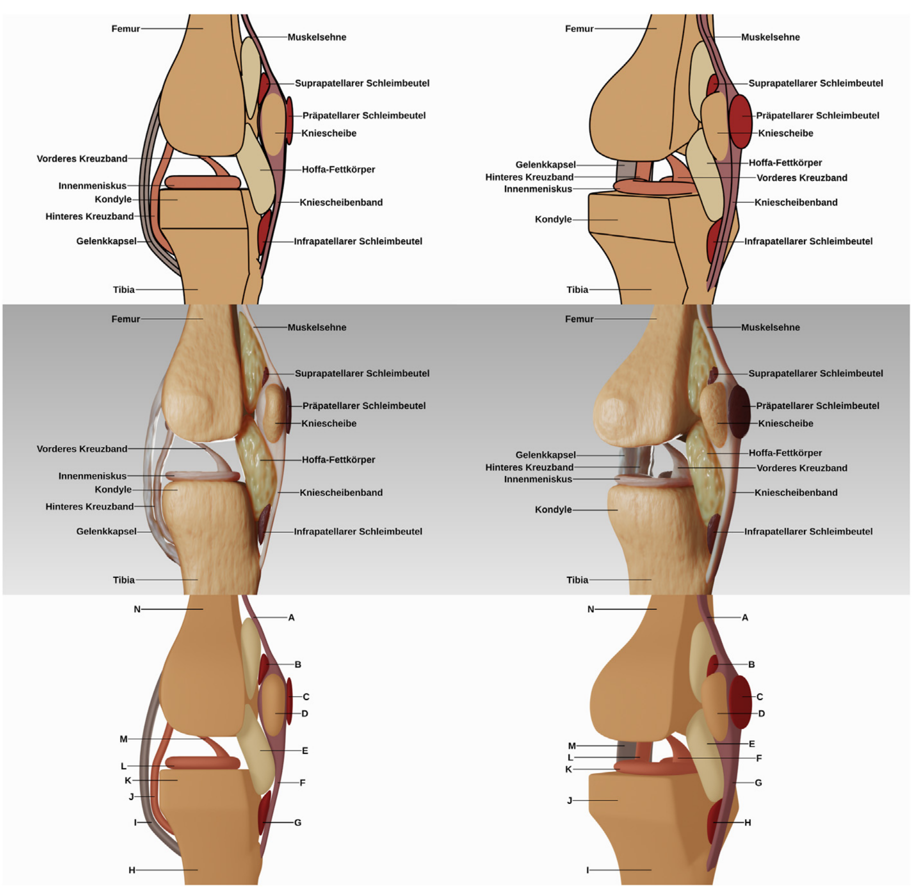

Fig. 3. The learning materials used in Experiment 2 based on Fig. 1 in Skulmowski and Rey (2020, p. 254), with additional information and reference gathered from Tillmann (2016) and Bammes (2009). The top row presents the knee joint as a schematic rendering, and the middle row shows the realistic style. On the left column, the orthographic side view is displayed, while the perspective view can be seen on the right column. The two retention tests are presented in the bottom row. Adapted from "The Realism Paradox: Realism Can Act as a Form of Signaling Despite Being Associated With Cognitive Load," by A. Skulmowski and G. D. Rey, 2020, *Human Behavior and Emerging Technologies*, *2*, p. 254 (https://doi.org/10.1002/hbe2.190). © 2020 Skulmowski and Rey (licensed under the Creative Commons Attribution License, http://creativecommons.org/licenses/by/4.0).

5.1. Method

5.1.1. Participants and design

After finding a small effect of the design factors in Experiment 1, stronger visual differences between the visualizations were decided upon for Experiment 2 in order to increase potential effects. Furthermore, the expected effect size was reduced to $\eta_p^2 = 0.06$, resulting in a target sample size of 125 participants for the 2 × 2 between-subjects design with the factors realism (schematic vs. realistic) and perspective (orthographic side view vs. dimensional perspective view) as computed with G*Power (Version 3.1.9.2).

The participants were 107 female and 18 male teacher students with the same participation criteria as in Experiment 1. All groups contained 31 participants except the group in which the realistic orthographic side view was presented (n = 32). Again, participants not meeting the participation criteria or who encountered difficulties were not included in any of the analyses.

5.1.2. Materials

The stimuli used as learning targets in Experiment 2 consisted of four versions of the knee joint anatomy visualizations that are shown in the two top rows of Fig. 3. The schematic version is displayed in the top row and the realistic version can be seen in the middle row. The left column presents the orthographic flat side view, while the perspective view is shown in the right column. According to the GSR model, the factor of realism was implemented similarly as in Experiment 1. However, the schematic version did not feature any kind of shading, as shading has been found to be a critical component of realism (Höst et al., 2022). The two test visualizations are included in the bottom row. Participants could achieve 14 points per test, resulting in a maximum total score of 28 points. The knee visualizations are based on the materials developed by Skulmowski and Rey (2020) with additional reference material.

Experiment 2 also used the extraneous cognitive load items adapted from Klepsch et al. (2017) as discussed in Experiment 1. The extraneous load items had a reliability of McDonald's ω = 0.87. The retention test had a reliability of ω = 0.85.

5.1.3. Procedure

The procedure was identical to Experiment 1, except that the learning phase lasted 60 s. The study was conducted in a PC laboratory.

5.2. Results

5.2.1. Cognitive load

The hypothesized interaction effect between realism and perspective did not reach significance with a nonparametric ANOVA (p = .708). Overall, the cognitive load ratings were rather low, resulting in a floor effect (see Fig. 4A).

5.2.2. Retention

A nonparametric ANOVA did not result in a significant interaction effect between the two factors (p = .579). Although the data follows the expected pattern (see Fig. 4B), this effect would have been of an even smaller magnitude than the related effect in Experiment 1.

6. Experiment 3

Despite the strong differences in the visualizations of Experiment 2, no significant interactions or differences were found. However, the retention data follows the pattern of the alignment hypothesis on the descriptive level. Nevertheless, the first two studies presented in this paper suggest that the dimensionality and perspective of visualizations have little impact on learning and cannot be used to reliably predict whether realism will have a positive or negative effect. Both studies have in common that the differences in the visualizations that were compared are of a primarily graphical nature that does not necessarily affect how the information is processed at a semantic level. In other words, superficial changes in shape and perspective do not appear to have a strong influence on learning despite the considerable visual differences they can cause. An experiment was conducted to assess whether visual features that communicate semantic knowledge are a superior way of improving learning with realistic visualizations. It was expected that a higher shape distinctness and color coding (or rather, color segmentation) will lead to higher learning outcomes. In addition, it was hypothesized that learning with nondistinct shapes will be facilitated using color coding in a compensatory effect (see Skulmowski & Rey, 2018), leading to a hypothesized interaction effect. Importantly, inverted results were hypothesized for cognitive load responses.

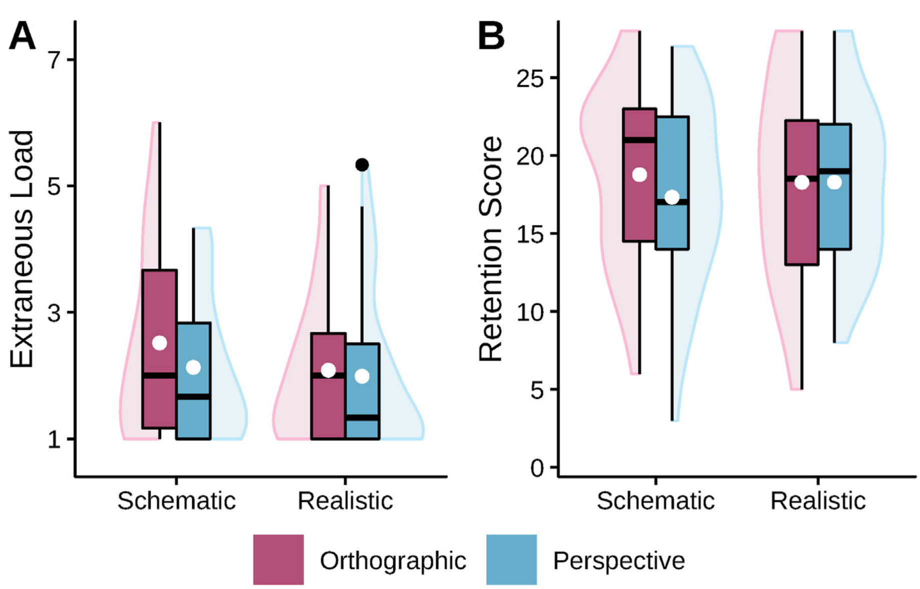

Fig. 4. Boxplots with violin elements of the data from Experiment 2. Fig. 4A: The averaged extraneous cognitive load data. Fig. 4B: The retention scores. White dots indicate the means of groups.

6.1. Method

6.1.1. Participants and design

The original study (Skulmowski, 2022b) assessing the effects of shape distinctness from which the materials for this study were adapted had 125 participants and did not find a significant main effect, but an interaction effect regarding cognitive load ratings of $\eta_p^2 = 0.08$. In addition to a number of improvements such as an optimized test and a more promising second factor of color coding (with effect sizes of $\eta_p^2 = 0.06$ in the study of Skulmowski & Rey, 2018, and $\eta_p^2 = 0.07$ in Skulmowski, 2022c), the target sample size was slightly increased to 152 to match an effect size of $\eta_p^2 = 0.05$ (power = 0.80). In order for a dataset to be included in the analyses, the responses needed to conform the participation conditions and data quality control measures as outlined in Experiment 1 (except that no question concerning prior knowledge was necessary due to the learning content being fictional).

As planned, data collection was stopped after 152 participants had completed the study.[1] Other data sets in which distractions or problems were indicated were excluded from the sample as in the two previous studies. Of the 152 participants, 128 were female and 24 were male. Participants were assigned to the four cells of the 2 × 2 between-subjects design with the factors shape distinctness (low vs. high) and color coding (without vs. with) using block randomization. There were 41 participants who learned with the visualization featuring a low shape distinctness and without color coding, 38 participants who used the version with a high shape distinctness and without color coding, 37 who were presented with a model with a low shape distinctness and with color coding, and 36 utilized a visualization with a high shape distinctness and with color coding.

6.1.2. Materials

The study reused renderings used by Skulmowski (2022b), and additional variations were created based the original source files. The method followed the approach presented by Skulmowski and Rey (2018) in which a fictional bone model is generated that can be shaped according to the requirements that a study may have. The model used by Skulmowski (2022b) was produced by applying an arbitrary texture as a deformer that shapes a spherical base mesh. The result is an organically-shaped object. Thus, the shape distinctness depends on the strength of this deformation and can be varied continuously. The renderings featuring a low shape distinctness (see the two top images in the left column of Fig. 5) show a rather blob-like appearance, while the model with a high shape distinctness (see the two top images in the right column of Fig. 5) displays easily recognizable shapes that resemble a bone structure. In order to assess whether a segmentation using color coding further helps learners, renderings without color (see the top row of Fig. 5) were compared with visualizations featuring color cues (see the middle row of Fig. 5). As a result, this study is focused on the geometry and shading dimensions, rather than the rendering dimension investigated in the previous experiments.

The learning tests were comprised of a retention test (see the bottom-left image in Fig. 5, $\omega = 0.86$) and a transfer test (see the bottom-right image in Fig. 5, $\omega = 0.87$), both with a maximum score of 15 points. Both of these models had a "medium" level of shape distinctness generated by setting the shape deformer of the base model to a level in-between the low distinctness and high distinctness versions used during the learning phase. In addition, the transfer model received further deformations requiring participants to apply their knowledge to this somewhat unfamiliar shape. In these tests, three components were labeled, but were not previously learned, and thus needed to be assigned the "NOT LEARNED" label. Again, the study used the same extraneous cognitive load items as the rest of the studies ($\omega = 0.90$).

6.1.3. Procedure

The procedure was identical to the previous experiments, except that the learning phase lasted 90 s and there was an additional transfer task on a separate page displayed after the retention test in a fixed order. This experiment was conducted online.

6.2. Results

6.2.1. Extraneous load

The cognitive load data (see Fig. 6A) were analyzed using a nonparametric ANOVA which resulted in a significant main effect of color coding, $F(1, 148) = 30.19$, $p < .001$, $\eta_p^2 = 0.17$. The other effects did not reach significance ($ps > .195$). Color coding reduced participants' cognitive load ratings as expected.

6.2.2. Retention

Retention scores (see Fig. 6B) also required a nonparametric ANOVA and revealed significant main effects of shape distinctness and color coding, $F(1, 148) = 4.38$, $p = .038$, $\eta_p^2 = 0.03$, and $F(1, 148) = 4.16$, $p = .043$, $\eta_p^2 = 0.03$. The interaction effect did not reach significance, $p = .195$. Therefore, shape distinctness and color coding increased retention performance as hypothesized. However, learning with nondistinct shapes was not enhanced by color coding in a compensatory manner.

6.2.3. Transfer

The transfer data (see Fig. 6C) did not result in significant main effects for shape distinctness and color coding ($p = .059$ and $p = .070$, respectively) using a nonparametric ANOVA. However, the data follow the same general trend of the retention data on the descriptive level. The interaction effect did not reach significance, $p = .393$.

7. General discussion

Three experiments were conducted to investigate how visual design features affect the cognitive processing of realistic visualizations. The first two experiments assessed the alignment hypothesis claiming that realism is particularly useful in situations that use the specific advantages of realistic imagery. In Experiment 1, the dimensionality of the objects presented in a visualization was manipulated together with the level of realism and revealed that subjective cognitive load was indeed highest if there is a mismatch between realism and dimensionality (e.g., if a flat cross-section is presented using realistic details) and lowest if there is an alignment (e.g., if a cross-section is presented in a schematic style). The retention data followed an inverted trend of this pattern on the descriptive level that did not reach significance, potentially due to a slight ceiling effect. However, Experiment 1 provided partial support for the alignment hypothesis.

Experiment 2 again tested the alignment hypothesis, but using the factor perspective instead of dimensionality. It was hypothesized that the realistic rendering would particularly help learners with the perspective view, as this layout could benefit from the availability of depth cues, while the schematic version would potentially suffice for memorizing the side view. On the descriptive level, the retention data follow this pattern, but this interaction effect did not reach significance.

The results of the first two experiments suggest that the alignment hypothesis may be correct as indicated by the cognitive load ratings in Experiment 1 and the general pattern of the results, but also indicate that this effect is not strong enough to explain the divergent results concerning realism found in the literature. Additional studies with far greater sample sizes are needed to precisely ascertain how strongly an alignment between realism, dimensionality, and perspective will impact learners. However, these promising results suggest that the design of visualizations will probably benefit from such an alignment, at least to a limited extent. Given some of the negative results of realism found in the

[1] The study was completed before Experiments 1 and 2, but is reported after these two in order to present an increase in visual changes between the experimental conditions from study to study.

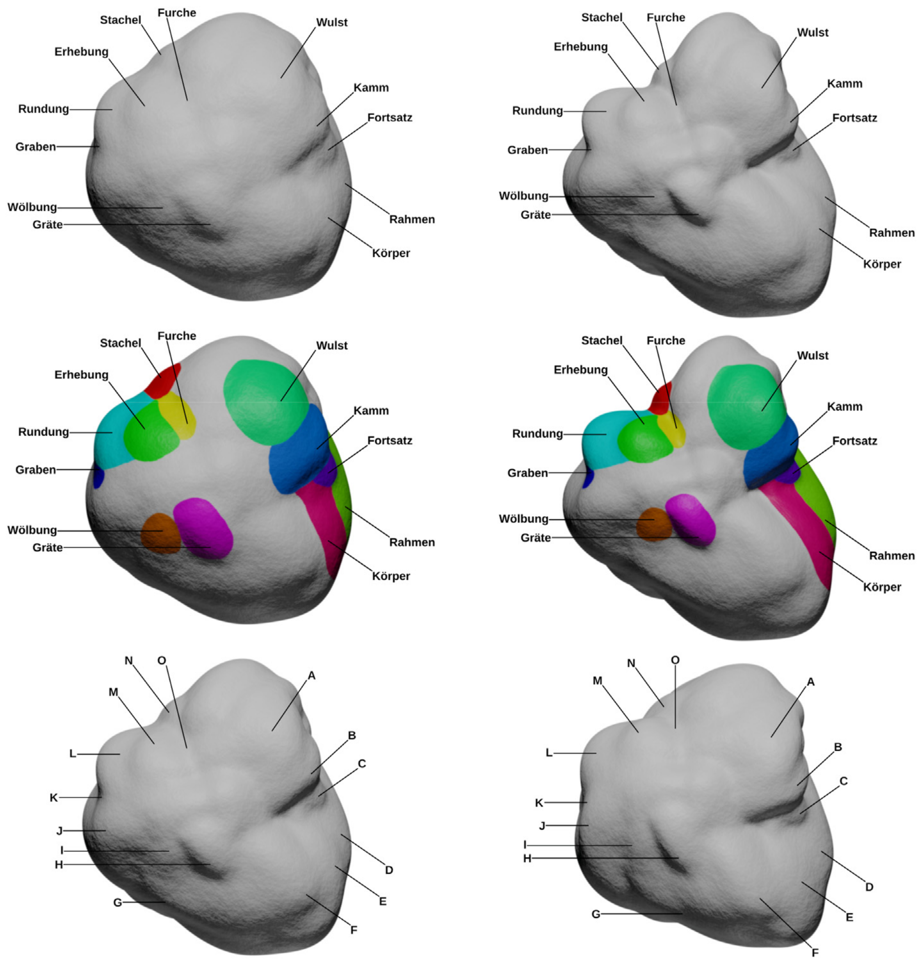

Fig. 5. Experiment 3 reused some visualizations of the second experiment in Skulmowski (2022b, p. 685). Additional visualizations were created using the original source files. The top and middle panel on the left column shows the model with a low shape distinctness, while the two panels on the right present the model with a high shape distinctness. Color coding was not included in the versions in the top row, but was applied to the model seen in the middle row. The bottom-left panel presents the retention test and the transfer test is shown in the bottom-right corner. The upper third of the figure is reproduced from Fig. 3c and d in Skulmowski (2022b, p. 685), the middle third of the figure is adapted from Fig. 3c and d in Skulmowski (2022b, p. 685) by applying color to the surfaces, and the lower third of the figure is an adaptation of Fig. 3e and f in Skulmowski (2022b, p. 685) created using a different rendering style and omitting the outline. Reproduced and adapted from "Realistic Visualizations Can Aid Transfer Performance: Do Distinctive Shapes and Descriptive Labels Contribute Towards Learning?," by A. Skulmowski, 2022, *Journal of Computer Assisted Learning*, *38*, p. 685 (https://doi.org/10.1111/jcal.12640). © 2022 Skulmowski (licensed under the Creative Commons Attribution License, http://creativecommons.org/licenses/by/4.0). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

literature (e.g., Scheiter et al., 2009), the cognitive design of realistic visualizations should make use of every single potential advantage. On a more general note, the first two experiments neither revealed an overall superiority of realistic nor of schematic visualizations. Considering the at times quite negative outlook on the use of realistic visualizations (e.g., Smallman & St. John, 2005), this result should be acknowledged in the design of visualizations. In the first two experiments, realism did not induce a higher cognitive load. Rather, realism even lowered extraneous cognitive load on the descriptive level. Although multiple studies have revealed that realism can raise the subjective cognitive load experienced by learners (e.g., Skulmowski, 2022a; 2023b). This may be a consequence of the realistic versions used in the present studies not introducing major surface detail through shading or complex geometry. According to the GSR model, the main dimension that the visualizations differ on is rendering, suggesting that changing the rendering mode to a more realistic one does not necessarily induce cognitive load.

Experiment 3 was conducted to get a clearer perspective on how to support learning with realistic visualizations by implementing semantically meaningful designs. While the other experiments primarily varied the shading and rendering dimensions of the GSR model, the last study assessed the effects of variations in the geometry and shading of realistic visualizations. Furthermore, the interaction between these two dimensions was tested. The experiment demonstrated that a geometry featuring a higher level of shape distinctness facilitates learning by lowering cognitive load. A higher shape distinctness also led to higher retention scores than a visualization showing a nondistinct shape. Similarly, separating elements of the model using color coding positively affected retention and indicates that shading can be used to lower cognitive load and enhance learning (thereby also confirming previous studies, e.g., Skulmowski & Rey, 2018). However, these two dimensions did not interact with each other in a significant manner on any of the learning outcomes.

The present studies extend previous research by allowing to distinguish between the effects that different design factors have on learning. In a previous study, Huk et al. (2010) found a positive effect of a realistic rendering style combined with a 3D perspective when compared with a schematic style presented as a side view. The results of Experiments 1 and 2 suggest that this advantage may only be in part attributed to the perspective and depth of the realistic rendering. Previous research has shown that the overall style has a far greater positive or negative effect on learning (e.g., Scheiter et al., 2009; Skulmowski, 2022b; Skulmowski, 2023b; Skulmowski & Rey, 2021) than minor design aspects (e.g., He et al., 2022; Keller et al., 2006). However, the semantic information conveyed using more distinct shapes has a major impact on learning. This could not be shown conclusively in previous studies that only found interaction effects of distinct shapes, such as that for realistic renderings, cognitive load is increased if less distinct shapes are presented to learners (Skulmowski, 2022b). When considered in conjunction with previous research, the present studies suggest that the influence of perspective is negligible, dimensionality can affect subjective perceptions of cognitive load, and distinct shapes and segmentation have a measurable effect on retention performance.

In sum, the experiments suggest that the effects of visual design choices that enhance or facilitate the cognitive processing of elements in realistic visualizations are largest if these design choices amplify semantic relationships. An alignment between the specific advantages that realism offers (such as depth cues and shadows) with the dimensionality of objects (rather than the flatness of cross-sections) may offer a minimal

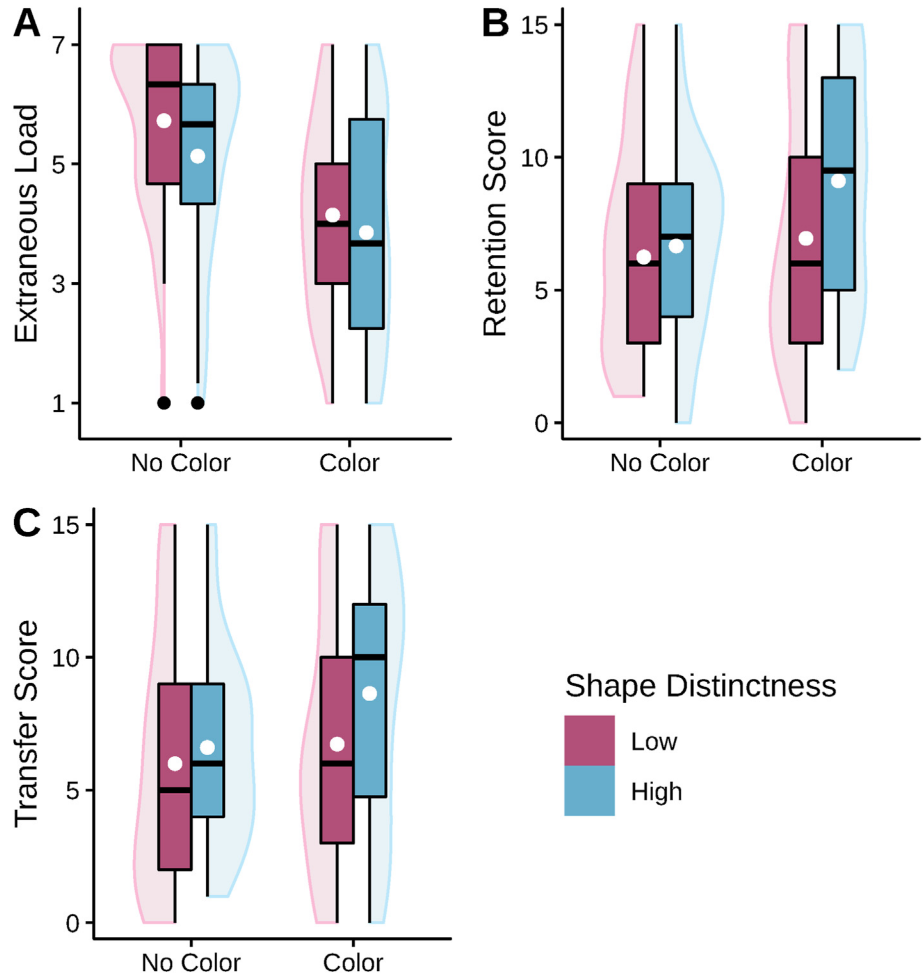

Fig. 6. Boxplots with violin elements of the data from Experiment 3. Fig. 6A: The averaged cognitive load data. Fig. 6B: The retention scores. Fig. 6C: The transfer scores. White dots indicate the means of groups.

advantage, but the magnitude of such effects appears to be rather small. Nevertheless, it might be advisable that visualizations should be designed with an alignment optimization in mind, at least when in doubt. However, Experiment 3 provides solid evidence for the claim that the design of realistic visualizations should emphasize distinct shapes and provide visual guides for segmentation. This does not necessarily be achieved through similarly bright colors as in the study, as the use of different shaders (such as muscle shaders displaying fibers and fat shaders highlighting glossiness) might be sufficient to make different parts in a visualization stand out from each other.

7.1. Implications for the design of XR education

The results of the present studies will be helpful for designers of static visualizations, but can also be the starting point of investigations of alignment factors in XR. For instance, when designing an application that primarily displays flat, orthographic cross-section views, it may be advisable to include less detailed models. However, for situations in which the models may be seen from different perspectives emphasizing their dimensionality, a higher level of realism may be advisable. Given the rather small impact of dimensionality and perspective on cognitive processing found in the present studies, it will be interesting to assess whether these effects increase as soon as learners are confronted with a multitude of different models in XR. Since learners in more comprehensive XR application are likely to go through the (cognitive) processes involved in the studies multiple times if several models are being presented, the advantages of an alignment between dimensionality, perspective, and realism may be more pronounced. More XR research is needed to confirm this hypothesis.

As recently discussed (Skulmowski, 2023b), the effects that visual design can have on learners using visualizations does not necessarily correspond to the impact of design in virtual reality (based on results of Liberman & Dubovi, 2022). Thus, it needs to be assessed whether perspective, dimensionality, and shape distinctness play a role in different XR implementations. For instance, the factor of perspective may be investigated in different forms in XR-based realism research. In virtual reality, learners will usually choose their own perspective by adjusting their head tilt, possibly making it difficult to interfere in this process by forcing a certain perspective on learners. However, even in virtual reality, the design of instructional applications could be optimized by finding ways in which the perspective of virtual objects could emphasize depth. Previous research suggests that so-called *canonical* views (such as typical side or top views rather than oblique or unfamiliar rotations) are easier to process for learners (Stull et al., 2009; see also Skulmowski & Rey, 2020). For mixed reality settings in which learners can rotate objects (or stand-in objects) themselves, it may be beneficial to give learners an indication of when a canonical view has been reached (see Stull et al., 2009), while also giving them the opportunity to understand the spatial structure using free rotation. Similar hypotheses could be derived for augmented reality research. In all of these settings, it should also be investigated whether the dimensionality of the objects have an effect (such as in the augmented reality study by Krüger et al., 2022).

7.2. Limitations and outlook

The experiments share several limitations with other similar studies (e.g., Skulmowski, 2022b; Skulmowski & Rey, 2021). One of these limitations of the methodology used for the studies is the relatively short learning time. This short time in which only one visualization is at the center of attention has proven itself as a method for investigating the effects of realism in instructional visualizations, but the effects found in the studies should be checked regarding their generalizability to other, more extensive settings. A longer learning time would not have been feasible for the present studies, as some studies already resulted in slight ceiling effects concerning retention performance that would only have been exacerbated with longer learning times. In addition, the countdown timer presented together with the instructional visualizations and the fact that some people may find the depicted organs somewhat "gross" should be considered. It is possible that the learning performance is higher without features that may distract learners, such as being disgusted by visualizations. However, the level of detail present in the visualizations used does not surpass the realism of medical visualizations that can be found in everyday life.

As there have been instances of age-related differences in how much learners can benefit from realism (Menendez et al., 2020, 2022), the aspect of age should be considered in future studies. While the present studies used a homogeneous sample of university students aged between 18 and 30 years, the studies should be replicated with different samples and age groups.

Due to the slight rotation of the visualizations used in Experiment 2, it may have been helpful to measure participants' spatial abilities. However, since the experiments by Huk et al. (2010) found no effects of spatial ability with their related experimental manipulation, it is unlikely that spatial ability had an influence on the present results.

7.3. Conclusion

The present studies provide evidence that using the strengths of realism, such as depth cues, can facilitate cognitive processing to a certain extent. However, the studies underline that in order for realism to have an impact on learning, the semantic content of visualizations needs to be increased through more distinct shapes or methods of visual segmentation. In summary, realism was not shown to be an obstacle for learning, but the optimization of learning with realistic visualizations may require profound changes in the design of visualizations.

Funding

No funding was received for conducting these studies.

Code availability

Not applicable.

Ethics approval

The experiments presented in this article do not require ethics review in Germany by law. The national and institutional guidelines have been followed.

Consent to participate

All participants provided their informed consent to participate in the experiments.

Consent for publication

Not applicable.

Data availability statement

The data of the studies is available from the author on request.

Declaration of competing interest

The author serves as an Editorial Board Member for the journals *Educational Psychology Review* and *Journal of Computer Assisted Learning* and has been a member of the Editorial Board of *Human Behavior and Emerging Technologies*.

References

Bammes, G. (2009). *Die Gestalt des Menschen*. Freiburg, Germany: Christophorus Verlag.

Belenky, D. M., & Schalk, L. (2014). The effects of idealized and grounded materials on learning, transfer, and interest: An organizing framework for categorizing external knowledge representations. *Educational Psychology Review, 26*(1), 27–50.

Berney, S., Bétrancourt, M., Molinari, G., & Hoyek, N. (2015). How spatial abilities and dynamic visualizations interplay when learning functional anatomy with 3D anatomical models. *Anatomical Sciences Education, 8*, 452–462.

Biederman, I. (1985). Human image understanding: Recent research and a theory. *Computer Vision, Graphics, and Image Processing, 32*, 29–73.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115–147.

Chen, S. C., Hsiao, M. S., & She, H. C. (2015). The effects of static versus dynamic 3D representations on 10th grade students' atomic orbital mental model construction: Evidence from eye movement behaviors. *Computers in Human Behavior, 53*, 169–180.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences, 14*, 69–110.

He, C., Gunalp, P., Meyerhoff, H. S., Rathbun, Z., Stieff, M., Franconeri, S. L., & Hegarty, M. (2022). Visual working memory for connected 3D objects: Effects of stimulus complexity, dimensionality and connectivity. *Cognitive Research: Principles and Implications, 7*, 19. https://doi.org/10.1186/s41235-022-00367-9

Höst, G., Schönborn, K. J., & Tibell, L. (2022). Visual images of the biological microcosmos: Viewers' perception of realism, preference, and desire to explore. *Frontiers in Education*. https://doi.org/10.3389/feduc.2022.933087

Huk, T., Steinke, M., & Floto, C. (2010). The educational value of visual cues and 3D-representational format in a computer animation under restricted and realistic conditions. *Instructional Science, 38*(5), 455–469.

JASP Team. (2022). *JASP (Version Version 0.16.1) [Computer software]*.

Keller, T., Gerjets, P., Scheiter, K., & Garsoffky, B. (2006). Information visualizations for knowledge acquisition: The impact of dimensionality and color coding. *Computers in Human Behavior, 22*(1), 43–65.

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology, 8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997

Krüger, J. M., Palzer, K., & Bodemer, D. (2022). Learning with augmented reality: Impact of dimensionality and spatial abilities. *Computers and Education Open, 3*, Article 100065. https://doi.org/10.1016/j.caeo.2021.100065

Lavie, N., Lin, Z., Zokaei, N., & Thoma, V. (2009). The role of perceptual load in object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1346–1358.

Leiner, D. J. (2021). *SoSci Survey [computer software]*.

Liberman, L., & Dubovi, I. (2022). The effect of the modality principle to support learning with virtual reality: An eye-tracking and electrodermal activity study. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12763

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Menendez, D., Rosengren, K. S., & Alibali, M. W. (2020). Do details bug you? Effects of perceptual richness in learning about biological change. *Applied Cognitive Psychology, 34*(5), 1101–1117.

Menendez, D., Rosengren, K. S., & Alibali, M. W. (2022). Detailed bugs or bugging details? The influence of perceptual richness across elementary school years. *Journal of Experimental Child Psychology, 213*, Article 105269.

Nebel, S., Beege, M., Schneider, S., & Rey, G. D. (2020). A review of photogrammetry and photorealistic 3D models in education from a psychological perspective. *Frontiers in Education, 5*, 144. https://doi.org/10.3389/feduc.2020.00144

Pett, D., & Wilson, T. (1996). Color research and its application to the design of instructional materials. *Educational Technology Research & Development, 44*(3), 19–35.

R Core Team. (2021). *R: A language and environment for statistical computing (version 4.0.5) [computer software]. Vienna, Austria*. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org.

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learning and Instruction, 19*(6), 481–494.

Skulmowski, A. (2022a). Is there an optimum of realism in computer-generated instructional visualizations? *Education and Information Technologies, 27*, 10309–10326.

Skulmowski, A. (2022b). Realistic visualizations can aid transfer performance: Do distinctive shapes and descriptive labels contribute towards learning? *Journal of Computer Assisted Learning, 38*, 681–691.

Skulmowski, A. (2022c). When color coding backfires: A guidance reversal effect when learning with realistic visualizations. *Education and Information Technologies, 27*, 4621–4636.

Skulmowski, A. (2023a). Guidelines for choosing cognitive load measures in perceptually rich environments. *Mind, Brain, and Education, 17*, 20–28.

Skulmowski, A. (2023b). Realistic details impact learners independently of split-attention effects. *Cognitive Processing*. https://doi.org/10.1007/s10339-022-01123-z

Skulmowski, A., Nebel, S., Remmele, M., & Rey, G. D. (2022). Is a preference for realism really naive after all? A cognitive model of learning with realistic visualizations. *Educational Psychology Review, 34*, 649–675.

Skulmowski, A., & Rey, G. D. (2018). Realistic details in visualizations require color cues to foster retention. *Computers & Education, 122*, 23–31.

Skulmowski, A., & Rey, G. D. (2020). The realism paradox: Realism can act as a form of signaling despite being associated with cognitive load. *Human Behavior and Emerging Technologies, 2*(3), 251–258.

Skulmowski, A., & Rey, G. D. (2021). Realism as a retrieval cue: Evidence for concreteness-specific effects of realistic, schematic, and verbal components of visualizations on learning and testing. *Human Behavior and Emerging Technologies, 3*(2), 283–295.

Skulmowski, A., & Xu, K. M. (2022). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review, 34*, 171–196.

Smallman, H. S., & St John, M. (2005). Naive realism: Misplaced faith in realistic displays. *Ergonomics In Design, 13*, 6–13.

Stull, A. T., Hegarty, M., & Mayer, R. E. (2009). Getting a handle on learning anatomy with interactive three-dimensional graphics. *Journal of Educational Psychology, 101*(4), 803–816.

Sweller, J., van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292.

Tillmann, B. N. (2016). *Atlas der Anatomie des Menschen* (3rd ed.). Berlin, Germany: Springer-Verlag.

Wickham, H. (2016). *ggplot2: Elegant Graphics for data analysis*. New York: Springer-Verlag.

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the ACM conference on human factors in computing systems (CHI '11). Vancouver, British columbia* (pp. 143–146). New York, NY: ACM Press (May 7–12, 2011).

Wu, H. K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education, 88*, 465–492.