Human Perception and Classification of AI-Generated Images: A Pre-Study based on a Sample from the Media Sector in Germany

Stephan Böhm

CAEBUS Center of Advanced E-Business Studies
RheinMain University of Applied Sciences
Wiesbaden, Germany
e-mail: stephan.boehm@hs-rm.de

***Abstract*—Recent advances in Generative Artificial Intelligence (AI) have significantly expanded and improved image generation and processing possibilities. Applications, such as DALL-E, Midjourney, and Stable Diffusion have simplified Generative AI for non-technicians and made it accessible to a broad audience. The quality of the generated images has steadily increased in recent months, with photo-realistic representations almost indistinguishable from real photos. AI-based image generation and editing methods are also becoming increasingly accessible for professional use, where high-quality image generation and editing were formerly reserved for specially trained personnel. However, the perception of Generative AI's results and potential depends not only on image quality. Human users may have reservations or a biased assessment of the performance of AI for image generation, for example, because they doubt the creativity of AI or fear the substitution of jobs. Against this background, a pre-study with a sample of N = 172 participants from the media sector in Germany is presented. The participants were asked about their attitudes towards image-generating AI and had to assess a test set of images regarding quality and type of generation. The results show that while minor differences in quality are observed, classification precision is almost independent of the quality rating and the participants' attitudes or experiences. The study supports the conclusion that even representatives from the media sector cannot systematically recognize AI-generated images based on image quality at the current performance level of image-generating Generative AI.**

Keywords–Generative AI; AI-based media disruption; AI-generated images; human perception of AI; identification of AI-generated images.

I. Introduction

Recent advances have significantly influenced the development of Generative AI in image generation and processing in machine learning and visual computing. In particular, the introduction of Generative Adversarial Networks (GANs) has played a crucial role in automatic image generation with computers. GANs have revolutionized the field by introducing a framework in which two neural networks, the generator and the discriminator, compete against each other to produce high-quality synthetic images [1]. Corresponding approaches to generating realistic images have proven extremely effective and have pushed the boundaries of what AI can achieve in image generation.

In addition, the emergence of Large Language Models (LLM) has significantly influenced the spread of AI technologies within a non-technical audience [2]. These models, such as GPT-3, GPT-4 and most recently GPT-4o in the ChatGPT application [3], have demonstrated and popularized the potential of using natural language prompts to a mass audience. In this context, creating, modifying, and editing images based on detailed descriptions in natural language has gained notoriety and spread rapidly [4].

By training with huge amounts of data, these models can now understand and interpret human input to produce corresponding visual results, which also democratizes the creation of images with the help of AI. Whereas in the past, more sophisticated types of digital image editing were reserved for experts trained in the operation of specialized software, impressive results can now be achieved by appropriately describing the results as part of the prompt engineering of an image-generating Generative AI.

Since the launch of ChatGPT in November 2022 [5], significant qualitative improvements have been achieved in generating images with AI. For example, ChatGPT and subsequent solutions have demonstrated remarkable capabilities in creating images that closely resemble authentic photographs, blurring the lines between human and AI-generated content [6]. These advances have meant that distinguishing between the two has become difficult, highlighting the rapid progress of AI technology in mimicking human creativity and perception.

However, there are still limitations in the professional use of AI-generated images. Issues, such as maintaining consistency of style, context, and coherence in the generated images remain an obstacle to the productive and regular integration of AI-generated content in various domains. Ongoing research and development work continuously addresses these challenges to improve the quality and authenticity of AI-generated images. Especially, it is becoming increasingly difficult for humans to distinguish real photographs from AI-generated images, which is reflected in the increased research interest in so-called ”deep fakes” [7]–[9]. However, the higher performance and greater difficulty distinguishing AI-generated content applies not only to photo-realistic images but also to creative works, such as illustrations and artworks.

Against this background, the remainder of this paper is structured as follows: After this introduction, Section II presents the research background on the attitude towards AI-generated content, image classification, and quality criteria before formulating the research questions of this study. Building on this, Section III presents this study's survey and test design.

Section IV then presents the results of this pre-study, followed by the conclusion in Section V. Finally, Section VI concludes with limitations and an outlook for further research.

II. Research Background and Objectives

This section provides a brief overview of related research in the perception and evaluation of AI-generated imagery and then narrows down the research questions of this study.

A. *Attitude towards AI-generated Content*

An important field of research on innovative Generative AI tools is how AI-generated images are perceived. The perception of AI-generated images and art by humans is a complex and evolving area of research. Research suggests that there is a bias towards such computer-generated art. Studies show that people tend to differentiate between AI and human-made art, often undervaluing the former [10]. This bias may be due to the perceived effort involved in creating art, as AI-generated artworks are sometimes seen as less effortful compared to traditional art forms [11]. However, efforts to anthropomorphize AI systems, e.g., by highlighting the role of human programmers and software as collaborators, may help to counteract this bias and facilitate the consideration of AI-generated outputs as genuine artworks [12].

In addition, the attribution of creativity to AI systems plays an important role in how AI-generated art is perceived. Studies have shown that people's attitudes towards AI-generated aesthetics are influenced by their perception of the AI's capabilities and creativity [13]. This could be because Generative AI represents models trained to uncover and replicate design patterns, and therefore, AI is denied the ability to create something novel. Furthermore, the perceived partnership between humans and AI in the creation process, where humans develop the code for AI algorithms and provide instructions to generate art, can increase the value and appreciation of AI-generated artworks [14].

The evaluation of AI-generated art is not only influenced by the artwork itself but also by the context in which it is presented. Factors, such as the explicit and implicit perception of AI-generated art in different cultural contexts can affect how these artworks are received [15]. Furthermore, understanding AI capabilities in generating images in different domains, such as clinical settings, may influence human perceptions of the quality and reliability of AI-generated content [16].

Another tension in the perception of Generative AI tools is that such innovative solutions can be perceived as support or opportunities to increase productivity, but also as a threat to the company's activities and a risk of job substitution [17]. With regard to the application of (Generative) AI in general, there are already studies that aim to estimate corresponding increases in productivity or implications for the workplace [18]–[20]. In the creative and media sector in particular, however, hardly any studies still examine such attitudes and correlations.

In summary, human perceptions of AI-generated images and artworks are complex and influenced by biases, perceived effort, attributions of creativity, cultural contexts, and understanding of AI's capabilities. As AI plays an increasing role in creative endeavors, further research is needed to investigate how these perceptions evolve and shape the interaction between humans and AI in creative industries like the media sector.

B. *Image Classification*

There are several different research approaches to the issue of recognizing AI-generated images. Firstly, there are technical approaches that relate, for example, to the analysis of image artifacts and pixel patterns resulting from the generation process [21]. Other approaches use machine learning methods (e.g., Contrastive Language-Image Pre-training (CLIP) [22], Convolutional Neural Networks, and Transfer Learning [23]) to differentiate between real and AI-generated images. In the context of research into deep fakes, several studies have already investigated how humans can distinguish real authentic images and videos – often related to the representation of people or human faces – from those that appear realistic but are faked AI-generated content using deep learning technology [24]–[27].

This study also aims to investigate human's ability to distinguish between images generated with and without AI. However, the focus here is less on the actual recognition performance or the ability to detect non-authentic image material but more on the relationships between the classification decision, perceived image quality, and the attitude towards Generative AI of employees in the media sector. However, there is still a need for research in the media sector, while more research has already been published for AI-generated artwork. For example, several studies have investigated people's ability to distinguish between human-generated and AI-generated art. Chamberlain et al. [10] found a bias towards computer-generated art and emphasized the difficulties distinguishing between human-generated and AI-generated artworks. Gangadharbatla [11] examined the impact of knowledge of art attribution in evaluation and focused on the accuracy of the distinction. Zhou and Kawabata [28], and Gu and Li [29] also investigated participants' ability to distinguish between human-created and AI-generated artworks, with different results in detail. The studies by Lyu et al. [30], and Natale and Henrickson [12] likewise reported mixed results, i.e., some participants correctly recognized AI-generated artworks, while others had difficulty doing so.

Horton et al. [31] emphasized that comparing human and AI-generated art can improve the perception of human creativity. In addition, Fortuna et al. [32] emphasized that individual evaluation schemes influence the differences in evaluating AI- and human-generated artworks. Another study by Ho [33] discussed social and ethical issues related to AI-generated art, while Rasrichai et al. [34] provided insights into how presumed knowledge of an artist's identity influences the evaluation of artworks. With regard to the use of images in the media sector, it is not so much individual attribution, uniqueness, or artistic impression that is important; rather, images are often used for visualization, explanation, and to create context. Therefore, the results from the art sector are transferable, but only to a limited extent. In conclusion, the issue of distinguishing between AI and human-generated imagery has so far been considered primarily from the perspective of art and artists, but there is still a need for research in the media.

C. *Image Quality Evaluation*

There are several approaches to evaluating the quality characteristics of an image based on the analysis of corresponding psychological factors and cognitive evaluation processes of works of art. For example, criteria for evaluating image quality could be derived from studies based on established theories of aesthetic judgment and the psychological processing of art. The model of Leder et al. [35] outlines stages of perceptual analysis, which includes initial reception and basic features, such as color and composition, to cognitive coping and evaluation, which includes more subjective and complex judgments, such as creativity and narrative understanding. Graf and Landwehr [36] propose a model that distinguishes between the pleasurable and interesting aspects of aesthetic experience. Their work is important for understanding how different aspects of an artwork, including its emotional impact and originality, contribute to the overall aesthetic evaluation.

From a simplified transfer of the findings of this work, relevant criteria for the present study can be derived for the qualitative evaluation of images, such as (1) detail and texture quality, (2) color harmony, (3) composition and structure, (4) creativity and originality, (5) emotional impact, and (6) narrative perception. These criteria have not been taken directly from the aforementioned research but are based on essential findings for the evaluation of works of art and transfer them to the quality assessment of images. For further details, it is referred to the corresponding literature [35][36].

D. Research Objectives

Based on the previous explanations and the identified research needs, the following research questions have been formulated for this pre-study:

- To what extent are Generative AI tools already widespread in the media sector sample, and how is the work-related impact of this new technology on the working environment perceived?
- How is the quality of AI-generated images perceived, and to what extent does this quality assessment influence the classification of images as AI-generated?
- To what extent is the precision of the classification of AI-generated images of the participants dependent on their experience with digital image processing, AI tools, and attitudes towards Generative AI?

This pre-study will assess these research questions in a sample of working adults from the media sector. The procedure and results are described in the following sections.

III. Survey and Test Design

A questionnaire was developed to answer the research questions defined in the previous section. The questionnaire contains parts on the participants' characteristics, experiences, and attitudes toward image generation by Generative AI, as well as a section in which AI and non-AI-generated (real) images are to be evaluated in terms of their quality and classified concerning the type of image generation. The questionnaire had no time restrictions for answering the questions, and the participants could decide how long they wanted to look at the pictures. The structure of this questionnaire is described in more detail below.

A. Survey Contents and Structure

The questionnaire was realized as an online questionnaire using the survey software Unipark [37]. The questionnaire was distributed via a link and answered in the web browser. The survey was divided into four sections:

- *Sample characteristics:* At the beginning of the questionnaire, basic demographic information, such as age, gender, educational qualifications, and employment status, was collected to analyze the demographic profile of the study participants.
- *Experience with digital image editing and Generative AI:* Then participants were asked about their experience with digital imaging and various AI applications for image generation. This involves determining the extent to which the participants have come into contact with digital image editing privately, during their education, or professionally and which specific AI tools they know and use.
- *Attitude towards the impact of Generative AI:* Next, the participants were asked to express their opinion on the impact of AI. This involves an assessment of potential job losses, productivity increases, threats to copyright, and the general quality of AI-generated images compared to human creation.
- *Evaluation of AI and non-AI generated (real) images:* The main part of the questionnaire focused on the evaluation of six different images generated either by humans or by AI. Participants were asked to evaluate various aspects of image quality, including detail, color harmony, composition, creativity, emotional impact, and narrative elements. They also had to assess whether the images shown were created by AI and how confident they were in their assessment.

The questionnaire concluded with individual overall assessments of the difficulty of the classification task and the importance of quality features.

B. Image Evaluation and Classification

For this part of the evaluation of images, a set of images had to be defined first. The Kaggle Data Set "AI-Generated Images vs. Real Images" [38] was used for this purpose. Three AI-generated and three non-AI-generated images were selected from the data set to keep the processing time acceptable for the participants. Because the motif could influence the evaluation, three pairs of images with similar compositions were used in each case. The first image was selected randomly, and then a similar composition with a contrary form of image generation was searched for in the data set. It was ensured that no well-known images by popular artists were used and that the images did not contain any watermarks or signatures of artists. The following images were selected for presentation:

- *Photo-realistic images of animals:* A lion in an unnatural pose (AI-generated, Image 1) and a parrot in close-up (real, Image 2).
- *Photo-realistic portraits:* A side portrait of a woman (real, Image 3) and a frontal portrait of a woman (AI-generated, Image 4).
- *Abstract landscapes:* Naive depiction of a country house (real, Image 5) and a colorful abstract valley with a river (AI-generated, Image 6).

The images in the dataset were crawled from the web and cannot be printed here due to unresolved copyrights. However, the filenames provided in the Appendix can identify them in the dataset.

Each image was presented in a separate section in high resolution in the online questionnaire. The respondents were first asked to evaluate the images in terms of image quality using the following criteria as discussed in Section II-C:

- *Detail and texture quality:* Evaluation of the image's perceived level of detail and texture.

- *Color harmony:* Evaluation of the harmony and appropriateness of the use of color.
- *Composition and structure:* Evaluation of the structural composition of the image.
- *Creativity and originality:* Evaluation of the creativity expressed in the image and its originality.
- *Emotional impact:* Determination of the extent to which the image is emotionally appealing.
- *Narrative perception:* Evaluation of whether the image tells a story or conveys a recognizable message.

Participants were also asked whether they thought it was AI-generated or non-AI-generated for each image. In addition, the certainty of the decision was to be indicated, and the quality criteria were to be ranked in terms of their importance in the classification decision, with at least one important criterion to be selected.

IV. Findings of the Study

A. Survey Implementation and Sample Characteristics

The survey was conducted via a panel provider in mid-May 2024. The panel included men and women over 18 years who live in Germany and are particularly media-savvy, i.e., come from media companies and media degree programs or have completed vocational training in the media sector. However, there were no filter questions to exclude participants. This was done against the background that the sample was narrowed down to the media sector, but in principle, everyone could participate in the questionnaire. A total of 189 participants completed the survey. Responses less than a quarter or three times as long as the median survey duration were excluded. As a result, 172 responses were left in the sample and analyzed further. As Table I shows, the study participants are predominantly men (60.5%) with a bachelor's, master's, or diploma degree (55.2%) who work as employees (79.7%). The sample is, therefore, not representative of the population in Germany or a specific, definable target group in the media sector. However, this pre-study focuses on fundamental relationships between attitudes towards generating an image with Generative AI and identifying AI-generated images. The results obtained, therefore, remain meaningful as an initial indication but can only be applied to the sector as a whole to a limited extent.

TABLE I. SAMPLE DEMOGRAPHICS

	Count	Percentage
Age (Years)		
< 25	6	3.5%
26-35	62	36.0%
36-50	50	29.1%
> 50	54	31.4%
Gender		
Male	104	60.5%
Female	68	39.5%
Highest Educational Qualification		
Vocational qualification	29	16.9%
Bachelor	45	26.2%
Master, Diploma, etc.	50	29.1%
Other	48	27.9%
Employment		
Employee	137	79.7%
Civil servant	5	2.9%
Self-employed	25	14.5%
Other	2	2.9%
Total	**172**	**100.0%**

B. Digital Imaging Experience and Use of Generative AI

Almost all of the participants have already gained experience with digital image editing in the private sphere or as a hobby (87.7%), in training and studies (68.0%) or at work or in a company (79.5%). These results initially show that knowledge in the field of digital imaging is not only reserved for specialists and experts in a professional context but is now also widely used in everyday life. Comprehensive experience in digital image editing (rather or very many) was found most often in the private sphere and hobbies (50.9%), while such an extent of experience in training and studies (39.0%), as well as at work (49.1%) was less stated. In terms of duration, most of the participants had a total of 6-10 years of experience with digital image editing (none: 18.6%, 0-2 years: 16.9%, 3-5 years: 19.8%, 6-10 years: 26.7%, 11-20 years: 12.8%, 21 years or more: 5.2%).

Table II shows the popularity and frequency of using AI-based applications for image creation and editing in the sample (a selection of tools known and used in Germany was chosen [39]). The best-known applications (the tool is used or at least known) are Adobe Firefly (66.7%), DALL-E (54.1%), Midjourney (53.8%), and Bing Image Creator (53.6%). Therefore, more than half of the respondents already know about image creation and editing methods with Generative AI. However, the proportion of those who have already used such applications is significantly lower. Only with Adobe Firefly, more than half of the participants in the study already gained experience of use (50.3%), while this otherwise fluctuates between 39.9% (Bing Image Creator) and 34.5% (Jasper Art). The proportion of those who use Generative AI applications almost daily is still below ten percent and highest for Adobe Firefly (8.8%) and DALL-E (8.1%). The high prevalence of Adobe applications can be explained by the fact that the people in the sample are media-savvy, and Adobe products are the industry standard in the media sector and creative industries.

TABLE II. POPULARITY AND USAGE FREQUENCY OF SELECTED AI TOOLS

	I do not use	I do know, but haven't used it yet	Very rare, only tried out so far	Irregularly, on occasion	Regularly, several times a week	Regularly, almost every day
DALL-E	45.9%	17.4%	8.1%	11.0%	9.3%	8.1%
Midjourney	46.2%	15.2%	9.9%	9.4%	13.5%	5.8%
Stable Diffusion	48.8%	15.7%	8.1%	11.0%	11.0%	5.2%
Adobe Firefly	33.3%	16.4%	10.5%	15.2%	15.8%	8.8%
Bing Image Creator	46.4%	13.7%	6.5%	14.3%	13.7%	5.4%
Jasper Art	46.8%	53.2%	5.8%	12.3%	10.5%	5.8%

In the next section of the questionnaire, the study participants were asked about their agreement with predetermined statements on the impact of using Generative AI tools for generating images ("To what extent do you agree with the following statements on the generation of images with AI?"). A 5-point Likert scale was used for the feedback ("Fully agree", ..., "Do not agree at all"). Figure 1 shows the results for this question as a percentage of the selected response options. For all six questions, it can initially be seen that around a third of respondents are still undecided about the impact the use of AI will have in this area.

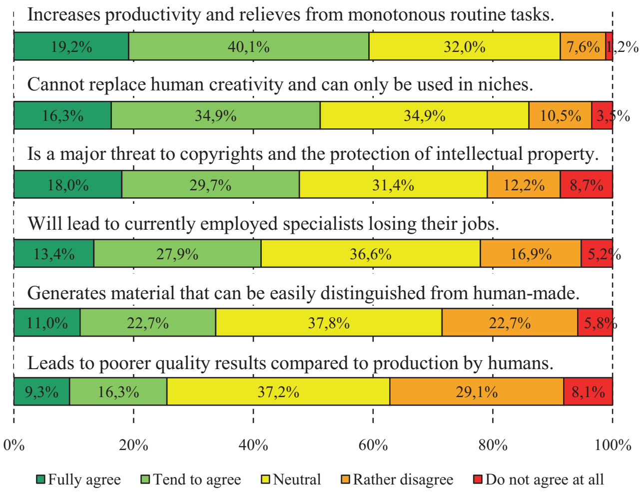

Figure 1. Respondents Agreement Level to Statements on the Impact of AI-based Image Generation.

However, a majority of the participants expect AI to increase productivity and relieve them of routine tasks. Among those who have formed an opinion, the prevailing attitude is that AI cannot replace human creativity and can only be used in niches. The stated agreement also outweighs disagreement regarding the threat to copyrights, the substitution of jobs, and, with a small difference only, that AI-generated images are easy to distinguish. This differs from the statement on the lower quality, which was rejected by a significantly larger proportion of respondents than agreed with. These results thus reflect the findings of other studies that were previously mentioned. Although a certain threat to jobs and copyrights is perceived as a result of image generation with Generative AI, most respondents assume an increase in productivity and expect that the limits of Generative AI lie where particular human creativity is important.

C. Results on the Image Classification Test

Participants were asked to answer questions about six test images in the next section of the questionnaire. In the first step, they were asked to rate the image quality in relation to the previously discussed criteria on a five-point scale (1 = very poor, ..., 5 = very good). Participants could also select “no response”. Table III shows the average ratings across all participants and the overall quality as the average of the six criteria values. The first interesting observation is that the three ”real art” images, i.e., those not generated with AI, received the highest overall quality values.

As Table IV shows, most respondents classified only one image as real or not AI-generated. This is a photo-realistic depiction of a parrot, characterized by a high level of detail and color richness. Two images tagged as real art in the image set were incorrectly classified as AI-generated by the respondents. Overall, the classification is largely independent of the image quality, which supports the previous observation in this study

TABLE III. RESPONDENTS’ ASSESSMENT OF THE TEST IMAGES BY THE QUALITY CRITERIA

Image	1	2	3	4	5	6
Generation	AI	Real	Real	AI	Real	AI
Overall quality (average)	3.58	4.28	4.06	3.46	3.70	3.67
Detail and texture quality	3.75	4.47	4.14	3.38	3.66	3.72
Color harmony	3.80	4.47	4.22	3.60	3.77	3.81
Composition and structure	3.57	4.39	4.18	3.49	3.73	3.64
Creativity and originality	3.69	4.11	3.89	3.43	3.74	3.72
Emotional effect	3.33	4.13	3.98	3.45	3.64	3.57
Narrative perception	3.36	4.10	3.95	3.45	3.67	3.59

that the respondents predominantly assume that AI does not generate images of poorer quality. It is also interesting to note in Table IV that most respondents were rather or very confident in their classification decisions, i.e., no major deviations in decision confidence between the images reported.

TABLE IV. RESPONDENTS’ CLASSIFICATION AND CERTAINTY ON AI GENERATION OF TEST IMAGES

Image	1	2	3	4	5	6
Generation	AI	Real	Real	AI	Real	AI
AI	93.6%	31.4%	65.7%	87.2%	69.6%	85.4%
Real	6.4%	68.6%	34.3%	12.8%	30.4%	14.6%
Majority	AI	Real	AI	AI	AI	AI
Very uncertain	1.7%	0.6%	1.2%	1.7%	0.6%	3.5%
Rather uncertain	11.6%	16.9%	19.8%	12.8%	23.3%	15.1%
Rather certain	54.7%	52.3%	52.3%	45.3%	44.2%	43.0%
Very certain	32.0%	30.2%	26.7%	40.1%	32.0%	38.4%

In addition to evaluating the images according to the perceived quality, the participants were asked to rank the quality criteria based on their importance for classifying the respective images as real art or AI-generated. Table V shows the results of this assessment of the importance of the criteria for the various images. From the different levels of importance of the individual criteria in the classification decision on AI generation for the various images, it can be deduced that this ranking strongly depends on the motif. In the first image, composition, structure, creativity, and originality are the most important decision criteria. This fits in with the fact that in this image, a lion is shown in a rather unnatural pose in front of an incongruous background. In the second image, detail, texture quality, and color harmony are the most important criteria, which also fits the motif, as a photographic close-up of a colorful parrot is shown here. The different importance of the criteria and the resulting motif-dependent evaluation profiles are visualized in Figure 2 for Image 1 and 2.

The importance of the criteria thus provides important clues for image-related decision-making. However, the image quality in this respect does not systematically influence the categorization as AI-generated. In Table V for Image 3, for example, the criteria detail and texture quality (3.23) and composition and structure (3.19) are the most important evaluation criteria and were also rated relatively well (4.14, 4.18) in Table III. Nevertheless, Image 3 was classified as AI-generated by the majority of the participants. This can be explained by examining the participants’ free text comments reported in the survey data. The decision to classify the image as AI-

TABLE V. RESPONDENTS' ASSESSMENT OF THE IMPORTANCE OF QUALITY CRITERIA FOR IMAGE CLASSIFICATION

Image	1	2	3	4	5	6
Generation	AI	Real	Real	AI	Real	AI
Detail and texture quality	2.56	3.28	3.23	3.60	2.92	2.84
Color harmony	2.24	3.51	2.68	3.07	3.49	3.41
Composition and structure	3.16	3.01	3.19	3.12	3.07	2.97
Creativity and originality	3.09	2.04	2.28	2.13	2.42	2.62
Emotional effect	2.14	1.83	2.16	1.91	1.98	2.01
Narrative perception	1.80	1.56	1.78	1.42	1.55	1.62

generated was evaluated with statements, such as "exaggerated idealization", "looks very edited on the face", "the skin is too perfect", "the natural is missing", or "looks artificial". These ratings are presumably because although this image is a photo-realistic portrait of a woman, it is a real art, not a photograph. Thus, the classification as Generative AI seems less about the perceived quality and more about certain inconsistencies as deviations between expected (photo) and perceived (not a photo) image features, where deviations from the expectations are interpreted as indications of AI generation.

Table V also shows that technical characteristics of the image (detail and texture quality, color harmony, composition, and structure) play a more important role in classification, while perceptions in terms of creativity and originality, emotional effect, and narrative perception are of lesser importance. A reason why primarily technical criteria were used in the image quality evaluation may also be because the participants were unaware of the task and the background of the creation of the pictures. For example, whether an original pose or a realistic depiction was required or the picture idea was not described. Future studies, therefore, should investigate further how the implementation of an image idea is perceived in images created with AI (prompt engineering) and without AI (traditional digital image creation and editing).

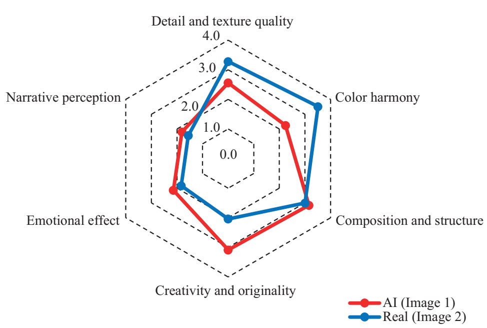

Figure 2. Importance Comparison of Quality Criteria Profile for Image 1 and 2.

D. Participant Characteristics and Image Classification

In the final step of the survey data analysis, several correlations were examined between participant characteristics and the image classification task. As part of this pre-study, simple correlation analyses (due to the partly ordinally scaled variables using Spearman correlation) and significant tests were carried out. Table VI shows the corresponding correlation between selected experience data with digital image editing during education (ExpEdu) and in the work environment (ExpWork), experience with AI tools (AITool), as well as the agreement values for the statements on substitution of jobs by AI (JobLoss), the increase in productivity (ProdImp), the limited potential of AI to replace human creativity (CreatLim), and the ease of differentiation (EasyDiff) of AI and non-AI images which were previously discussed in Section IV-B.

TABLE VI. CORRELATION MATRIX FOR SELECTED EXPERIENCE AND STATEMENTS ON IA IMPACT

	ExpEdu	ExpWork	AITool	JobLoss	ProdImp	CreatLim	EasyDiff
ExpEdu	–						
ExpWork	0.757**	–					
AITool	0.691**	0.642**	–				
JobLoss	0.075	0.026	-0.06	–			
ProdImp	0.230**	0.298**	0.286**	0.028	–		
CreatLim	0.141	0.157*	0.075	0.193*	0.190*	–	
EasyDiff	0.371**	0.396**	0.441**	0.037	0.275**	0.302**	–

Correlation is significant at the * 0.05/** 0.01 level (2-tailed).

Significant strong correlations can be found between the intensity of the use of AI tools and experience with digital image processing in education and the work environment. The significant weak correlation between expectations of increased productivity and the corresponding experience with digital image editing and AI tools is interesting and plausible. The significant but very weak correlation between the assessment that Generative AI will lead to job losses and the agreement with the statement that AI cannot replace human creativity is unexpected and remarkable. The coincidence of these contradictory statements in the participants' opinions could indicate that the two statements tended to be supported by people with a rather negative or skeptical attitude toward AI technology.

The level of agreement with the limited creativity of Generative AI also correlates very weakly with the extent of the participants' work experience and their agreement with the impact of Generative AI on their working environment. It is interesting to note that the assessment of the ease of distinguishing AI-generated images correlates with almost all other experience and agreement values. The assessment of differentiability is most strongly influenced by the intensity of usage of AI tools. This is plausible, as participants who regularly and frequently use AI tools are expected to be best able to assess the possibilities and results.

The following will examine the influences of the participant characteristics on the test persons' classification results of the pictures. Figure 3 shows the frequency of the number of correct classifications by the participants. On average, 3.99 images were correctly classified by the subjects as AI-generated or not AI-generated. The distribution in the figure indicates that most probably random differences rather than systematic differences are responsible for the differences in the precision of the classification decision.

This assumption is strengthened when the results of the correlation analysis in Table VII are considered. In addition to the variables of the study described above, the experience with

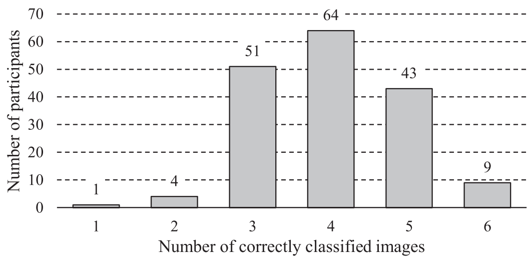

Figure 3. Frequency of Correct Image Classifications.

digital image processing in the private sector (ExpPriv), the total duration of the experience (ExpDur), and the number of correctly classified images (CorrClas) are also listed here. First of all, it can be seen that all experience-related characteristics correlate significantly and considerably with each other. Unexpectedly, however, there is a significant but very weak negative correlation between the number of correctly classified images and experience with digital image processing in training and education, as well as the intensity of the use of AI tools.

This could be explained by the fact that participants with extensive experience also know that very high-quality results can now be achieved with Generative AI and, therefore, considered AI generation to be possible for almost all of the images inspected. The results are nevertheless surprising and indicate that even with extensive experience in digital image processing, it was not possible to classify the test images presented systematically.

TABLE VII. CORRELATION MATRIX FOR EXPERIENCE AND CORRECT IMAGE CLASSIFICATION

	ExpPriv	ExpEdu	ExpWork	ExpDur	AITool	CorrClas
ExpPriv	–					
ExpEdu	0.653**	–				
ExpWork	0.634**	0.757**	–			
ExpDur	0.536**	0.324**	0.460**	–		
AITool	0.589**	0.691**	0.642**	0.285**	–	
CorrClas	-0.138	-0.163*	-0.136	-0.074	-0.180*	–

Correlation is significant at the * 0.05/** 0.01 level (2-tailed).

In a last consideration, the correlations between the agreement values for the statements of the AI impact and the correct image classification are shown in Table VIII. This table also lists the variables for the agreement values on the statement of a threat to intellectual property (IPThreat) and the poorer quality of AI-generated images (PoorQual). There are also no clearly interpretable relationships here, even though some correlations may indicate a certain basic attitude toward AI. There is a significant, moderate correlation between the perceived risk for Intellectual Property (IP) and the substitution of jobs. The perception of poor quality of AI-generated image material correlates very weakly with job substitution and (weakly) with IP risk and irreplaceable human creativity. However, there is only a significant, very weak positive correlation with the perceived IP risk regarding the number of correctly classified images. This could indicate that those participants who have dealt more extensively with the Generative AI procedures and understand the problem of reuse of design patterns by AI (that relates to IP problems) were also able to achieve slightly better classification results.

TABLE VIII. CORRELATION MATRIX FOR STATEMENTS ON IA IMPACT AND IMAGE CLASSIFICATION

	JobLoss	ProdImp	IPThreat	CreatLim	PoorQual	EasyDiff	CorrClas
JobLoss	–						
ProdImp	0.028	–					
IPThreat	0.404**	-0.186*	–				
CreatLim	0.193*	0.190*	0.168*	–			
PoorQual	0.178*	-0.100	0.330**	0.356**	–		
EasyDiff	0.037	0.275**	0.086	0.302**	0.325**	–	
CorrClas	0.053	-0.045	0.157*	0.03	0.033	-0.075	–

Correlation is significant at the * 0.05/** 0.01 level (2-tailed).

As a result, it can be stated that there is no statistical evidence within the scope of the study that certain experience with digital image editing or a high-intensity use of AI tools systematically improves the ability to correctly assign the AI-generated images in the set of images presented. There are also no clear correlations between certain attitudes towards the impact of AI and the classification result. However, it should be noted that this may be due to the composition of the small sample or the selection of motifs, and therefore, the results of this pre-study show a tendency but cannot yet be generalized.

V. Conclusions

This preliminary study has provided important findings on the reception of image-generating Generative AI in the German media sector. The following results can be summarized concerning the research questions formulated at the beginning:

- *Use and awareness of Generative AI:* Less than two years after the launch of ChatGPT, about one-half of the participants in this sample from the German media sector are familiar with Generative AI tools for digital image creation and editing. The most common AI tools, such as DALL-E or Adobe Firefly, are already used almost daily by nearly one in ten of those surveyed.
- *Impact of Generative AI:* Around a third of the participants have not yet formed a clear opinion on the effects of Generative AI on their working environment. However, for those with an opinion, the majority agrees with the statements that AI increases productivity and relieves the burden of routine tasks but cannot replace human creativity. The performance of Generative AI is already perceived as very high: The participants are almost undecided as to whether AI-generated images are still distinguishable. Only a minority within the sample perceives AI-generated images as characterized by poorer quality.
- *Quality and recognition of AI-generated images:* While the non-AI-generated images are assigned a slightly higher quality, only one out of three non-AI-generated real artwork images is recognized correctly by the majority of participants. The test shows that quality is not used to draw conclusions about AI generation, and no specific quality criteria are important for the classification decision. Rather, the importance of these criteria varies depending on the motif. Overall, the participants seem to pay more attention to

inconsistencies in the composition of images when identifying AI-generated images.
- *Factors influencing image classification:* The correct recognition of images in the image set presented cannot be explained systematically by the experience or attitudes of the participants, although there was a tendency to overestimate the proportion of AI-generated images. However, it is interesting to note that the correlations examined reveal some relationships between more skeptical or positive attitudes toward the impact of AI.

Based on these results, it should be noted that there is still a great openness towards using Generative AI. However, there are already skeptical perspectives on its use, which could increase if negative expectations prove true. For example, fears regarding the risks of copyrighting an IP threat must be effectively countered. It is difficult and will certainly become even more difficult to distinguish AI-based images from the creative work of humans by the end product. Thus, it can be expected that the human element in creative collaboration with AI and the added value of a human expert must be explained and emphasized more to customers in future media productions.

VI. Limitations and Outlook

The results of this pre-study are based on a sample obtained via a panel. The users received an incentive for their participation. Although participants from the media sector were specifically contacted for participation, there were no filter questions or quotas to obtain a representative sample for the media sector in Germany. Against this background, the results can only be generalized to a limited extent. The test is also subject to several limitations. With only six images presented, the participants were exposed to a very small test set. The choice of motifs may also have influenced the results, as the selection was not purely random but rather pairs of similar compositions of AI-generated and non-AI-generated images.

However, based on this study's results, whether larger and more representative samples or more comprehensive and randomly selected image tests could generate more meaningful findings is questionable. The study results indicate that with the current state of image generation with Generative AI, even experts are often unable to make a reliable decision about the type of image generation based on the images produced or their quality. Rather, subsequent studies should focus on the image generation process. Therefore, future studies should consider the underlying goals or idea of image generation and let participants evaluate the resulting images in relation to the image idea. In addition to a binary setup (with and without AI), it could be interesting to investigate how collaboration between humans and AI affects the production process and the results. The design of such human-AI collaboration processes in the media and creative sector appears to be an important field of research that has remained largely unexplored.

Appendix

The following information specifies the images from the Kaggle dataset "AI-Generated Images vs Real Images" [38] used in this study:

- Image 1: AI-generated,
filename: 41b6d9592db18a15b1e32dfd50.jpg.
- Image 2: Real,
filename: shouts-animals-watch-baby-hemingway.jpg.
- Image 3: Real,
filename: portrait075a-819x1024.jpg.
- Image 4: AI-generated,
filename: 52520977911_33437880be_z.jpg.
- Image 5: Real, filename:
taxture-scenery-poster-500x500.jpg.
- Image 6: AI-generated, filename:
clgjlgjec001a08k0bhi51i88.jpg.

References

[1] S.-C. Huang and T.-H. Le, "Generative adversarial network," in *Principles and Labs for Deep Learning*, S.-C. Huang and T.-H. Le, Eds., Elsevier, 2021, pp. 255–281, ISBN: 9780323901987. DOI: 10.1016/b978-0-323-90198-7.00011-2.

[2] Y. Wang, Y. Pan, M. Yan, Z. Su, and T. H. Luan, "A Survey on ChatGPT: AI–Generated Contents, Challenges, and Solutions," *IEEE Open Journal of the Computer Society*, vol. 4, pp. 280–302, 2023. DOI: 10.1109/OJCS.2023.3300321.

[3] H. Dong and S. Xie, *Large Language Models (LLMs): Deployment, Tokenomics and Sustainability*, May 27, 2024. [Online]. Available: http://arxiv.org/pdf/2405.17147v1 [retrieved: 05/31/2024].

[4] T. B. Brown *et al.*, *Language Models are Few-Shot Learners*, 2020. DOI: 10.48550/arxiv.2005.14165.

[5] M. Ghassemi *et al.*, "ChatGPT one year on: who is using it, how and why?" *Nature*, vol. 624, no. 7990, pp. 39–41, 2023. DOI: 10.1038/d41586-023-03798-6.

[6] G. Eysenbach, "The Role of ChatGPT, Generative Language Models, and Artificial Intelligence in Medical Education: A Conversation With ChatGPT and a Call for Papers," *JMIR Medical Education*, vol. 9, e46885, 2023, ISSN: 2369-3762. DOI: 10.2196/46885. [Online]. Available: https://mededu.jmir.org/2023/1/e46885/ [retrieved: 05/31/2024].

[7] B. Khoo, R. C.-W. Phan, and C.-H. Lim, "Deepfake attribution: On the source identification of artificially generated images," *WIREs Data Mining and Knowledge Discovery*, vol. 12, no. 3, e1438, 2022, ISSN: 1942-4787. DOI: 10.1002/widm.1438. [Online]. Available: https://wires.onlinelibrary.wiley.com/doi/10.1002/widm.1438 [retrieved: 05/31/2024].

[8] Y. Ju, S. Jia, L. Ke, H. Xue, K. Nagano, and S. Lyu, "Fusing Global and Local Features for Generalized AI-Synthesized Image Detection," in *2022 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2022, pp. 3465–3469, ISBN: 978-1-6654-9620-9. DOI: 10.1109/icip46576.2022.9897820.

[9] C. Becker and R. Laycock, *Embracing Deepfakes and AI-generated images in Neuroscience Research*. 2023. DOI: 10.22541/au.168122346.61187955/v2.

[10] R. Chamberlain, C. Mullin, B. Scheerlinck, and J. Wagemans, "Putting the art in artificial: Aesthetic responses to computer-generated art," *Psychology of Aesthetics, Creativity, and the Arts*, vol. 12, no. 2, pp. 177–192, 2018, ISSN: 1931-3896. DOI: 10.1037/aca0000136.

[11] H. Gangadharbatla, "The Role of AI Attribution Knowledge in the Evaluation of Artwork," *Empirical Studies of the Arts*, vol. 40, no. 2, pp. 125–142, 2022, ISSN: 0276-2374. DOI: 10.1177/0276237421994697.

[12] S. Natale and L. Henrickson, "The Lovelace effect: Perceptions of creativity in machines," *New Media & Society*, vol. 26, no. 4, pp. 1909–1926, 2024, ISSN: 1461-4448. DOI: 10.1177/14614448221077278.

[13] D. B. Shank, C. Stefanik, C. Stuhlsatz, K. Kacirek, and A. M. Belfi, "AI composer bias: Listeners like music less when they think it was composed by an AI," *Journal of Experimental Psychology: Applied*, vol. 29, no. 3, pp. 676–692, 2023. DOI: 10.1037/xap0000447.

[14] L. Bellaiche *et al.*, "Humans versus AI: whether and why we prefer human-created compared to AI-created artwork,"

Cognitive Research: Principles and Implications, vol. 8, no. 1, p. 42, 2023. DOI: 10.1186/s41235-023-00499-6.

[15] E. Cetinic and J. She, *Understanding and Creating Art with AI: Review and Outlook*, Feb. 18, 2021. [Online]. Available: http://arxiv.org/pdf/2102.09109v1 [retrieved: 05/31/2024].

[16] N. Nishida *et al.*, "Artificial intelligence (AI) models for the ultrasonographic diagnosis of liver tumors and comparison of diagnostic accuracies between AI and human experts," *Journal of Gastroenterology*, vol. 57, no. 4, pp. 309–321, 2022. DOI: 10.1007/s00535-022-01849-9.

[17] Z. Epstein *et al.*, "Art and the science of generative AI: A deeper dive," *Science*, vol. 380, no. 6650, pp. 1110–1111, 2023, ISSN: 0036-8075. DOI: 10.1126/science.adh4451. [Online]. Available: http://arxiv.org/pdf/2306.04141v1 [retrieved: 05/31/2024].

[18] M. Mirbabaie, F. Brünker, N. R. J. Möllmann Frick, and S. Stieglitz, "The rise of artificial intelligence – understanding the AI identity threat at the workplace," *Electronic Markets*, vol. 32, no. 1, pp. 73–99, 2022, ISSN: 1019-6781. DOI: 10.1007/s12525-021-00496-x.

[19] M. Xia, "Co-working with AI is a Double-sword in Technostress? An Integrative Review of Human-AI Collaboration from a Holistic Process of Technostress," *SHS Web of Conferences*, vol. 155, p. 03 022, 2023. DOI: 10.1051/shsconf/202315503022.

[20] D. Czarnitzki, G. P. Fernández, and C. Rammer, "Artificial Intelligence and Firm-Level Productivity," *SSRN Electronic Journal*, 2022. DOI: 10.2139/ssrn.4049824.

[21] F. Martin-Rodriguez, R. Garcia-Mojon, and M. Fernandez-Barciela, "Detection of AI-Created Images Using Pixel-Wise Feature Extraction and Convolutional Neural Networks," *Sensors*, vol. 23, no. 22, 2023. DOI: 10.3390/s23229037.

[22] A. G. Moskowitz, T. Gaona, and J. Peterson, *Detecting AI-Generated Images via CLIP*, 2024. DOI: 10.48550/arXiv.2404.08788.

[23] S. S. Baraheem and T. V. Nguyen, "AI vs. AI: Can AI Detect AI-Generated Images?" *Journal of Imaging*, vol. 9, no. 10, 2023. DOI: 10.3390/jimaging9100199.

[24] S. D. Bray, S. D. Johnson, and B. Kleinberg, "Testing human ability to detect 'deepfake' images of human faces," *Journal of Cybersecurity*, vol. 9, no. 1, 2023, ISSN: 2057-2085. DOI: 10.1093/cybsec/tyad011.

[25] Z. Liu, X. Qi, and P. Torr, *Global Texture Enhancement for Fake Face Detection in the Wild*, Feb. 1, 2020. [Online]. Available: http://arxiv.org/pdf/2002.00133v3 [retrieved: 05/31/2024].

[26] S. J. Nightingale and H. Farid, "AI-synthesized faces are indistinguishable from real faces and more trustworthy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 119, no. 8, 2022. DOI: 10.1073/pnas.2120481119.

[27] B. Shen, B. RichardWebster, A. O'Toole, K. Bowyer, and W. J. Scheirer, *A Study of the Human Perception of Synthetic Faces*, Aug. 11, 2021. [Online]. Available: http://arxiv.org/pdf/2111.04230v1 [retrieved: 05/31/2024].

[28] Y. Zhou and H. Kawabata, "Eyes can tell: Assessment of implicit attitudes toward AI art," *i-Perception*, vol. 14, no. 5, 2023. DOI: 10.1177/20416695231209846.

[29] L. Gu and Y. Li, "Who made the paintings: Artists or artificial intelligence? The effects of identity on liking and purchase intention," *Frontiers in Psychology*, vol. 13, p. 941 163, 2022, ISSN: 1664-1078. DOI: 10.3389/fpsyg.2022.941163.

[30] Y. Lyu, X. Wang, R. Lin, and J. Wu, "Communication in Human–AI Co-Creation: Perceptual Analysis of Paintings Generated by Text-to-Image System," *Applied Sciences*, vol. 12, no. 22, p. 11 312, 2022. DOI: 10.3390/app122211312.

[31] C. B. Horton, M. W. White, and S. S. Iyengar, *Will AI Art Devalue Human Creativity?* 2023. DOI: 10.21203/rs.3.rs-2987022/v1.

[32] P. Fortuna, A. Modliński, and M. McNeill, "Creators Matter. Perception and Pricing of Art Made by Human, Cyborgs and Humanoid Robots," *Empirical Studies of the Arts*, vol. 41, no. 2, pp. 331–351, 2023, ISSN: 0276-2374. DOI: 10.1177/02762374221143717.

[33] S. C. Y. Ho, "From Development to Dissemination: Social and Ethical Issues with Text-to-Image AI-Generated Art," *Proceedings of the Canadian Conference on Artificial Intelligence*, 2023. DOI: 10.21428/594757db.acad9d77.

[34] K. Rasrichai, T. Chantarutai, and C. Kerdvibulvech, "Recent Roles of Artificial Intelligence Artists in Art Circulation," *Digital Society*, vol. 2, no. 2, 2023, ISSN: 2731-4650. DOI: 10.1007/s44206-023-00044-4.

[35] H. Leder, B. Belke, A. Oeberst, and D. Augustin, "A model of aesthetic appreciation and aesthetic judgments," *British Journal of Psychology*, vol. 95, no. Pt 4, pp. 489–508, 2004, ISSN: 0007-1269. DOI: 10.1348/0007126042369811.

[36] L. K. M. Graf and J. R. Landwehr, "A dual-process perspective on fluency-based aesthetics: the pleasure-interest model of aesthetic liking," *Personality and Social Psychology Review*, vol. 19, no. 4, pp. 395–410, 2015. DOI: 10.1177/1088868315574978.

[37] Unipark, *Online survey software: Surveys made easy with Unipark*, 2024. [Online]. Available: https://www.unipark.com/en/ [retrieved: 05/31/2024].

[38] C. Bowman, *Kaggle dataset: AI Generated Images vs Real Images: Web scraped images: AI and Real. Can you tell the difference?* 2024. [Online]. Available: https://www.kaggle.com/datasets/cashbowman/ai-generated-images-vs-real-images [retrieved: 05/31/2024].

[39] M. Benning, *KI-Bilder erstellen: Top 12 Bild-Generatoren*, Feb. 23, 2024. [Online]. Available: https://mind-force.de/marketing/ki-bilder-erstellen-bild-generatoren/ [retrieved: 05/31/2024].