

Can We Tell the Difference? Human Sensitivity to AI-Generated vs. Human-Made Portraits

Mats Nissen

s5214920

Department of Psychology, University of Groningen

PSBE3E-BT15.2024-2025.3: Bachelor Thesis

Group number: 43

Supervisor: Ben Gützkow

Second evaluator: Niklas Neumann

In collaboration with: Linde Van Egmond, Laura Hordijk, Filip Fukas, Jaas Maaskant, Robin Fischer

Month 06, 2025

Abstract

The increasing sophistication of artificial intelligence (AI) presents new challenges in how humans perceive and evaluate AI-generated content. This study investigated whether individuals can accurately identify AI-generated versus human-made portrait paintings and examined factors that influence this judgment accuracy. Participants ($N = 232$) were randomly assigned to a training or control group and completed a task on classifying portraits as either human-made or AI-generated. The training manipulation consisted of interleaved practice to facilitate inductive learning. Along with this, questionnaires were employed to measure Theory of Mind (ToM) abilities and Assessment orientation as part of the Regulatory Mode Theory. Results showed that participants in the training condition performed better than those in the control group, indicating that judgment accuracy is malleable and can be improved through the right training intervention. ToM scores predicted better judgment accuracy specifically for human-made portraits, suggesting that social cognitive abilities are relevant when interpreting socially rich stimuli. However, no interaction was found between ToM and training, nor was there a significant effect of Assessment on training benefit. These findings highlight both the potential of targeted interventions to enhance judgment accuracy and the specific role of ToM in recognizing human intention in visual art. However, the practical application of these findings is limited, highlighting that further research is needed to improve our understanding of how humans evaluate AI-generated works.

Keywords: Artificial intelligence, inductive learning, decision making, Theory of Mind, Assessment orientation

Can We Tell the Difference? Human Sensitivity to AI-Generated vs. Human-Made Portraits

Back in October 2018, Christie's auction house sold a painting by the French art collective "Obvious" for \$432,500 (Cohn, 2018). Special about this painting was that it was the first time that a painting generated by artificial intelligence (AI) was sold at an auction house (Cohn, 2018). What is more, it even exceeded the estimates of the prize by a large margin, with the original estimates being in the range of \$7,000 to \$10,000 (Cohn, 2018). Subsequently, protesting voices among the art community have surfaced, stating that the work of artists is exploited by the AI companies for the training of their algorithms (Veltman, 2025). With the rise of AI programs capable of producing compelling artworks, a philosophical debate has emerged around the source of artistic value and whether AI-generated art can be appreciated in the same way as human-made creations. In the case of the painting sold at Christie's, transparency about the creation process was given (Cohn, 2018). However, this level of openness may not be guaranteed in future cases, raising legitimate concerns. What are the implications for the art world if AI is used to replicate the creative efforts of human artists without disclosure? And are we already at the point where AI-generated works are indistinguishable from those made by humans? These questions form the basis of the current investigation. The relevance and complexity of this issue, coupled with increasing public interest, underscore the need for empirical research. As AI-generated art is a relatively new phenomenon, the existing body of literature is still limited. Consequently, this study takes an exploratory approach aimed at identifying the psychological factors that influence individuals' ability to distinguish between AI-generated and human-made artworks.

As AI technology continues to evolve, limitations in its current design make it possible to detect and differentiate its output from that of humans. An AI program like Midjourney works with a so-called diffusion model that involves training with huge datasets (Bergmann & Stryker, 2024). A large quantity of images paired with text descriptions is fed to the program in order for it to statistically learn associations between meaning and graphical features like shapes, colors, or composition (Bergmann & Stryker, 2024). During training, the real images are then gradually manipulated with small elements of random noise (Bergmann & Stryker, 2024). "Random noise" in this context can best be

compared to the fuzzy static of an old TV that is just random pixelation. The program then learns how to reverse this noise back to the original (Bergmann & Stryker, 2024). This learning process in turn enables the program to start from noise exclusively when presented with a text prompt, essentially starting from scratch to create the final picture (Bergmann & Stryker, 2024). However, since this process is not free of flaws, the task for the future will be to find ways to effectively sensitize people to these irregularities to identify AI-generated work. A plausible option would be to enhance people's ability to detect these inconsistencies through targeted exposure or training.

People frequently acquire knowledge about categories, patterns, or concepts simply by encountering repeated examples in everyday life. This type of learning, known as inductive learning, allows individuals to form generalizations based on observed instances and is central to how humans naturally categorize and navigate the world (Bruner, Goodnow, & Austin, 1956). Research by Kang and Pashler (2011) provides compelling evidence for how the type of exposure influences the effectiveness of inductive learning in the context of artworks. In their study, participants learned to identify different painters' styles more successfully when paintings were shown in an interleaved instead of massed fashion. This suggests that interleaved practice, through mere exposure to contrasting examples, enhances people's ability to notice subtle differences, making it a particularly effective strategy for inductive learning. Inductive learning is typically implicit and passive, relying on accumulated experience to recognize regularities or differences across stimuli (Ashby & Maddox, 2005). Facial recognition illustrates this very well, as individuals are exposed to countless faces, gradually acquiring the ability to distinguish between them based on subtle differences in configuration and expression (Valentin et al., 1994). Eventually, humans become remarkably adept at recognizing subtle variations in facial features, enabling them to reliably identify individual faces with high accuracy (Bruce & Young, 1986). So, especially for the artistic depiction of humans, people could potentially be sensitized very effectively to differences in the creation process. Accordingly, one focus of this project is to identify psychological factors that contribute to more accurate discrimination of facial features in AI-generated and human-made portraits.

In addition to the effects of simple exposure, differences in social cognition may be predictive of someone's ability to distinguish AI-generated from human-made portraits, given that portraits intentionally depict humans and are thus rich in social content. Theory of Mind (ToM) is a cognitive ability that normally develops around the ages of four to five and involves understanding that other people have their own thoughts and perspectives that can differ from one's own (Navarro, 2022; Westby & Robinson, 2014). It helps to predict the behavior of others based on inferred mental states and is ultimately important for social interactions and empathy (Navarro, 2022). Research has highlighted the role of ToM in interpreting visual cues that convey social information. For instance, the ability to follow and respond to gaze cues, an important social signal, improves with higher ToM abilities, as it involves not just recognizing where someone is looking but also understanding the underlying mental state behind the gaze (Frischen, Bayliss, & Tipper, 2007). This heightened ability to interpret gaze direction allows individuals with stronger ToM abilities to make more accurate judgments in social scenarios (Bayliss & Tipper, 2005). In this way, ToM enables improved social cognition but also visual attention and the ability to interpret subtle cues within visual scenes (Schilbach et al., 2006). A property that might ultimately translate to spotting inconsistencies between AI-generated and human-made art.

Even when individuals have traits that support accurate judgments, their effectiveness may depend on the degree to which a person is motivated to carefully evaluate the problem at hand. Assessment as part of the Regulatory Mode Theory (Kruglanski et al., 2000) is a stable motivational orientation that involves a reflective process within self-regulation. Individuals with high Assessment orientation want to carefully compare different possibilities, such as strategies or objectives, by considering their pros and cons (Kruglanski et al., 2000). This comparison helps determine which option is more effective or appropriate in a given context (Kruglanski et al., 2000). Research indicates that high Assessment orientation promotes careful decision-making, leading to greater accuracy, albeit often requiring more time to complete a task (Mauro et al., 2009). Accordingly, the effect of Assessment could also extend to the interpretation of portraits and influence how many details a person considers to accurately evaluate the paintings. While our initial aim is to identify psychological traits linked to overall

performance in detecting flaws in AI-generated artworks, the role of Assessment is based on the idea that individuals high in this trait are more motivated to scrutinize and evaluate every potential imperfection.

Present Study

With this study, we aim to find out how well people can distinguish between AI-generated and human-made art and what factors influence this judgment accuracy. To examine this, we will implement a training condition to test whether inductive learning through interleaved practice improves the ability to recognize subtle inconsistencies between the two types of images. Investigating individual differences in psychological constructs will allow us to further explore how this ability is affected. Our focus will be specifically on portraits, as they contain socially rich content that potentially interacts with particular cognitive processes.

The effectiveness of interleaved practice has already been established by previous research (Kang & Pashler, 2011). We therefore firstly hypothesize that (H1): people who have been exposed to examples of AI-generated and human-made portraits are better at distinguishing between these two categories of stimuli compared to people who have no prior exposure to similar examples.

Recognizing subtle differences in the portraits likely draws on mechanisms related to understanding others' mental states. We therefore propose that ToM, as a component of social cognition, influences judgment accuracy, leading to our second hypothesis, that (H2): people with higher ToM abilities are better at distinguishing between AI-generated and human-made portraits. Furthermore, considering the potential contributions that both ToM and inductive learning might have on judgment accuracy, we thirdly hypothesize that (H3): the effect of stimulus exposure in H1 is moderated by ToM, leading to improved learning effects for individuals who have higher ToM abilities.

Given that precise consideration in decision making can be expected of someone who is measured high in Assessment orientation, we think that this has the potential to improve learning effects as well. We therefore come to our fourth hypothesis, that (H4): the effect of stimulus exposure in H1 is moderated by Assessment, resulting in better learning effects for individuals who are measured high in Assessment orientation.

Method

Participants

The sample in this study consisted of $N=232$ participants who completed the study. The sample was assembled through convenience sampling with $n=46$ participants coming from the authors' social network who voluntarily took part. The remaining $n=186$ participants were recruited through the SONA system, which research institutions employ to manage study participation. Students registered through SONA participated in exchange for course credits as part of their academic program. Additionally, five participants were excluded from the analysis for suspected negligence, finishing the study in under seven minutes. Two more cases were kept out due to missing score values. Prior to data collection, the study underwent an ethical review and was approved by the Ethics Committee of the Faculty of Behavioural and Social Sciences at the University of Groningen.

Procedure and Measures

The study was conducted as an experiment, hosted online on Qualtrics. Participants were instructed on the purpose of the study, how data would be collected and handled, and the rights of participants. Based on that information, participants could give informed consent, including confirming that they were at least 16 years old. As part of Qualtrics systems, information was also collected about the participants' browsers, namely the browser type and version, operating system, and screen resolution. The study consists of two parts, a questionnaire part and an experimental part using an inductive learning paradigm. First, participants were asked to fill out a number of scales and items.

Theory of Mind

To test the participant's ToM abilities, a shortened version of the Reading the Mind in the Eyes Test (RMET) was applied (Chander et al., 2020). This test especially lends itself to the current research context, since this task measures the ability to recognize complex emotional and mental states by presenting photographs of the eye region of faces. Participants are asked to select which of four descriptive words best matches the mental state conveyed by the eyes (eg. doubtful, affectionate, playful, aghast). The RMET is widely used to detect subtle differences in social-cognitive functioning (Baron-

Cohen et al., 2001). Chander et al. (2020) developed a shorter version of this test by applying statistical methods to identify and exclude less effective test items. Due to practical constraints, the current study employed a more concise 8-item version. While this shorter format may come with some reduction in measurement precision, it was assumed that the impact on the overall results would be minimal.

Assessment

For the measurement of Assessment, nine items from the Assessment scale of the Regulatory Mode Questionnaire (Kruglanski et al., 2000) were utilized. The full Assessment scale features twelve items, but three were dropped as they pertain to social interaction, which was deemed not relevant for the purpose of this study. Items on this scale are measured using a six-point scale (1 = *strongly disagree*, 6 = *strongly agree*). Examples of these items include: “I like evaluating other people's plans.” and “I often critique work done by myself or others.”. All items of the Assessment scale used in this study are listed in Appendix B. The following sections present additional measures that, while not included in this particular analysis, were part of the overall study procedure.

Cognitive and Affective Empathy

To measure cognitive and affective empathy, four items from the Questionnaire of Cognitive and Affective Empathy (Reniers et al., 2011) were used. The four items are: “I often get deeply involved with the feelings of a character in a film, play, or novel.”, “I can tell if someone is masking their true emotion”, “I get very upset when I see someone cry.”, “I find it easy to put myself in somebody else’s shoes.” This questionnaire, and therefore these four items, uses a four-point Likert scale (1= *strongly disagree*, 4 = *strongly agree*).

Cognitive Reflection

Items from the Cognitive Reflection Test (Frederick, 2005) were used to measure preference for analytical or intuitive cognitive style. Scoring low on this scale, reflected by relying on the first reaction that emerges, would indicate a preference for an intuitive style and vice versa. This test consists of three mathematical items, each requiring a number as an answer. The items were stated as follows: “If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? A:

____ minutes,” and ”In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? A: ____ days”. For the third item, we used a reformulation of De Neys et al. (2013) due to a high probability of possible familiarity with the item. This item is worded as follows: “A magazine and a banana cost €2.90 in total. The magazine costs €2.00 more than the banana. How much does the banana cost? A: ____ cents”.

Experimental setup

With facial recognition being a very sophisticated process that enables humans to differentiate between subtle differences in facial features, we assume that investigating paintings rich in facial information could produce the most meaningful effect in our study. The AI portraits were created using the program Midjourney. By employing carefully formulated, detailed prompts, portrait images could be generated. For example; “*A woman standing in the rain, face tilted upward, eyes closed, conveying relief and sadness simultaneously, muted blues and grays, impressionistic brushwork.*”. The full list of prompts is provided in Appendix C. The human-made portraits were collected from websites with an open-access policy. These platforms provide visual materials that are freely accessible and reusable for research and educational purposes. A complete list of the paintings made by humans is provided in Appendix D. After completion of the questionnaire, participants were randomly assigned to the experimental condition or the control condition. The experimental condition featured the inductive training paradigm, whereas the control condition received no training. Participants in the training condition received a set of 74 portraits through interleaved practice to familiarize participants with the variation within the dataset and to enhance their evaluative skills. The control group received 36 portraits, but without any prior training or practice. This allowed for the evaluation of the effect of visual training on judgment accuracy. The portraits given to the participants were equally distributed between AI-generated and human-made portraits. Additionally, a proportional distribution based on the year of creation was chosen. In this way, a balanced dataset was compiled.

Results

Preliminary Analysis

The random assignment to the experimental conditions resulted in two groups of comparable sizes ($N = 109$ for the training group, $N = 123$ for the control group). Descriptive statistics regarding the relevant variables, as well as correlations, are listed in Table 1 of Appendix A. Judgment accuracy was named “percent correct” and was calculated as the percentage of correctly classified artworks across all trials for each participant. Participants correctly identified an average of 60.31% ($SD = 11.00$) of the artworks as either AI-generated or human-made. While the experimental condition showed an average of 64.84% ($SD = 10.79$) correct identifications, the control group averaged out at 56.29% ($SD = 9.56$) correct identifications. Results from the correlational analysis confirm a significant positive association between condition and percent correct, $r(230) = .389, p < .001$. Across the full sample, participants correctly identified 59.7% ($SD = 15.00$) of the AI-generated artworks and 60.11% ($SD = 13.63$) of the human-made artworks, indicating a comparable level of performance across both stimulus types.

Performance on the RMET averaged at 5.71 correct responses ($SD = 1.67$), while Assessment scores averaged 4.19 ($SD = 0.60$). RMET scores were significantly associated with percent correct, $r(230) = .243, p < .001$, indicating that participants with better performances in deducing emotional states from the eyes tend to have better judgment accuracy. Furthermore, results suggest better judgment accuracy when a participant scores high in Assessment orientation, indicated by the significant positive correlation between Assessment scores and percent correct, $r(230) = .152, p = .021$.

Main analysis

In order to investigate the hypotheses, the data was analyzed with a linear regression model. We constructed a model in which percent correct was regressed on the condition (dummy-coded). This initial model for the investigation of the main effect had a significant result, $R^2 = .15, F(1, 230) = 40.97, p < .001$, showing that the condition alone can explain a significant amount of variance in the judgment accuracy levels, supporting H1. The full coefficient table for this model can be seen in Table 2.

For the second hypothesis, we extended the model by including the centered RMET scores as a second predictor and the interaction between condition and RMET scores as a third predictor. Overall the model increased in significance, $R^2 = .2$, $F(3, 228) = 18.75$, $p < .001$, explaining significantly more variance than condition alone. The condition is, however, still the most significant predictor in this model, $B = 8.19$, $SE = 1.31$, $t(228) = 6.26$, $p < .001$, 95% CI [5.61, 10.77], showing that the training still contributes the most to explain variance in judgment accuracy. The centered RMET scores could also significantly contribute to the predictive power of the model, $B = 1.24$, $SE = 0.55$, $t(228) = 2.26$, $p = .025$, 95% CI [0.16, 2.32], giving support to H2. The interaction term, on the other hand, proved to be non-significant, $B = 0.36$, $SE = 0.78$, $t(228) = 0.46$, $p = .648$, 95% CI [-1.18, 1.9], and therefore H3 is not supported. The complete output of coefficients and corresponding test statistics can be seen in Table 3. An illustration of the interaction can be seen in Figure 1.

The fourth hypothesis has been investigated in the same manner by including the centered Assessment scores as a second predictor and the interaction between condition and Assessment scores as the third predictor. Compared to the initial model, including only condition as a predictor, the overall significance of the model only slightly increased, $R^2 = .17$, $F(3, 228) = 15.9$, $p < .001$, indicating only a modest improvement of predictive power. This first impression is confirmed when looking at the individual predictors, since only the condition was significant, $B = 8.35$, $SE = 1.33$, $t(228) = 6.3$, $p < .001$, 95% CI [5.74, 10.97], while the other terms were not, including the interaction effect, $B = 0.29$, $SE = 0.25$, $t(228) = 1.16$, $p = .248$, 95% CI [-0.2, 0.78]. With this non-significant interaction effect, we therefore fail to support H4. Figure 2 provides a visual representation of this interaction model. Table 4 presents the regression coefficients for this model along with the relevant test statistics.

Further Analysis

To deepen our understanding of the findings above, we conducted further analyses focusing specifically on the predictive power of RMET scores. Two models were constructed, of which the first predicts percent correct for judging AI-generated portraits and the second predicts percent correct for judging human-made portraits. In both models, the centered RMET scores, the condition, and the

interaction term out of the two made up the three predictors. For both models, the full list of coefficients, including their respective test statistics, is given in Tables 5 and 6. A comparison of the coefficients for the RMET scores in both models shows a striking difference. Whereas for judging AI-generated art the RMET scores show essentially a near-null effect, $B = 0.02$, $SE = 0.14$, $t(228) = 0.1$, $p = 0.891$, 95% CI [-0.26, 0.3], when it comes to judging human-made art, RMET scores are a highly significant predictor, $B = 0.44$, $SE = 0.12$, $t(228) = 3.6$, $p < .001$, 95% CI [0.2, 0.69]. In contrast, the interaction term remained non-significant across both models, while the condition term consistently retained its statistical significance. Collectively, these findings indicate that RMET scores are able to predict better judgment accuracy exclusively for human-made but not for AI-generated portraits.

Discussion

This current investigation was designed to evaluate people's ability to distinguish between AI-generated and human-made portraits and to identify factors that enhance judgment accuracy. Central to this design was the hypothesis that inductive learning through interleaved practice would improve judgment accuracy, based on the idea that being exposed to different types of portraits helps people learn to spot the subtle differences between AI-generated and human-made art. To evaluate this, we analyzed the data with a linear regression model in which judgment accuracy was regressed on the condition. Descriptive statistics complemented this analysis. In accordance with our hypothesis, we found a significant model in which the training condition could account for 15% of the variance in judgment accuracy. This is consistent with the findings from Kang & Pashler (2012), in which interleaved practice proved useful to identify differences in the styles of paintings. This further solidifies the effectiveness of interleaved practice in enhancing familiarity with specific stimuli, ultimately improving the recognition of subtle details and the ability to distinguish between stimulus types. Furthermore, the significance of the results shows that the ability to distinguish between AI-generated and human-made art is malleable. This is a foundational yet important insight for future research, as it suggests that judgment accuracy is not fixed and can potentially be improved through the optimization of training interventions. The next section examines whether social cognition could provide a meaningful basis for shaping these interventions.

To represent an aspect of social cognition, we took measurements on ToM by employing the RMET. We hypothesized that people with higher ToM abilities would be better at distinguishing AI-generated from human-made art, built upon the reasoning that ToM could help in interpreting subtle social cues present in portraits. The results of a linear regression analysis gave support to this hypothesis, suggesting that ToM ability in individuals is able to predict judgment accuracy. This is in line with previous research findings that illustrated how social cognition is involved when judging visual cues that are rich in social content (Bayliss & Tipper, 2005; Frischen, Bayliss, & Tipper, 2007; Schilbach et al., 2006). These findings support our theoretical assumption that the content of portraits is meaningfully linked to a social-cognitive ability such as ToM. Individuals with higher ToM abilities are likely more attuned to the social cues conveyed in faces, which may enhance their ability to make accurate judgments. Consequently, individuals with lower ToM abilities fall into a critical group that should be the focus of policy guidelines because they are more at risk for making wrong judgments. Overall, these findings show us that stimulus content should be considered for investigating how perceptual judgment is shaped. Future research should account for such features when examining different forms of human-AI interactions. Further implications are restricted by the small increase in explanatory power of our model and the fact that the training condition is still the better predictor. It indicates that ToM contributes only marginally to judgment accuracy. Additional findings from our investigation underscore this point.

Based on the assumption that training and ToM are interconnected, ToM was hypothesized to moderate the effectiveness of the training, with greater training benefits expected for individuals exhibiting higher ToM abilities. We failed to find support for this hypothesis. This indicates that ToM has no additional influence on the extent to which individuals benefit from the training. A likely explanation for this non-significant result would be that both the training and ToM contribute independently to the prediction of judgment accuracy, therefore not interacting in any way. Although it is possible that ToM would influence other forms of training, we can at least deduce that for inductive learning through interleaved practice, there is no meaningful relationship. Additional findings from our analysis place further constraints on the implications for ToM.

After analyzing how ToM predicts judgment accuracy for human-made and AI-generated portraits, respectively, we have to conclude that the positive effect of ToM seems to differ from what we theorized. Previous predictions that we postulated for the effect of ToM were built on the assumption that there are flaws in the AI process that can be detected, meaning that higher social cognition makes one more likely to detect the elements of an AI picture that do not match the profile of usual social stimuli. Statistical analysis showed us, however, that ToM predicts judgment accuracy exclusively for judging human-made portraits. A central question that emerges from this is not only why ToM strongly predicts judgment accuracy for human-made portraits, but also why this predictive power does not extend to AI-generated portraits. Interpreting these findings, therefore, has to involve contrasting AI-generated and human-made portraits and how underlying features of these pictures might be related to causing high predictive power for judgment accuracy of only one of the two stimulus categories. A factor that potentially explains the difference in predictive power is that AI-generated portraits are created by an algorithm, whereas human-made portraits are a product of the artist's knowledge and craftsmanship as well as the client's expectations of the painting's effect. The shared contribution of these elements makes human-made portraits a product of a certain intentionality that individuals with high ToM abilities can pick up on, explaining the high predictive power for correctly judging human-made portraits. AI-generated portraits, on the other hand, do not show these characteristics of intentionality. Based on this, it is likely that these characteristics of intentionality serve individuals with high ToM abilities as informative markers to make accurate judgments. In the absence of these characteristics, individuals lack diagnostic cues to guide their judgments, meaning that those with high ToM abilities no longer have an advantage over those with lower ToM abilities in making accurate evaluations. The role of ToM as a predictor of judgment accuracy, therefore, appears to be limited. While we initially expected to identify individuals with low ToM abilities to be especially at risk of making wrong judgments about AI-generated art, our findings suggest otherwise. In practical terms, this indicates that individuals with higher ToM abilities are more likely to confidently and correctly recognize human-made art. However, when it comes to evaluating AI-generated images, ToM offers no advantage. As a result, we are unable to identify

a distinct risk group based on ToM ability for misjudging AI-generated art, limiting the utility of ToM in this context. These findings highlight the need to explore other individual differences.

Another candidate for a potential connection with judgment accuracy was Assessment as part of the Regulatory Mode Theory. We hypothesized that Assessment has a moderating effect on the training benefit, resulting in more training benefit for the individuals who are measured high in Assessment orientation. This prediction was based on the idea that people high in Assessment orientation are more inclined to critically evaluate options and search for accuracy, which may enhance their ability to recognize and make use of subtle distinctions learned during training. After investigating a linear regression model, we found no evidence that is able to support this hypothesis. The non-significance of both the main and interaction effect is best explained by the analysis having accurately reflected the absence of a meaningful influence of Assessment. A possible explanation for this null result is that the nature of our task may not have been sensitive to individual differences in Assessment. It is plausible that such differences become more apparent in tasks that allow for extended and careful decision making. In contrast, our task emphasized quick, intuitive judgments, which may have limited the influence of Assessment. Since a real-world scenario does not necessarily require quick judgment, further research has to be conducted to clarify the role of Assessment.

Limitations

The current study also exhibits a number of limitations. Although we were able to have sufficient statistical power through a large sample, the participants were recruited through convenience sampling and primarily consisted of students from similar backgrounds, which limits the representativeness of the sample. Another risk of this study design revolves around the creation of the stimuli. Generating AI images through prompts does not always produce perfect results, raising the question of whether studies in this context should use stimuli that reflect the average output of AI systems or instead represent a human's attempt to craft images that maximize the quality of AI-generated art. This threatens the representativeness of the stimuli, and together with the drawbacks of our sampling method, the external

validity of our study is very diminished. Furthermore, the reliability and validity of certain measures are in question since we had to shorten questionnaires in favor of practical concerns.

Conclusion and Future Directions

With the current project, we made a good first attempt to investigate how humans interact with AI products. We set out to explore several factors that potentially influence judgment accuracy for distinguishing between AI-generated and human-made portraits. With the significance of the training effect and also the involvement of ToM, we can primarily show that judgment accuracy in this context is variable and therefore likely susceptible to change with the right interventions. However, because the background of this study extended beyond the identification of cognitive learning processes, but also started out with the broader question of whether we can protect the integrity of art, we have to limit our findings for the moment since none of the results are promising enough to take this as a basis for an effective intervention strategy. What is important to consider is that, for this context, we would ideally want to create means by which individuals become near perfect in judgment accuracy, since the mere doubt of the integrity of someone's work can already be enough for the perceived value to decrease. The statistical analysis showed us, however, that for the trained participants, still a third of the stimuli are judged incorrectly. In an everyday situation, this means that no matter whether a person received training or not, there is a substantial risk for the stimuli to be judged in the wrong way. Future research should aim at refining the training methods by perhaps using longer training phases or a different method altogether, to find large enough effect sizes that suggest adequate practical relevance. With the marginal effect of ToM, we have the proof that psychological constructs are able to have an influence on judgment accuracy, albeit rather small. In future projects, the focus can shift to other psychological constructs that potentially facilitate judgment accuracy. The results on ToM also show that identifying these psychological constructs should ideally be directed by the content of the stimuli. In that way, it can also be fruitful to extend this line of research to other forms of artistic paintings besides portraits. The results from this attempt underscore once more the complexity and the need to refine our understanding of AI

related topics. With the rise of these technologies, it is therefore imperative to keep up with this development and find ways to sensitize people about the characteristics of AI-generated products.

References

- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, 56, 149–178. <https://doi.org/10.1146/annurev.psych.56.091103.070217>
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test Revised Version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42(2), 241–251. <https://doi.org/10.1111/1469-7610.00715>
- Bayliss, A. P., & Tipper, S. P. (2005). Gaze cuing of attention to biological motion. *Visual Cognition*, 12(2), 141–157. <https://doi.org/10.1080/13506280444000292>
- Bergmann, D., & Stryker, C. (2024, August 21). *Diffusion models for image generation – A comprehensive guide*. IBM. https://www.ibm.com/think/topics/diffusion-models?utm_source=chatgpt.com
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77(3), 305–327. <https://doi.org/10.1111/j.2044-8295.1986.tb02199.x>
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. Wiley.
- Chander, R. J., Grainger, S. A., Crawford, J. D., Mather, K. A., Numbers, K., Cleary, R., Kochan, N. A., Brodaty, H., Henry, J. D., & Sachdev, P. S. (2020). Development of a short-form version of the Reading the Mind in the Eyes Test for assessing theory of mind in older adults. *International Journal of Geriatric Psychiatry*, 35(11), 1322–1330. <https://doi.org/10.1002/gps.5369>
- Cohn, G. (2018, October 25). *AI art sold at Christie's: A new frontier in art and auctions*. The New York Times. <https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html>
- De Neys, W., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: Cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, 20(2), 269–273. <https://doi.org/10.3758/s13423-013-0384-5>

- Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19(4), 25–42. <https://doi.org/10.1257/089533005775196732>
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133(4), 694–724. <https://doi.org/10.1037/0033-2909.133.4.694>
- Kang, S. H. K., & Pashler, H. (2011). Learning Painting Styles: Spacing is Advantageous when it Promotes Discriminative Contrast. *Applied Cognitive Psychology*, 26(1), 97–103. <https://doi.org/10.1002/acp.1801>
- Kruglanski, A. W., Thompson, E. P., Higgins, E. T., Atash, M. N., Pierro, A., Shah, J. Y., & Spiegel, S. (2000). To “do the right thing” or to “just do it”: Locomotion and assessment as distinct self-regulatory imperatives. *Journal of Personality and Social Psychology*, 79(5), 793–815. <https://doi.org/10.1037/0022-3514.79.5.793>
- Mauro, R., Pierro, A., Mannetti, L., Higgins, E. T., & Kruglanski, A. W. (2009). The Perfect Mix: Regulatory Complementarity and the Speed–Accuracy Balance in Group Performance. *Psychological Science*, 20(6), 681–685. <https://doi.org/10.1111/j.1467-9280.2009.02363.x>
- Navarro, E. (2022). What is theory of mind? A psychometric study of theory of mind and intelligence. *Cognitive Psychology*, 136, 1–18. <https://doi.org/10.1016/j.cogpsych.2022.101495>
- Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A Questionnaire of Cognitive and Affective Empathy. *Journal of Personality Assessment*, 93(1), 84–95. <https://doi.org/10.1080/00223891.2010.528484>
- Schilbach, L., Eickhoff, S. B., Mojsisch, A., & Vogeley, K. (2006). Minds in the eyes: An fMRI study of the interaction of gaze and mental states. *NeuroImage*, 32(1), 336–347. <https://doi.org/10.1016/j.neuroimage.2006.03.014>
- Valentin, D., Abdi, H., O'Toole, A. J., & Cottrell, G. W. (1994). Connectionist models of face processing: A survey. *Pattern Recognition*, 27(9), 1209–1230. [https://doi.org/10.1016/0031-3203\(94\)90100-7](https://doi.org/10.1016/0031-3203(94)90100-7)

Veltman, C. (2025, February 17). *Christie's AI art auction protests*. NPR.

<https://www.npr.org/2025/02/17/nx-s1-5296911/christies-ai-art-auction-protests>

Westby, C., & Robinson, L. (2014). A developmental perspective for promoting theory of mind. *Topics in Language Disorders*, 34(4), 362–382. <https://doi.org/10.1097/TLD.0000000000000035>

Appendix A Tables and Figures

Table 1

Correlational Analysis and Descriptive Statistics

Variables	<i>M</i>	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8.
1. Condition	–	–	–							
2. Empathy	3.121	0.469	.025	–						
3. Assessment	4.189	0.602	.056	.273**	–					
4. RMET	5.711	1.672	.075	.090	.017	–				
5. CRT	1.948	1.022	.141*	.063	-.070	.194*	–			
6. % Correct (AI)	59.7	15	.307**	.167*	.163*	.099	.129	–		
7. % Correct (Human)	60.106	13.628	.299**	.070	.068	.302**	.022	.200*	–	
8. % Correct (Total)	60.309	10.996	.389**	.167*	.152*	.243**	.094	.790**	.748**	–

Note. $n = 232$

* $p < .05$

** $p < .001$

Table 2***Coefficient Table for Model with Condition as Predictor for Judgment Accuracy***

Model		Unstandardized	Standard Error	Standardized ^a	t	p	95% CI	
							Lower	Upper
H ₀	(Intercept)	60.309	0.722		83.537	< .001	58.887	61.732
H ₁	(Intercept)	56.293	0.915		61.491	< .001	54.489	58.097
	Condition (Training)	8.548	1.336		6.400	< .001	5.917	11.180

^a Standardized coefficients can only be computed for continuous predictors.

Table 3***Coefficient Table for Model with Condition, RMET Score and Interaction as Predictors for Judgment Accuracy***

Model		Unstandardized	Standard Error	Standardized ^a	t	p	95% CI	
							Lower	Upper
H ₀	(Intercept)	60.309	0.722		83.537	< .001	58.887	61.732
H ₁	(Intercept)	56.437	0.896		62.982	< .001	54.672	58.203
	RMET_centered	1.239	0.548	0.188	2.263	0.025	0.160	2.318
	Condition (Training)	8.190	1.308		6.263	< .001	5.613	10.767
	RMET_centered * Condition (Training)	0.358	0.783		0.457	0.648	-1.184	1.900

^a Standardized coefficients can only be computed for continuous predictors.

Table 4***Coefficient Table for Model with Condition, Assessment Score and Interaction as Predictors for Judgment Accuracy***

Model		Unstandardized	Standard Error	Standardized ^a	t	p	95% CI	
							Lower	Upper
H ₀	(Intercept)	60.309	0.722		83.537	< .001	58.887	61.732
H ₁	(Intercept)	56.328	0.908		62.013	< .001	54.538	58.118
	As_Total_centered	0.149	0.158	0.074	0.946	0.345	-0.162	0.460
	Condition (Training)	8.352	1.327		6.296	< .001	5.738	10.965
	As_Total_centered * Condition (Training)	0.290	0.250		1.159	0.248	-0.203	0.783

^a Standardized coefficients can only be computed for continuous predictors.

Table 5

Coefficient Table for Model with Condition, RMET Score and Interaction as Predictors for Judgment Accuracy for AI-generated Portraits

Model		Unstandardized	Standard Error	Standardized ^a	t	p	95% CI	
							Lower	Upper
H ₀	(Intercept)	10.746	0.177		60.620	< .001	10.396	11.095
H ₁	(Intercept)	9.970	0.233		42.877	< .001	9.512	10.428
	RMET_centered	0.019	0.142	0.012	0.137	0.891	-0.261	0.299
	Condition (Training)	1.623	0.339		4.783	< .001	0.955	2.292
	RMET_centered * Condition (Training)	0.212	0.203		1.042	0.299	-0.189	0.612

^a Standardized coefficients can only be computed for continuous predictors.

Table 6

Coefficient Table for Model with Condition, RMET Score and Interaction as Predictors for Judgment Accuracy for Human-made Portraits

Model		Unstandardized	Standard Error	Standardized ^a	t	p	95% CI	
							Lower	Upper
H ₀	(Intercept)	10.819	0.161		67.169	< .001	10.502	11.136
H ₁	(Intercept)	10.182	0.204		50.014	< .001	9.781	10.583
	RMET_centered	0.444	0.124	0.303	3.569	< .001	0.199	0.689
	Condition (Training)	1.364	0.297		4.590	< .001	0.778	1.949
	RMET_centered * Condition (Training)	-0.065	0.178		-0.366	0.715	-0.415	0.285

^a Standardized coefficients can only be computed for continuous predictors.

Figure 1

Interaction Effect of RMET Score and Training Condition on Judgment Accuracy

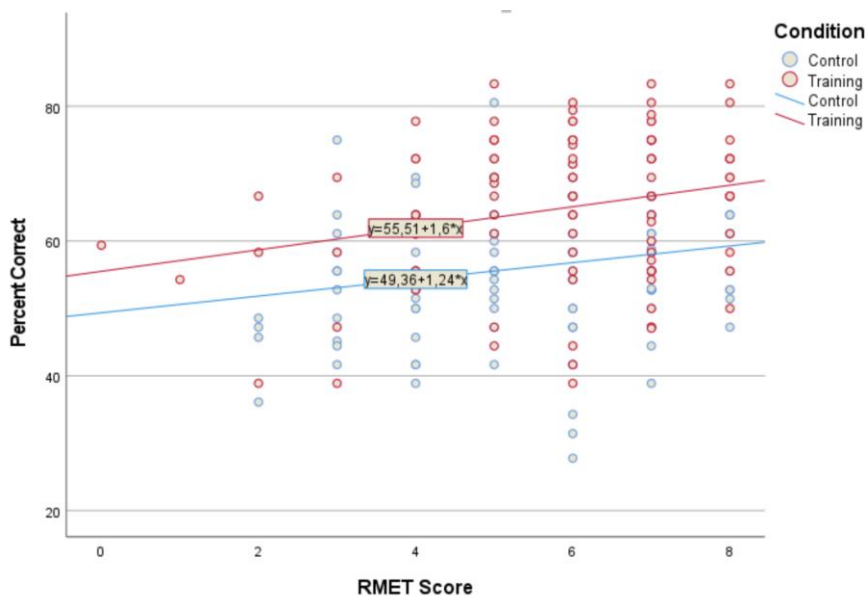
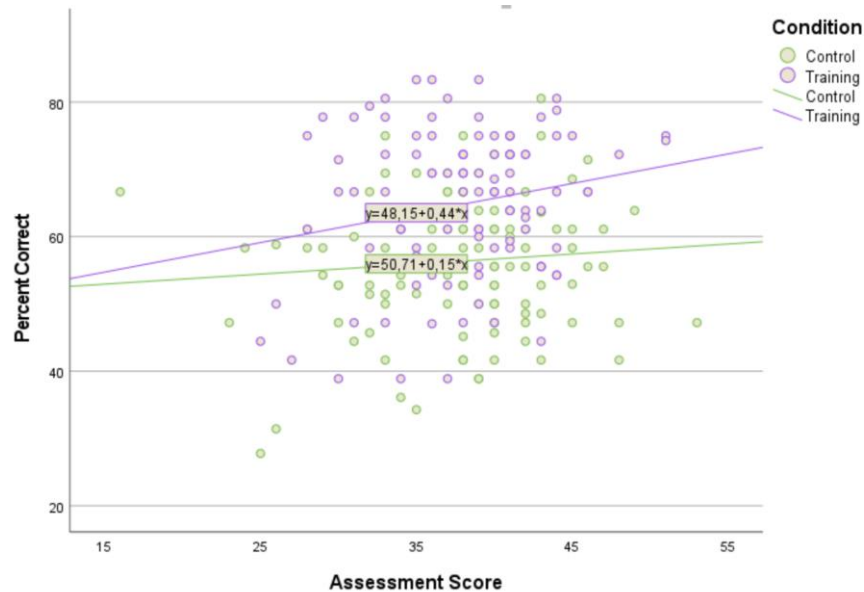


Figure 2

Interaction Effect of Assessment Score and Training Condition on Judgment Accuracy



Appendix B
Assessment Scale Items

“I spend a great deal of time taking inventory of my positive and negative characteristics.”
“I like evaluating other people's plans.”
“I often compare myself with other people.”
“I often critique work done by myself or others.”
“I am a critical person.”
“I am very self-critical and self-conscious about what I am saying.”
“I often think that other people's choices and decisions are wrong.”
“I rarely analyze the conversations I have had with others after they occur.”
“When I meet a new person I usually evaluate how well he or she is doing on various dimensions (e.g., looks, achievements, social status, clothes).”

Appendix C
AI-made Picture Prompts

Modern AI portraits

A bold, textured acrylic portrait of a woman in her early 40s, her face slightly asymmetrical with deep lines around her mouth and forehead. Her short, wavy auburn hair is painted with visible, thick brushstrokes, and her eyes are expressive but not overly detailed. The artist has left imperfections in the paint--small cracks, uneven layering, and drips near the edges. The background is rough, with streaks of mixed colors, resembling a weathered wall. Some areas of the painting look scraped away, revealing layers beneath. Urban contemporary realism, raw and unpolished, textured acrylic on canvas

A portrait painting in the Renaissance style depicting a young noble person

A portrait of a woman with flushed cheeks and wide eyes, overwhelmed by joy, intense color contrasts and dynamic brushwork, inspired by Gauguin

A portrait of a young man, likely in his late teens or early twenties. He is light-skinned, with short, dark brown hair styled neatly. He is facing the viewer, and looking directly at the observer. His expression is neutral and contemplative. He has a somewhat serious, or perhaps thoughtful, expression. His features are well-defined, with a pointed chin and prominent brow. He wears a light teal-blue coat or jacket with a dark, almost black, collar. There's a suggestion of a lighter-colored shirt underneath the jacket. He has a slightly prominent nose and full, lightly colored lips. The body type appears slender, with a classic, somewhat formal posture. The setting is a plain, light gray background, characteristic of a studio portrait. The lighting is directed from the front, casting subtle shadows on the face and the collar of his jacket, creating a classical painting atmosphere. The composition is centered and in a three-quarter view, with a close-up perspective of the subject's face. The color palette is limited to soft, muted tones of grays, browns, and pale blues, mimicking the look of an oil painting. The brushstrokes are visible, contributing to the style of a painted portrait. The overall atmosphere is calm, intimate, and formal

A Renaissance-style oil painting of a ugly woman wearing a pyjama. She sits next to a baby bed covered in white lace, holding a book. Her expression is calm and serious. The background consists of a soft blue house. The lighting is soft and realistic, with rich textures in the clothing and skin

A raw, expressive oil painting of a middle-aged man with deep wrinkles and a slightly crooked nose, painted in a contemporary abstract realism style. His face is full of character, with loose, gestural brushstrokes defining his strong jaw and furrowed brow. The colors are slightly exaggerated--warm reds and cool blues blending into the skin tones, creating a sense of movement. The paint is layered unevenly, with some areas thick and textured while others are thin and washed out, giving an unfinished, spontaneous feel. Some distortion in proportions adds a unique, handcrafted quality. Modern figurative painting, oil on canvas, thick and textured, imperfect yet deeply expressive

A woman standing in the rain, face tilted upward, eyes closed, conveying relief and sadness simultaneously, muted blues and grays, impressionistic brushwork

An emotional portrait painting in impressionism style

An expressive Impressionist oil painting of a young woman with auburn hair in an elegant loose bun. Her deep green eyes are full of life, and her fair skin has a warm, rosy glow. Painted in the style of Mary Cassatt, with loose but deliberate brushstrokes that capture movement and emotion. She wears a pastel blue dress with delicate ruffles, and the background consists of soft, blended colors, suggesting a sunlit garden. The oil paint is thick and textured, visible in every stroke, adding depth and richness to the portrait. The lighting is soft and natural, creating a sense of intimacy and warmth. Impressionist realism, highly detailed, museum-quality oil painting

An old, Impressionist-era oil painting of a man in his forties, with a tired but kind expression. His face is not perfectly symmetrical, and the brushwork is loose and expressive rather than hyper-detailed. His hair is a mix of gray and brown, painted in quick, gestural strokes. The canvas is old and has fine cracks throughout, especially in the lighter areas. The background is softly blurred with faded greens and blues, resembling an outdoor garden scene that has lost some vibrancy over time. The varnish has yellowed, and some edges of the canvas have frayed slightly. The oil paint has built up in certain areas, creating a rich but uneven texture. Aged Impressionist oil on canvas, soft, painterly, and authentically worn

Bold, expressive painting of a cool, rebellious young woman with a black bob haircut, biting her lip with a nonchalant gaze. She wears a red leather jacket, slightly oversized, giving her a tough, confident look. The brushstrokes are rough and unrefined, with a raw, energetic feel

Create a painting of a happy grandma in Modern art style

do you wanna make a self portrait painting of a brown tinted woman with a beautiful dress on. You can see the strokes of the brush and the structure of the painting canvas

do you wanna make a self portrait painting of a old man with a beard sitting on a chair and elderly clothes. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting canvas

do you wanna make a self portrait painting of a old man with a beard sitting on a chair and elderly clothes. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting canvas

do you wanna make a self portrait painting of this kid. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting canvas. It has to be Renaissance style

Expressive, painterly portrait of a 1 to 2-year-old baby with a sweet, innocent face and soft blonde hair. The painting is created with bold, rough brushstrokes, giving it a raw and textured appearance. The colors are rich and slightly abstract, with streaky, unrefined lines that emphasize movement and emotion rather than fine details. The style is loose and impressionistic, almost like an energetic, modern take on classic portraiture. The colours are dark

<https://s.mj.run/FmIxIRN8fXM> Paint a similar picture with emotions

impressionist style self potrait of a contemplating male from that era

impressionist style self potrait of a contemplating ugly male from that era

Paint a lithograph painting inspired by Lucian Freud's self portrait 'Reflection' from 1996. The protagonist should also look thoughtfully out of the picture frame. There are next to no details. Only very thick shadings that show the major wrinkles and complexions of the face. Be careful to make the painting really dark and gloomy, include sholders and upper torso and have no detail. The background is supposed to be covered in dark lines

Paint a self portrait inspired by Pablo Picassos art style. Paint a man with black hair, combed to the side, a green jacket with a white collar shirt below. The person has a relatively large nose and big eyes. The background is brown. Really emphasize that this is Picassos art style. Show that every line the artist drew is understandable and visible for the viewer

Portrait of a man, mid-40s, light complexion, European or North-European ethnicity, in a studio setting, with a thoughtful expression. He's wearing a dark navy blue, collared jacket-like garment, likely made of a slightly thick fabric with visible texture and brushstrokes showing the technique used. The jacket is in a loose fit, suggesting a comfortable, informal style. Beneath the jacket, a light-colored, possibly white or cream-colored collared shirt is visible. The man has short, neatly combed graying hair, a mustache and short trimmed beard. He is looking off to his right, his gaze slightly directed away from the viewer. He has a slightly serious but not stern expression. The posture is relaxed, but the body is angled slightly towards the viewer. The palette knife or artist's tool is present in the lower-right quadrant of the image, and its dark brown/black color contrasts with the muted tones of the background. The background is a neutral grayish-beige. Lighting is soft, studio-like, with subtle shading and highlighting the features of the subject. The composition is a close-up portrait, focusing on the subject from the shoulders up. The style of the image is reminiscent of early-20th-century portrait painting, with a focus on capturing the subject's mood and personality. oil painting style, detailed brushstrokes visible, realistic portrait, muted color palette, studio lighting, subtle shadows, 8k resolution, portrait painting, careful detail in facial features, subtly rendered details in clothing and environment

Portrait painting in baroque style

Realistic oil painting portrait of a stern-looking man wearing a black hat, with a well-groomed mustache, dressed in vintage attire. The lighting is dramatic with soft shadows, and the brushstrokes are expressive yet detailed. The color palette is warm and natural, featuring earthy tones. The style resembles late 19th-century European realism, with a textured painterly effect. The background is subtle and muted, ensuring focus on the man's intense gaze

Non-modern AI portraits

17th century classicism oil portrait

A 17th-century Baroque painting of an elderly woman with a solemn expression. She wears a traditional black dress with a large white pleated collar and a delicate white bonnet covering her hair. Her wrinkled face and sad eyes convey wisdom and experience. She holds a small cloth in her hands, which are aged and realistic. The background is dark and subtly lit, enhancing the chiaroscuro effect. The painting style resembles that of Gerrit Dou or Rembrandt, with soft lighting, rich textures, and a warm color palette

A baroque portrait painting of a young large woman, capturing a moment of contemplative solitude with a subtle undertone of bathroom humour. The woman is dressed in a classic outfit, her hand delicately

covering her nose as if she grossed out by a sickening smell. The painting style mimics the textured brushstrokes and moody lighting of baroque art, suggesting a blend of nostalgia and wit. The image should evoke a sense of timeless elegance while also hinting at an amusing, light-hearted secret, perfect for a quirky and sophisticated bathroom decor piece

A coloured expressionist oil portrait with an old man holding an onion

A modern portrait of a happy old men in the 19th century

A painting of a woman with a dress sitting in front of natural scenery, painted in the Renaissance style.

The painting incorporates characteristics of sfumato and shows idealized beauty.

A portrait painting in Renaissance style incorporating the characteristics of Sfumato and Chiaroscuro to depict a noble person with idealized beauty and showing characteristics that make the painting seem real

A portrait painting in Renaissance style incorporating the characteristics of Sfumato and Chiaroscuro to depict a trader or merchant with simple and neutral facial expressions and showing characteristics that make the painting seem real. Also make the faces with low detailedness and few variations in color.

A Renaissance-style oil painting of a bearded man wearing a black robe and a black beret. He sits at a table covered in red fabric, holding a small piece of paper. His expression is calm and serious. The background consists of a deep green curtain on one side and a blue wall with red cords. The lighting is soft and realistic, with rich textures in the clothing and skin. The painting style resembles that of Hans Holbein the Younger, with fine details and a balanced composition.

A romanticism style portrait of a young guy with short blonde hair who is holding a rapier pointing to the floor and the other hand is behind his belt

A striking Baroque-style oil painting of a 40-year-old man with a strong, angular face, deep-set blue eyes, and a thick, well-groomed beard. His expression is serious, his gaze piercing. Painted in the style of Caravaggio, the artwork features dramatic chiaroscuro lighting, with deep shadows and warm highlights that sculpt his face. He wears a dark velvet cloak with a subtle gold embroidery trim. The brushwork is visible in the fine details of his skin, capturing every imperfection, subtle wrinkles, and the natural sheen of oil paint. The background is almost black, with a faint hint of an ornate interior. This masterpiece looks like it belongs in a museum, highly detailed and lifelike.

An aged oil painting of a middle-aged nobleman with a strong, weathered face, painted in the style of Hans Holbein the Younger. His gaze is solemn, and his deep-set brown eyes show wisdom. His skin has natural imperfections--wrinkles, slight discoloration, and rough texture--captured with loose yet deliberate brushstrokes. The painting's surface has visible cracks and worn edges, with faded pigments in certain areas. He wears a dark velvet coat with a gold clasp, and the background is a muted brown with faint traces of an old interior. The brushstrokes are visible in the fabric and skin, with a slightly uneven texture. Oil on aged wooden panel, 16th-century realism, museum-quality but aged.

An old, faded oil painting of a young woman in the style of 19th-century Romanticism. Her expression is melancholic, her blue eyes slightly unfocused, as if lost in thought. Her hair is loosely curled, painted with soft, impressionistic strokes rather than precise details. The paint has aged, showing fine cracks across her forehead and faded colors in her pale pink dress. The background was once vibrant but has darkened with age, leaving an uneven, worn texture. The varnish has yellowed, giving the entire piece a warm, antique tone. Some areas of the canvas show slight warping, making it look truly vintage. Aged oil on canvas, soft focus, timeworn masterpiece. Little detail in the face and somewhat ugly person. Make the face less smooth and perfect.

do you wanna make a self portrait painting in of a woman with an old linen cloth to her. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting canvas

do you wanna make a self portrait painting of a old man with a beard and wearing elderly clothes. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting canvas. It has to be Renaissance style.

do you wanna make a self portrait painting of a old man with a beard and wearing elderly clothes. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting canvas. It has to be Renaissance style.

do you wanna make a self portrait painting of a tinted woman wearing elderly clothes. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting canvas. It has to be Renaissance style.

do you wanna make a self portrait painting of a woman sitting on a wooden chair and elderly clothes. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting canvas

do you wanna make a self portrait painting in baroque style. You can see that it's a painting from the 16th century. You can see the strokes of the brush and the structure of the painting

European nobleman portrait, 17th or 18th century style, middle-aged man, light complexion, with powdered, shoulder-length light gray wig, looking directly at viewer, slightly serious expression, wearing rich, dark reddish-purple velvet doublet, ornamental gold embroidery and trim, fitted jacket over light-colored armor, lace cuffs and cravat, light grayish-white, formal shirt, gloves, holding a helmet in his left hand, furred cloak in deep blue, fur lining visible on the shoulders and around the cloak. Gentleman poses formally, standing, slightly angled toward the viewer, leaning slightly forward; light tan background; soft, rich colors, with subtle shading and highlights, creating a sense of depth and realism, subtle lighting, focal point on the subject. A backdrop of a light tan and brown landscape, possibly a classical architectural structure or ruins. Classic portraiture style, Baroque period, oil on canvas, detailed brushstrokes and texture, fine details in clothing and facial features, perspective and composition are precise and traditional.

<https://s.mj.run/R3nIoC02qGM> Can you paint a painting like this but then a man with a costume on? But the same style

<https://s.mj.run/tFCOKavql7Q> can you paint a painting, looking like this painting?

make a 17th century classicism portrait of two older woman working on a field

make a portrait from 18th century depicting a priest deep in prayer

old oil portrait of an old japanese couple, the woman is looking sideways

Paint the portrait of an african man, painted ca. 1525 - 1530. Really try to incorporate the distinct style of this art epoche. make it less realistic and less detailed

Painting in Baroque-style painting of a noblewoman with delicate facial features and a subtle, confident smile. She wears an elaborate ruffled lace collar and a richly decorated hat with golden embroidery. The portrait has a dark, moody background, emphasizing the subject's soft, luminous skin. The artwork is inspired by Flemish Baroque painters like Peter Paul Rubens, featuring warm skin tones, and dramatic chiaroscuro lighting. The painting has low contrast, not detailed, looking like it's an old painting made in 1600. you should be able to see well the details of the paint brushes. And the texture. The painting was made in 1533 and the art style should be not very detailed and very simple. Her face should be dark and not lighted out.

Please create a drawing of a self portrait, by a thirteen yearold of the 13th century. The portrait is made on rugged paper with a bit of text in the top right corner. It features the self-portrait of the side profile including upper body, featuring torso, and arms up until the hips, of a 13 year old boy. The boy wears a simple head which looks more like a piece of cloth redesigned to fit the head. He has big round cheeks and a slightly retracting chin. The forehead of the boy is covered in hair. The drawing was made with a silverpen. The drawing features a young boy with long hair and slightly oversized clothing.

portrait painting of a musician and composer from the 18th century

portrait painting of a woman with a dress on from 18th century

portrait painting that has been painted between 1400 and 1800

Portrait of a middle-aged man, European features, likely 50s, bald head, wearing a dark-colored, long-sleeved coat with a visible fur collar in a light brownish-gold tone. A white lace or linen ruff is around his neck. The subject's expression is neutral, direct, and somewhat serious. He appears to be looking directly at the viewer, with a somewhat pensive or contemplative look. The painting style is Baroque/Italian Renaissance, emphasizing realistic details and the use of light and shadow. The background is a plain, dark brownish-gray, drawing focus on the subject. The lighting is focused on the subject's face and upper body, creating a subtle contrast between the subject and the background. The colors are muted, with a rich

palette of earth tones. The texture of the paint is visible, suggesting brushstrokes and an oil painting technique. The composition is centered, showing a bust-length portrait. The perspective is straightforward. Detailed brushstrokes create a sense of depth and volume on the subject's face and clothing. High detail, oil painting, realistic portrait, Baroque style, 17th century, European portrait.
sandro botticelli old renaissance portrait
sandro botticelli old renaissance portrait

Appendix D

Human-made pictures

Modern human-made portraits

Afrikaanse Vrouw Met Hoofddoek - Bianca ter Riet
Angelica - Daniel Graves
Bernadette - Kim van der Enden
Der Maler Selbstportrait - Ernst Ludwig Kirchner
Egyptian Girl - Alexej von Jawlensky
Girl Arranging Her Hair – Abbott Handerson Thayer
Jünglingskopf - Alexej van Jawlensky
Ketsia - Daniel Sprick
Levi - Kim van der Enden
Francis Bacon – Lucian Freud
Mulati Nieu - Karl Merilaid
Portrait - Franz Marz
Portrait - Simon Davis
Portrait of a child - A. J. Alper
Portrait study of a woman - Ignacy Marek
Portrait young male in profile with glasses - Elena Zorina
Portrait of Artist's daughter – William Merrit Chase
Selbstporträt - Ernst Eitner
Self portrait - Raymond McIntyre
Self-portrait – Lois Mailou Jones
Self-portrait - Pablo Picasso
Self-Portrait: Reflection - Lucian Freud
Tashi in Red - A.J. Alper
The Two Fridas - Frida Kahlo

Non-modern human-made portraits

A bearded man wearing a hat - Rembrandt van Rijn
Admiral Harry Paulet – Francis Cotes
André François Alloys de Theys d'Herculais - Nicolas de Largillierre
Bildnis Einer Alten Frau - Gerrit Dou
Countess of Derby – George Romney
Ellen Maurice - Marcus Gheeraerts the Younger
Female Martyr Saint – Carlo Francesco Nuvolone
Filippo Archinto – Tiziano Vecellio
George Luks – Robert Henri
Girl at the Window - Eugen von Blaas
Gustave Boyer – Paul Cézanne
Honeysuckle Bower - Michael John
Chancellor Leonhard von Eck – Barthel Beham
Charles Rousseau Burney - Thomas Gainsborough
Lady Guildford - Hans Holbein the Younger

Madonna and Child - Giovanni Bellini
Monk in Prayer - French Painter
Nell Gwyn - Peter Lely
Portia - John Everett Millais
Portrait Luise Polácková - Stanislav Kasparides
Portrait of a Woman - Adriaen Hanneman
Portrait of a Young Woman - Lorenzo di Credi
Portrait of an African Man - Jan Jansz Mostaert
Portrait of Pope Clement IX - Carlo Maratta
Selbstporträt - Albrecht Dürer
Self Portrait - George Spencer Watson
Self Portrait - Giovanni Battista Salvi da Sassoferrato
Self Portrait - Milan Thomka Mitrovský
Self Portrait - William-Adolphe Bouguereau
Self-portrait - Anna Dorothea Therbusch
Self-Portrait with a Hat - Paul Raud
Self-Portrait with the Colosseum, Rome - Maerten van Heemskerck
Self-portrait at 69 Years - Francisco de Goya
Selfportrait - Antonín Procházka
The Venetian Girl - Frank Duveneck
Toussaint Lemaistre - Jean-Baptiste-Camille Corot

Appendix E

Declaration of AI usage

No content generated by AI technologies has been presented as my own work.

I acknowledge the use of ChatGPT (<https://chatgpt.com/?model=auto>); to generate materials for background research and self-study in the drafting of this assessment.

Example:

Prompt: “Give me a list of sources that show research on theory of mind. Specifically I would like to have sources that show that high levels of Theory of Mind are associated with better judging capabilities. The context of judging should be preferably that of judging visual stimuli but it can be others as well”

Output:

Research indicates that higher levels of Theory of Mind (ToM)—the ability to attribute mental states to oneself and others—are associated with enhanced judgment capabilities, particularly in interpreting visual stimuli. Below is a curated list of studies that explore this relationship:

□ Theory of Mind and Visual Judgment

1. Visual Perspective Taking and ToM Brain Regions

This study examined brain activation during visual perspective-taking tasks, highlighting that areas like the right temporo-parietal junction (rTPJ) and medial prefrontal cortex (mPFC) are

involved in processing another's visual perspective. These regions are also crucial for ToM, suggesting a link between ToM abilities and visual judgment accuracy.

2. **Gaze Following and ToM in Visual Scenes**

Research demonstrated that individuals with higher ToM abilities are more adept at interpreting gaze cues in visual scenes. Participants were better at determining where an observer was looking, especially when the gaze was directed at animate objects, indicating that ToM enhances the interpretation of visual attention.

3. **Shot Scale in Film and ToM Tendency**

An experiment involving film viewers found that close-up shots of characters' faces increased viewers' tendency to recognize mental states, thereby enhancing their ToM tendency. However, this effect did not extend to improving general ToM ability, suggesting that while visual stimuli can influence ToM tendency, they may not necessarily enhance overall ToM abilities.

4. **Theory of Mind and Metaethical Judgment**

This study explored how ToM network activity correlates with metaethical judgments, which involve assessing the objectivity or subjectivity of moral statements. Findings indicated that ToM-related brain regions are engaged during these judgments, suggesting that ToM abilities influence complex evaluative processes beyond basic visual stimuli.

□ **Additional Resources**

- **Theory of Mind and Moral Judgment**

This article discusses how ToM influences moral decision-making, particularly in understanding others' beliefs and intentions, which can extend to interpreting visual cues in moral contexts.

- **Neural Basis of ToM in Blind Individuals**

Research on congenitally blind individuals revealed that ToM-related brain regions are active even without visual experience, suggesting that these areas process abstract representations of others' mental states, which can influence judgments in non-visual contexts. [PubMed](#)

These studies collectively suggest that higher ToM abilities are associated with improved judgment capabilities, particularly in interpreting visual stimuli. While some effects are specific to certain contexts, the overarching theme is that understanding others' mental states enhances one's ability to make accurate judgments in various situations.

I followed up on this with the following prompt: “give me a reference list for the second bullet point”

Output:

Here is a reference list for the second bullet point, "**Gaze Following and ToM in Visual Scenes**":

1. **Senju, A., & Johnson, M. H. (2009).** The eye gaze cueing effect: A developmental study. *Developmental Science*, 12(5), 663–671. <https://doi.org/10.1111/j.1467-7687.2009.00833.x>

2. **Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007).** Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133(4), 694-724. <https://doi.org/10.1037/0033-2909.133.4.694>
3. **Hood, B. M., & Macrae, C. N. (2008).** Examining the role of attention in gaze cueing: Evidence from a visual search task. *Journal of Experimental Social Psychology*, 44(2), 519-522. <https://doi.org/10.1016/j.jesp.2007.09.003>
4. **Schilbach, L., et al. (2006).** Minds in the eyes: An fMRI study of the interaction of gaze and mental states. *NeuroImage*, 32(1), 336-347. <https://doi.org/10.1016/j.neuroimage.2006.03.014>
5. **Bayliss, A. P., & Tipper, S. P. (2005).** Gaze cuing of attention to biological motion. *Visual Cognition*, 12(2), 141-157. <https://doi.org/10.1080/1350628044000292>

These sources examine how gaze cues and Theory of Mind abilities are related in terms of interpreting visual attention, and they explore how ToM impacts judgment, particularly in relation to visual stimuli and social contexts.

Ultimately, I used the sources #2, #4, and #5 in my paper. Roughly in this way I curated several bits of background information for my final product.

I acknowledge the use of ChatGPT (<https://chatgpt.com/?model=auto>); to generate materials that are included in my work in a modified form.

Examples:

Prompt: Improve this: "A central question that emerges from this is not only why ToM has high predictive power for human-made portraits but also why this in turn is not the case for AI-generated portraits."

Output:

Sure! Here's an improved version with clearer structure and phrasing:

"A central question that arises is not only why ToM strongly predicts performance for human-made portraits, but also why this predictive power does not extend to AI-generated portraits."

If you'd like a slightly more formal or academic tone, here's an alternative:

"This raises a key question: why does ToM exhibit strong predictive value for human-made portraits, but not for AI-generated ones?"

Let me know if you want it tailored further for a specific section (e.g., discussion, abstract).

I used the first suggestion in my text while keeping the beginning of my sentence. In that way I used only parts of the output for my final version.

Prompt: "Central to this design was the hypothesized positive effect that inductive learning through interleaved practice would have on judgment accuracy." Rephrase this differently so that the reason for this hypothesis is stated

Output:

Certainly! Here's a clearer rephrasing that includes the reasoning behind the hypothesis:

"Central to this design was the hypothesis that interleaved practice would improve judgment accuracy, based on the idea that comparing different types of portraits side by side helps people learn to spot the subtle differences between AI- and human-made art."

Let me know if you want a more formal or more concise version!

I altered this version slightly to come to this final version that is included in my text:

"Central to this design was the hypothesis that inductive learning through interleaved practice would improve judgment accuracy, based on the idea that being exposed to different types of portraits helps people learn to spot the subtle differences between AI-generated and human-made art."

In this way, I used ChatGPT to improve my wording/phrasing/structure of certain sections while still incorporating my own ideas as to what needs to be included.