Are visual artists better visual perspective takers? An exploratory study with an unexpected outcome.

Steven Samuel.

City, University of London, Department of Psychology, School of Health and Psychological Sciences, UK.

Word count: 5778 (excluding title page, abstract, tables, captions)

Address for correspondence: Steven Samuel, Department of Psychology, City, University of London, Northampton Square EC1V 0HB. Email: steven.samuel@city.ac.uk

Abstract

Visual perspective taking (VPT) has been argued to elicit image-like representations of other people's visual experiences. Separately, it has been demonstrated that there are inter-individual differences in the ability to successfully take other people's visual perspectives. In the present study, adults were asked to judge how long two lines appeared visually from the point of view of an agent. The lines were of identical length, but the agent was always closer to one of the lines than the other, meaning that the closer line should be judged as appearing visually longer. It was hypothesised that adults with experience in the visual arts would perform better at this task for one or both of two reasons: 1) they should be more familiar with the knowledge that the closer an object is the larger it appears visually (i.e., the retinal image is larger), and 2) they might be able to "draw" an image-like representation that more accurately reflects the effect of distance on perceived size. Consistent with previous experiments with this paradigm, adults generally failed to judge the closest line as appearing longer; indeed, as many judged this line would appear visually shorter. Crucially, increasing experience in the visual arts failed to improve the accuracy of VPT judgments; even a group of professional illustrators failed to recognise that the line closest to the agent would appear longer than the line furthest from the agent. These results are discussed in the context of the processes and representation types potentially involved in VPT.

Visual perspective taking (VPT) concerns the ability to represent and/or make judgments about the viewpoint of another person or from another location (Samuel et al., 2023). VPT occurs when a perspective taker (henceforth the *observer*) needs to understand whether something is in the field of view of another person (henceforth the *agent*). This might occur for example just before the observer asks the agent to "please pass the salt," to first make sure the agent is aware of the salt. VPT also occurs when we want to understand how things might appear visually different to the agent, such as whether the table number is 6 or 9 depending on which side of the table you might be on. For the simpler, binary problem concerning whether an object is seen, often termed Level 1 perspective taking, research has shown that observers can 'draw' a line from the agent to the target object and conclude the target object is seen if the line is unbroken (Michelon & Zacks, 2006). For questions of the relative appearance of a target object, also known as Level 2 perspective taking (Flavell et al., 1981; Masangkay et al., 1974) there are many different candidate processes, each subject to individual differences both in terms of process selection and performance (for a review see Samuel et al., 2023).

Since vision concerns the perception of features such as colour, depth, edge, and size, one way in which VPT might occur is through the generation of a representation that captures this type of content. For example, Eleanor Ward and her colleagues refer to VPT as generating a "quasi-perceptual" representation related to mental imagery (Ward et al., 2019; Ward et al., 2020), a view that appears similar to that of Moll and Kadipasaoglu (2013), who talk about VPT representations as "snapshots" in the "optical" sense. Others have talked about representations that simulate "sensory consequences" (Kessler & Thomson, 2010), about how others' "visual experiences" influence the perspective-taker's own (Capozzi et al., 2014), and about "literally seeing the world through another's eyes" (Zhao et al., 2015). Some of the language used has been criticised for suggesting access to others' visual input (for

discussion see Cole & Millett, 2019; Cole et al., 2020; Cole et al., 2022). However, it is clear that descriptions like these suggest representations can be likened to an image, possibly in a similar way to how a photograph can capture a scene. A cognitive representation of other's perceptions that simulated that perception would be a powerful tool. However, a problem arises in that such images are not "collected" from the agent but must have their content created by the observer; observers will sometimes commit errors because they cannot corroborate their representation with the ground truth of the agent's visual experience. It follows that variability in the knowledge and/or expertise of the observer in the rendering of visual perspective into images could then impact the fidelity of the representation, and consequently the degree of success achieved in the VPT task. Similarly, while image-like representations contrast with rule-based or propositional alternatives that may themselves code for magnitudes through numbers, without the need to posit depictive mental imagery (Pylyshyn, 2001; Pylyshyn, 1973), these too should be influenced by the perspective-taker's own knowledge and experience. The present study investigates whether experience with visual arts makes people better at 1) *understanding* that objects closer to an agent will appear visually larger to her and or 2) *depicting* how much larger closer objects are in any image-like representation of the agent's visual experience.

A visual artist will often need to render the three dimensional world in a flat, 2D image, and doing so requires that they depict objects veridically or *proximally*, closer to a retinal image, rather than with post-encoding corrections (visual constancies) such as depth processing (Perdreau & Cavanagh, 2013). For example, a cow should be depicted as larger than a boy if they are side by side, but the further the cow recedes into the background the smaller it should appear, until eventually it must be depicted a smaller than the boy. It may therefore be that experience studying and producing visual art, particularly illustration (drawing, painting, and so on) will predict more accurate visual perspective taking when the

task involves this type of understanding. This could occur via one of two ways. Minimally, visual artists should be less likely to erroneously indicate that objects closer to an agent would appear smaller because they should have a more practical understanding that closer objects are visually oversized relative to objects further away. Alternatively, or *additionally*, visual artists might render a more accurate mental image of an agent's visual perspective, much as they render their own perspective more accurately on a canvas. However, a caveat on any hypothesis that visual artists are better visual perspective takers is that artists work with their *own* perspective, not other people's, and experience from the former may not extend to social cognition.

Prior research has found that adults are surprisingly poor at VPT when the problem concerns the relationship between the sizes and distance of objects from other perspectives. In an earlier multi-experiment VPT study by Samuel, Hagspiel, Eacott, et al. (2021), adults were shown an image of an agent looking at two horizontal lines of equal length and asked to indicate, using sliders, how long each line appeared to the agent (see Figure 1, left panel). To the agent, who was always to one side of both lines, the closer line appears about twice as long visually as the further line, as verified by a photograph taken by the agent (see Figure 1, right panel). Despite being asked to ignore the agent's knowledge (and their own) that the two lines were of equal length and to focus instead on how long the lines appeared *visually*, participants were not only unlikely to judge the closer line to appear longer, but they were equally likely to commit the unexpected error of judging that the *further* line appeared longer. To control for the possibility that participants may have been basing responses on the agent's knowledge rather than visual experience, a condition was introduced in which the agent was replaced with a camera and participants were asked how long each line would appear visually in a photograph the camera took. Despite the fact that the photo has no knowledge of the true length of the lines and is a 2D rendering of 3D space that *must* depict nearer objects as larger,

the accuracy of participants' judgments did not improve. Overall, the study revealed that adults find it difficult to infer the relative sizes of objects from other visual perspectives.

To explain this surprising pattern of results, Samuel, Hagspiel, Eacott, et al. (2021) speculated that many participants may have applied an erroneous "folk optics" rule, such as the belief that objects that are further away are somehow visually extended to compensate for their stretch into depth. Adults generally are quite susceptible to folk theories of perspective, as exemplified by the Venus Effect, whereby an agent who looks at an observer in a mirror is misunderstood as seeing their own face in the mirror despite the observer, agent and mirror not being along the same line of sight (Bertamini et al., 2003; Bertamini & Soranzo, 2018). This effect is exploited by visual artists and in film to present the illusion of an agent looking at themselves in a mirror, something that would otherwise be difficult without the artist or camera also having to appear in the mirror. Samuel and colleagues also speculated that visual artists may be less susceptible to folk optics strategies and might therefore be less likely to judge that the line furthest to the agent would appear visually longer to her.

The present study investigated whether experience with visual arts improves VPT accuracy. This was tested by recruiting participants to perform the VPT task designed by Samuel and colleagues, measuring their accuracy on the task, and relating this to a separate measure of their knowledge and experience of visual arts. Specifically, participants judged the length of each line as they appeared visually from the agent's viewpoint, and a "VPT Ratio" for each participant was calculated by dividing the length judgment for the longest line by the length judgment for the shortest line (as judged by the participant). Each ratio was then scored as positive or negative; a positive ratio was applied to a correct response (that the closer line appears visually longer), a negative score an incorrect one (that the closer line appears visually shorter). This ratio was then correlated with a self-report measure of each participant's visual Art Experience. A positive correlation would support the hypothesis that

visual artists are better visual perspective takers. This could be because experience with visual arts makes one less susceptible to the erroneous view that objects further away could appear larger. Alternatively, it could be because artists are practised at computing the proximal size of objects given an idea of their distance from a given viewpoint, and that this can improve the fidelity of any image-like representation of the two lines from the agent's perspective. These two hypotheses differ in terms of whether they are an advantage of an artist's *knowledge*, which should suffice at least to prevent the obviously erroneous closer-is-shorter response, or an advantage of an artist's ability to *depict* the relative size of each line. The latter would appear to trump the former; that is, if artists can judge how much larger the closer line appears, then it follows that they will not commit the error of judging the closer line to appear shorter. The former, however, should act minimally as a backstop to prevent the reverse error from occurring (i.e., the closer-is-shorter error). Initially, one experiment was planned where adults would be recruited without a priori filtering for art experience (henceforth Experiment 1a), but low levels of such experience in this initial sample led to the direct recruitment of professional illustrators in Experiment 1b in order to provide a broader distribution of scores for analyses.

Experiments 1a and 1b

Method

Participants

Experiment 1a. Initially designed as a single, stand-alone experiment, A power analysis using G*Power (version 3.1.9.6) found that 84 participants were required for an 80% chance to detect a medium effect size ($r = .03$) with a bivariate correlation, using an alpha of .05 and a two-tailed test. A hundred participants were recruited using *Prolific Academic* (www.prolific.co) with the following pre-screeners used within the platform: Age 18-45, a

minimum of 10 previous studies and maximum 1000 with *Prolific*, currently located in the UK, normal or corrected-to-normal vision, English as first language, and UK as the country the participant spent the most time in before the age of 18. The English language requirements were necessary to understand the nuances of the questionnaires in the study. All participants passed all four attention checks (three in the Interpersonal Reactivity Index, one in the Art Experience questionnaire). The data from two participants were removed for outlying VPT Ratios (-24 and -12.5, both over 3.5 the inter-quartile range from the mean), two more for giving zero for both line length questions on the sliders, and one for failing to complete any of the tasks. Nine further participants were replaced for giving a zero-length judgment for one of the two lines. Of the final 86 participants (Mean Age = 31 years) 40 identified as male and 46 as female.

Following data collection, it became clear that there was not enough variance in the sample to permit meaningful inferences from correlations with the Art Experience score, with 52 of 86 participants (60%) reporting scores of zero. This limitation was addressed in Experiment 1b. As a result, the data from Experiment 1a are not analysed separately but as part of a combined analysis with Experiment 1b so that inferences about Art Experience could be more reliable.

Experiment 1b. Since the majority of participants in Experiment 1a indicated that they had no visual arts experience beyond compulsory education, Experiment 1b was designed to augment the size of the Experiment 1a sample and test more directly the possibility that visual artists are better visual perspective takers. Experiment 1b was identical to Experiment 1a except for an additional participant eligibility requirement and the use of a between-subjects design on the new subset of data. The two groups were named Artists and Musicians. For the Artists, experience in creating illustrations not merely recreationally but as

part of one's work was a requirement, together with an explicit indication that the person did *not* play a musical instrument. This was done using the pre-screening in *Prolific*, which has a specific pre-screening questions for both of these criteria. Limited numbers of eligible participants for the Artists group in particular (92 in total) meant that an a priori power analysis was omitted, as it was unlikely to be sufficient for an 80% chance to detect a medium effect size with a between-subjects *t*-test. For the Musicians, the pre-screening features of *Prolific* were again used to recruit those who had declared at least five years' experience playing a musical instrument but had also explicitly indicated *no* experience in creating illustrations. This group was recruited as a comparison with Artists because of their similarity with visual artists in terms of creativity but difference in experience with *visual* arts specifically. The requirement in Experiment 1a to have spent most of one's time prior to the age of 18 in the UK and to have a current UK location were kept because widening these parameters did not add enough potential participants to outweigh the risk that some people may not have English as their first language but declared they did to be eligible for more studies. Finally, anyone who had participated in Experiment 1a was automatically excluded from Experiment 1b. A total of 58 participants were recruited for Experiment 1b. Two participants were excluded for failing one of the four attention check questions each. The data from two further participants, one from each group, were excluded from the analysis for outlying scores (ratios of -23 and +29.5). Final numbers were 27 participants in the Artists group (*M* = 33.3 years, *SD* = 6.8 years, 17 female, 8 male, 2 non-binary) and 27 participants in the Musicians group (*M* = 33.5 years, *SD* = 5.8 years, 14 female, 12 male, 1 non-binary). The final sample resulted in a reduced chance (43%) of detecting a medium effect size (*d* = 0.5) with a two-tailed between-subject t-test and an alpha of .05. Results should therefore be interpreted with caution.

Materials and Procedure

Visual perspective-taking task. The VPT task was closely modelled on that used in Samuel, Hagspiel, Eacott, et al. (2021). Participants were shown an image of a woman looking at two lines on a wall (see Figure 1). Above the image was the following: "Please look at the following picture in which there is a woman looking at two lines on a wall and answer the question below." The question below the image was as follows: "The woman in the photo knows that both lines on the wall are the same length. However, how long does each line actually *appear* from her visual perspective? Please drag the sliders below to answer, starting with the line on the [LEFT/RIGHT]." The position of the woman relative to the lines (left or right) and the line that the first (top) slider was related to (left or right) was counterbalanced. Each slider started in the leftmost, zero position, and dragging the slider to the right increased the length judgment, which ran up to a maximum of 100. The absolute values of line length judgments were not of interest, only the relative judgment, and the 0-100 scale was therefore not based on any real-world metric. Participants saw the number change as they dragged the slider. Both sliders were available at the same time, just below the image, which remained on-screen throughout (that is, participants did not need to rely on memory of the image, they could refer to it at leisure). No time limit was placed on responses. If a slider was not dragged at all (i.e., if it was left on zero) the participant was told that the response was not complete, and they were not allowed to continue until each slider had been moved. Participants could, however, first drag and then return the slider to a zero position. If this occurred, that participant's data were removed from the analysis, as a zero-length judgment suggested the participant erroneously judged that the line/s in question were not visible to the agent. This would suggest inattention, as the text above the image made explicit that the agent was "looking at two lines on a wall" (not merely facing them).
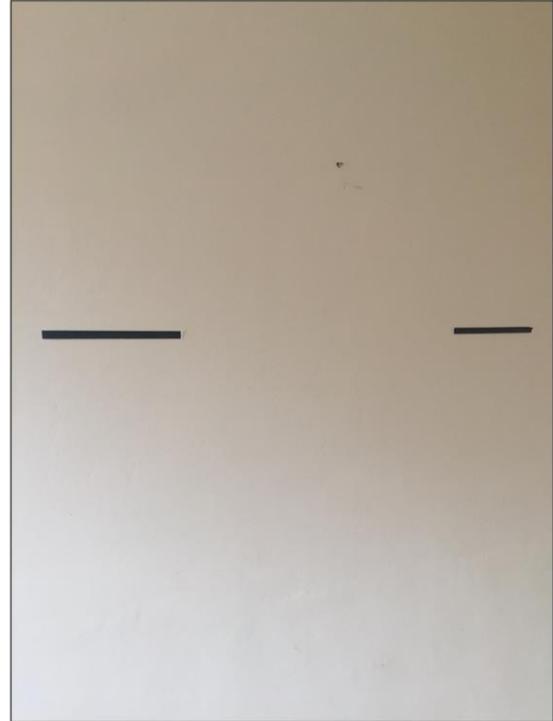
Participant's view   Agent's view

Figure 1. The agent view image was never shown to participants. It was a photo taken by the agent where she stood in the 'Participant's view' image, and demonstrates that, visually, the line closest to the agent appeared approximately twice as long as the line furthest from the agent.

Interpersonal Reactivity Index (IRI). Following the VPT task, participants completed the Interpersonal Reactivity Index (IRI: Davis, 1983). The IRI is a 28-item multiple choice questionnaire, consisting of four seven-item subscales (Fantasy, Perspective Taking, Empathic Concern, and Personal Distress). Higher scores on the Perspective Taking subscale have previously been associated with weaker egocentric biases in perspective taking in both belief-reasoning tasks (Meert et al., 2017) and VPT (Bukowski & Samson, 2017). The IRI was included in the study for two reasons. First, it provided a measure of cognitive and affective perspective taking that itself might correlate with performance on the VPT task.

Second, it provided a means of checking that any correlations between the VPT Ratio and Art Experience were not confounded with increased cognitive and/or affective perspective taking in those with higher Art Experience scores. Of particular interest in the present study were two of the IRI subscales. The first, Perspective Taking, is the tendency to adopt the psychological point of view of others. An example question from this subscale is: "I sometimes find it difficult to see things from the 'other guy's' point of view". The second, Empathic Concern, pertains to feelings of sympathy and concern for others, and measures affective empathy. An example question from this subscale is: "I often have tender, concerned feelings for people less fortunate than me." The Fantasy and Personal Distress subscales were retained for exploratory analyses but were judged a priori to be less relevant to VPT.

Art Experience. The Art Experience survey followed the IRI and was loosely based on the 'Art Experience' component of the Assessment of Art Attributes (AAA) devised by Chatterjee et al. (2010). The form it took in the present study was a short, three-question survey with a multiple-choice response. The questions were: 1) "How many studio art classes have you taken which were not compulsory (that is, you were not made to do them at school but chose to do them?)"; 2) "How many art theory or aesthetics classes have you taken which were not compulsory (that is, you were not made to do them at school but chose to do them?)"; and 3) "In the average week how many hours do you spend making visual art?". The multiple-choice response options for all three questions were "1","2", "3", "4", "5", and "6 or above", and these were summed to create an Art Experience score which ranged from 0-18. The questions were designed to capture participants' knowledge of visual arts and their propensity to learn and practise visual arts, i.e. , their enthusiasm and engagement with visual arts, through the measurement of optional rather than compulsory school-aged activities.

Analyses

The VPT Ratio was zero-centred, such that indicating precisely the same length for each line would equal a VPT ratio of zero. A VPT ratio of 0.2, for example, would indicate a judgment that the line closest to the agent appeared 20% longer to the agent than the line furthest from the agent. Conversely, a negative VPT ratio of -0.2 would indicate a judgment that the line closest to the agent appeared 20% *shorter* than the line furthest from the agent. A ratio was also preferred because it captured the difference between length judgments of 20-40 and 80-100 (for example), which a simple difference score would render as the same outcome (20) without capturing the fact that the first judgments suggest one line is double the length of the other and the second judgments that one was only one quarter longer than the other.

For the IRI, scores for Perspective Taking and Empathic Concerns were calculated separately, and reverse coding applied where necessary. Scores were able to range from zero to 28, with the latter indicating greater cognitive and affective perspective taking respectively.

The Art Experience score was calculated as the sum of the three individual questions, where the lowest possible score was zero and the highest 18. Note that 6 was a ceiling score on each question, so scores of 18 could mask potentially higher scores. However, only two participants achieved this maximum score.

All analyses were conducted using JASP version 0.16.4 and are parametric where normality tests were non-significant, non-parametric otherwise. All Bayesian analyses use the default Cauchy .707 prior.

Results

The means, standard deviations, and ranges of the VPT Ratio, Art Experience total (and its subscales), and the four scales of the IRI are displayed in Table 1. The mean VPT

Ratio was negative across the sample as a whole and for Artists and Musicians separately, indicating even before inferential statistics that adults were generally unsuccessful in understanding that the line closest to the agent should appear longer.

Table 1. Means and Standard Deviations from the Visual Perspective Taking (VPT) task, Art questionnaire, and Interpersonal Reactivity Index (IRI).

			M	*SD*	Min, Max
Exps. 1a & 1b	VPT	VPT Ratio	-0.28	1.61	-8, +4
	Art	Art Experience	1.94	3.64	0, 18
		- Art classes	0.61	1.45	0, 6
		- Theory classes	0.34	1.07	0, 6
		- Hours making art	0.99	1.79	0, 6
	IRI	Perspective Taking	17.39	4.58	3, 27
		Empathic Concern	18.89	4.79	5, 28
Experiment 1b					
Artists	VPT	VPT Ratio	-0.23	0.95	-4.08, 1.03
	Art	Art Experience	4.22	4.73	0, 18
		- Art classes	1.00	1.78	0, 6
		- Theory classes	0.56	1.48	0, 6
		- Hours making art	2.67	2.54	0, 6
	IRI	Perspective Taking	17.30	4.08	8, 25
		Empathic Concern	19.04	4.02	10, 27
Musicians	VPT	Ratio	-0.54	1.10	-4, 0.78
	Art	Art Experience	0.44	1.12	0, 5
		- Art classes	0.19	0.56	0, 2
		- Theory classes	0.07	0.39	0, 2
		- Hours making art	0.19	0.40	0, 1
	IRI	Perspective Taking	17.63	4.21	3, 23
		Empathic Concern	17.15	5.55	5, 28

Table 2. Results of Spearman's Correlations for all data (Exps. 1 & 2 combined). Bayes Factor analyses are based on Pearson's r *analyses.*

	Art Experience	Perspective Taking	Empathic Concern
VPT Ratio	$\rho(140) = .01$ $p = .96,$ $BF_{10} = 0.122$	$\rho(140) = -.04$ $p = .61,$ $BF_{10} = 0.283$	$\rho(140) = .12$ $p = .17,$ $BF_{10} = 0.107$

Experiment 1b (Artists and Musicians). This analysis concerned only the subset of participants who took part in Experiment 1b, that is Artists and Musicians, and excluded the wider sample from Experiment 1a. As expected, Artists had higher Art Experience scores (*M* = 4.2, *Mdn* = 3, *SD* = 4.7) than Musicians (*M* = 0.4, *Mdn* = 0, *SD* = 1.1), $U(54) = 599, p < .001$. This confirmed that the Artists did indeed have more experience creating illustrations than the Musicians. Importantly, Artists and Musicians did not differ on the Perspective Taking subscale, $U(54) = 347, p = .77$, or Empathic Concern subscale, $t(52) = 1.431, p = .16$. This meant that any group differences in VPT performance would be unlikely to be driven by differences in cognitive or affective perspective taking, although again the underpowered design means this too should be interpreted with caution.

The VPT Ratio scores for Artists was zero or negative depending on the measure of central tendency (*M* = -0.23, *Mdn* = 0). This suggests that the Artists group failed to identify that the line closest to the agent would appear longer than the line closest to the agent; indeed, the mean score suggests they erroneously judged the line closest to the agent to appear 23% *shorter* to the agent than the line furthest from the agent. The VPT ratio scores

for Musicians was negative by both measures of central tendency (M = -0.54, Mdn = -1.13); the mean suggests that this group erroneously judged the line closest to the agent to appear 54% shorter to the agent than the line furthest from the agent. Thus, neither group successfully understood the agent's visual experience of the lines, and the comparison of these two groups was therefore determining if one group failed by a greater amount than the other. There was no significant group difference on VPT Ratio, $U(54)$ = 425, p = .3. The effect size of this non-significant difference, calculated from an independent samples t-test, was d = 0.275. Of the Artists, 37% judged the further line to be longer, 37% the closer line to appear longer, and 26% judged both lines to appear identical. Of the Musicians, 55.6% judged the further line to appear longer, 33.3% the closer line to appear longer, and 11.1% judged both lines to appear identical.

Experiments 1a & 1b combined. To maximise power, the data from the first sample, collected for Experiment 1a, was now combined with the data collected from the sample for Experiment 1b. This increased the N to 140. A power analysis using G*Power suggested approximately a 95% chance to detect a medium effect size using correlational analyses with this sample size.

Spearman's correlations found no evidence of a relationship between VPT Ratio and any of Art Experience, Perspective Taking, or Empathic Concern (see Table 2 and Figure 2). Bayes Factor analyses based on Pearson's r also favoured the null hypothesis, suggesting the data for all three correlations were at least three times as likely under the null hypothesis than the alternative, exceeding the minimum threshold for a meaningful null result (Dienes, 2014).

A Wilcoxon signed-rank test also found no significant difference between length judgments of the line closest to the agent (M = 45, Mdn = 45, SD = 27.3) and the line furthest from the agent (M = 46.5, Mdn = 45, SD = 25.6), $W(140)$ = 3491.500, Z = 0.515, p = .61. Of

the 140 participants, 40% produced a negative ratio (the belief that the line closest to the agent would appear shorter to the agent), 46% a positive ratio, and 14% judged both lines to appear the same length to the agent. A Bayesian analysis found that the data were eight times more likely under the null hypothesis that there is no difference, $BF_{10} = 0.125$. It is noteworthy that, like the analysis for Experiment 1b alone, the mean for the further line was greater than the mean for the closer line, indicating that there was a (non-significant) erroneous judgment that the line closest to the agent appeared shorter to the agent than the line furthest from the agent[1].

Subset of participants with Art Experience > 0.

Given the high number of zero scores on Art Experience, data from the subset participants who scored at least 1 out of 18 on Art Experience (61 participants) were analysed separately. This less-skewed distribution might reveal patterns that the large dataset could not, though with the disadvantage of a loss of power. Again, there was no evidence of a relationship between Art Experience and VPT Ratio, $p = .71$, $\rho(61) = .05$. Bayes Factor analyses, again based on Pearson's r, still found meaningful support for the null, $BF_{10} = 0.169$. Thus, the absence of a relationship between Art Experience and VPT Ratio in the larger dataset is unlikely to be explained by the high number of zero Art Experience scores in the full sample.

---

[1] There was also no evidence of a sex difference, $U(133) = 2389.500$, $p = .73$; men ($M = -0.46$), women ($M = -0.14$)
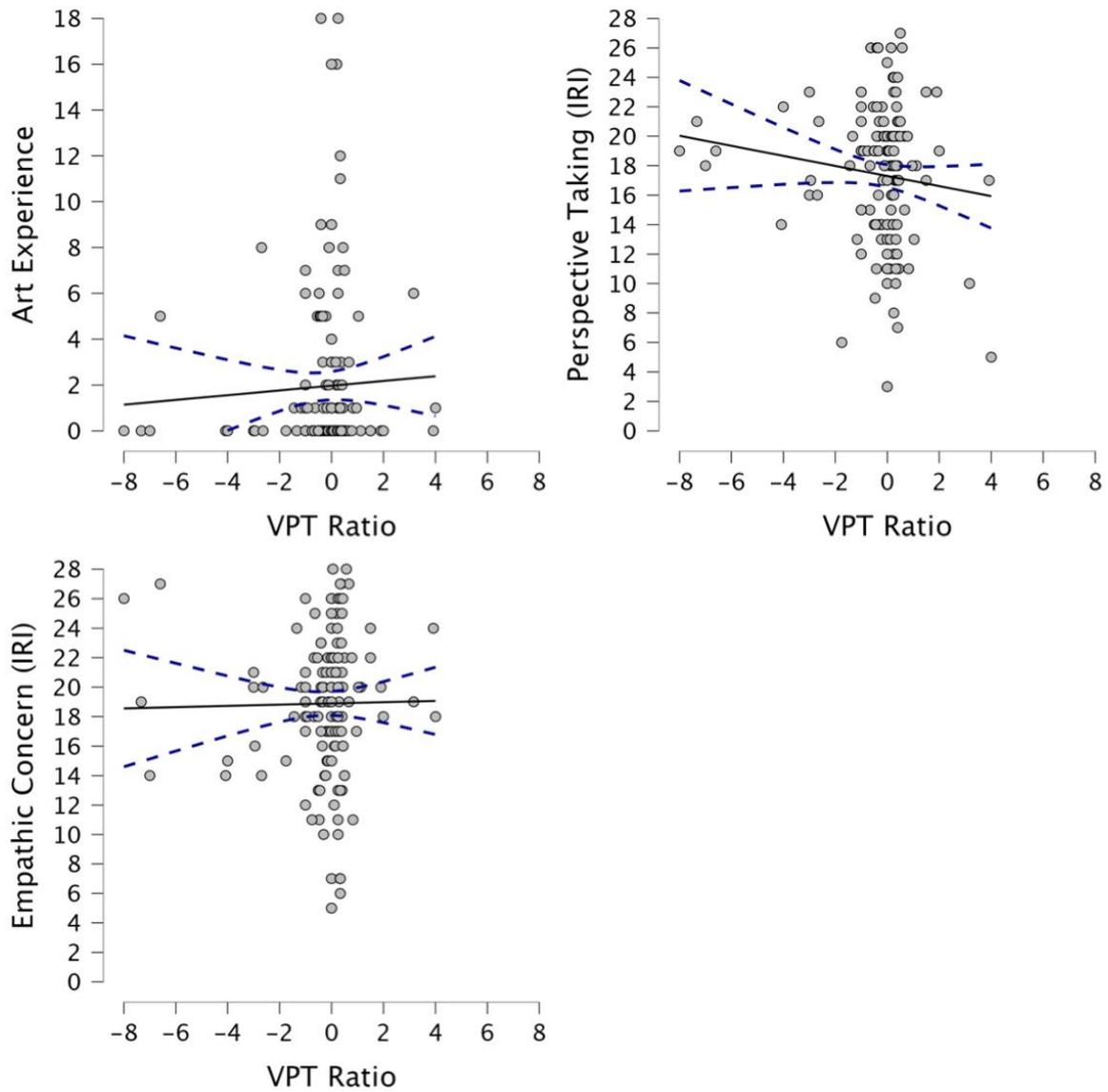
Figure 2. Scatterplots of relationships with VPT Ratio with 95% confidence intervals. A VPT Ratio of zero means both lines were judged to appear the same to the agent. Data from Experiments 1a and 1b combined (N = 140)

Discussion

The results of the present study were consistent with those of previous versions of this VPT task, which also found that adults failed to judge that the closer line to the agent would appear longer (Samuel, Hagspiel, Eacott, et al., 2021). The current study therefore replicates

that finding but also extends it. Visual artists should be much more adept at understanding that increased proximity correlates with increased relative size, and at accurately depicting three dimensional scenes in two dimensional formats. However, artists fared no better than non-artists. How can this be explained?

First, it could be that visual artists understand no better than non-artists the fact that closer objects take up a greater area in a flat image than identical but more distant objects. This seems unlikely, given that the accurate rendering of the 3D world into a flat, 2D image requires that objects appear naturally-sized relative to their location in the depicted space (Matthews & Adams, 2008; Perdreau & Cavanagh, 2011). There are artistic styles which do not seek to recreate perspective accurately, such as cubism, but it would be unusual for a visual artist not to have considered the real relationship between distance and size at some point in their education and practice, such as in still life drawing.

Second, it could be that representations generated in the task may not be image-like in the first place and therefore skill at "drawing" any such images is irrelevant. By this explanation, artists *are* better at depicting how objects further away seem smaller than non-artists, but this task did not entail depiction and therefore this skill was not useful. However, this alone could not to explain how artists were as likely to fall foul of the closer-as-shorter error also made by non-artists. That is, this knowledge "backstop" described in the Introduction would still need to fail for artists to perform like non-artists.

Third, it could be that visual artists *are* more adept at rendering the relationship between size and distance, and *do* understand that closer objects appear (visually) larger, at least in the context of their own perception, but do not extend this knowledge and experience to representations of *others'* perceptions. In essence, it could be that visual artists do not conceive of a VPT problem as a problem related in any way to their expertise as artists. This would mean that artists would effectively behave as non-artists in VPT. This would be a more

complete explanation of artists' failure to outperform non-artists in this task than either of the previous candidate explanations. This account does not require that artists are no better at depicting the world than non-artists and does not require that artists are unaware that closer objects generate larger retinal images. It only requires that artists do not connect a VPT problem with the domain of visual art. This therefore seems simultaneously the fullest and most parsimonious account of results.

Note that this disconnect between art and VPT does not imply that individual differences in observers cannot influence VPT performance; there are many reports that they do in both adults and children, including (but not limited to) observers' social class (Dietze & Knowles, 2020), bilingualism (Fan et al., 2015; Goetz, 2003), schizotypy (Langdon & Coltheart, 2001), executive functioning (Lin et al., 2010; Wardlow, 2013), and culture (Wu & Keysar, 2007). Rather, it appears that *some* individual experiences are not connected with VPT. Art experience would seem to be one such candidate. One potential account for this is that experiences that improve VPT performance may need to do one or both of 1) influence one's skill or propensity to take others' perspectives more broadly, or 2) influence the processing efficiency of VPT. The former has been invoked in explanations of how factors such as social class and bilingualism might enhance perspective taking, and the latter serves to explain a role of executive functioning. Art experience would not appear a priori to fall into either category. Indeed, the results of the IRI, which showed no advantage of visual artists over musicians in either perspective taking or empathic concern, provides evidence against art experience falling into the first category[2].

---

[2] And across the sample as a whole, Spearman's correlations found no evidence of a relationship between Art Experience scores and either the Perspective Taking or Empathic Concern scores (both ps > .1).

More broadly, these results serve to make clearer still that adults are poor visual perspective takers. This is problematic for theories of VPT that posit the spontaneous generation of mental images that simulate others' perceptions (Ward et al., 2019; Ward et al., 2020). As discussed elsewhere (e.g. Cole & Millett, 2019; Cole et al., 2020; Cole et al., 2022), such theories also do not explain why adults believe the line closest to the agent should appear *shorter*, (Samuel, Hagspiel, Eacott, et al., 2021), imagine others see things that they do not (Samuel et al., 2022), and resist encouragement to represent even their own visual experience as a flat image (Samuel, Hagspiel, Cole, et al., 2021). Indeed, perhaps the strongest evidence against such theories is the sheer inaccuracy of the majority of participants on VPT tasks that rely on understanding purely perceptual differences between perspectives.

Perhaps the most surprising feature of the results from this particular paradigm both in the past and as replicated here is that many adults judge that the item furthest away will appear the larger of the two. One can understand, for example, a judgment that the lines appear identical to the agent regardless of where they are viewed from, because it is true that they remain identical in absolute size. It may also be difficult to identify *how much* longer the closer line appears. However, there are no obvious circumstances whereby the further away something is from someone, the *larger* it should appear to them, except through the application of an erroneous rule. That approximately as many participants commit this "reverse" error as answer accurately suggests that adults, visual artists or not, are poor visual perspective takers. Indeed, the proportion of participants who correctly judged the closer line to appear longer did not exceed 46% in the full sample (and even smaller, 37%, in the artists alone). The best explanation for this range of results would therefore appear to be the same as the explanation offered originally, namely that adults apply idiosyncratic, intuitive, and sometimes erroneous folk optic rules to this VPT problem. One such rule could be that objects further away are "stretched" by the visual system to compensate for their loss in size.

The results of the present study suggest that even visual artists fall back on folk optic strategies in VPT.

This study is subject to a number of limitations. The first concerns the relative dearth of participants in the sample who had experience with the visual arts beyond that required in the course of compulsory school-aged education. This was addressed with two separate analyses: one with the subset of participants excluding those with an Art Experience score of zero, and one with a group of professional illustrators. In each case the result with the wider sample was reproduced. The illustrators even registered a mean *negative* VPT ratio. The second is that this was an exploratory rather than confirmatory study, and conclusions must be tempered as a result. A pre-registered follow-up is usually advisable in such instances, particularly if statistical power is a concern, and this was considered as a next step. However, the exceedingly small effect sizes, the favouring of the null in Bayesian tests, and crucially the fact that results frequently patterned in the *opposite* direction from what was hypothesised (including with the professional illustrators), suggests that further tests would similarly fail to reveal meaningful support for the experimental hypothesis.

In conclusion, increased experience of the creation of visual arts is unrelated to accuracy in a VPT task where an understanding of perspective (nearer = larger; further = smaller) is necessary for accurate performance. These results suggest that VPT tasks which pivot on visual rather than categorical differences between the self- and other-perspective are performed similarly by artists and non-artists alike, suggesting that even when artistic experience could potentially aid performance it may not be considered relevant to VPT.

References

Bertamini, M., Latto, R., & Spooner, A. (2003). The Venus effect: people's understanding of mirror reflections in paintings. *Perception, 32*(5), 593-599.

Bertamini, M., & Soranzo, A. (2018). Reasoning about visibility in mirrors: A comparison between a human observer and a camera. *Perception, 47*(8), 821-832.

Bukowski, H., & Samson, D. (2017). New insights into the inter-individual variability in perspective taking. *Vision, 1*(1), 8.

Capozzi, F., Cavallo, A., Furlanetto, T., & Becchio, C. (2014). Altercentric intrusions from multiple perspectives: Beyond dyads. *PloS one, 9*(12), e114210.

Chatterjee, A., Widick, P., Sternschein, R., Smith, W. B., & Bromberger, B. (2010). The assessment of art attributes. *Empirical Studies of the Arts, 28*(2), 207-222.

Cole, G. G., & Millett, A. C. (2019). The closing of the theory of mind: A critique of perspective-taking. *Psychonomic bulletin & review*, 1-16.

Cole, G. G., Millett, A. C., Samuel, S., & Eacott, M. J. (2020). Perspective taking: In search of a theory. *Vision, 4*(2), 30.

Cole, G. G., Samuel, S., & Eacott, M. J. (2022). A return of mental imagery: The pictorial theory of visual perspective-taking. *Consciousness and cognition, 102*, 103352.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a

multidimensional approach. *Journal of personality and social psychology, 44*(1), 113.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in*

Psychology, 5, 781.

Dietze, P., & Knowles, E. D. (2020). Social Class Predicts Emotion Perception and

Perspective-Taking Performance in Adults. *Personality and social psychology*

bulletin, 0146167220914116.

Fan, S. P., Liberman, Z., Keysar, B., & Kinzler, K. D. (2015). The exposure advantage: Early

exposure to a multilingual environment promotes effective communication.

Psychological Science, 26(7), 1090-1097.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge

about visual perception: Further evidence for the Level 1–Level 2 distinction.

Developmental Psychology, 17(1), 99-103.

Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Bilingualism:*

Language and Cognition, 6(1), 1-15.

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking:

embodied transformation versus sensorimotor interference. *Cognition, 114*(1), 72-88.

Langdon, R., & Coltheart, M. (2001). Visual perspective-taking and schizotypy: evidence for a simulation-based account of mentalizing in normal adults. *Cognition, 82*(1), 1-26.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of experimental social psychology, 46*(3), 551-556.

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development*, 357-366.

Matthews, W. J., & Adams, A. (2008). Another reason why adults find it hard to draw accurately. *Perception, 37*(4), 628-630.

Meert, G., Wang, J., & Samson, D. (2017). Efficient belief tracking in adults: The role of task instruction, low-level associative processes and dispositional social functioning. *Cognition, 168*, 91-98.

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception & psychophysics, 68*(2), 327-337.

Moll, H., & Kadipasaoglu, D. (2013). The primacy of social over visual perspective-taking. *Frontiers in human neuroscience, 7*, 558.

Perdreau, F., & Cavanagh, P. (2011). Do artists see their retinas? *Frontiers in human neuroscience, 5*, 171.

Perdreau, F., & Cavanagh, P. (2013). Is artists' perception more veridical? *Frontiers in neuroscience, 7*, 6.

Pylyshyn, Z. (2001). Is the imagery debate over? If so, what was it about. In E. Dupoux (Ed.), *Language, brain, and cognitive development: essays in honor of Jacques Mehler. MIT Press, Cambridge, MA* (pp. 59-83).

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological bulletin, 80*(1), 1-24.

Samuel, S., Cole, G. G., & Eacott, M. J. (2023). It's not you, it's me: a review of individual differences in visuospatial perspective taking. *Perspectives on Psychological Science, 18*(2), 293-308.

Samuel, S., Eacott, M. J., & Cole, G. G. (2022). Visual perspective taking without visual perspective taking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(7), 959.

Samuel, S., Hagspiel, K., Cole, G. G., & Eacott, M. J. (2021). 'Seeing'proximal representations: Testing attitudes to the relationship between vision and images. *PloS one, 16*(8), e0256658.

Samuel, S., Hagspiel, K., Eacott, M. J., & Cole, G. G. (2021). Visual perspective-taking and image-like representations: We don't see it. *Cognition, 210*, Article 104607.

Ward, E., Ganis, G., & Bach, P. (2019). Spontaneous vicarious perception of the content of another's visual perspective. *Current Biology, 29*(5), 874-880. e874.

Ward, E., Ganis, G., McDonough, K. L., & Bach, P. (2020). Perspective taking as virtual navigation? Perceptual simulation of what others see reflects their location in space but not their gaze. *Cognition, 199*, 104241.

Wardlow, L. (2013). Individual differences in speakers' perspective taking: The roles of executive control and working memory. *Psychonomic bulletin & review, 20*(4), 766-772.

Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. *Psychological Science, 18*(7), 600-606.

Zhao, X., Cusimano, C. J., & Malle, B. F. (2015). In Search of Triggering Conditions for Spontaneous Visual Perspective Taking. CogSci,