Objective Versus Subjective Measures of Face-Drawing Accuracy and Their Relations With Perceptual Constancies

Justin Ostrofsky
The Richard Stockton College of New Jersey

Dale J. Cohen
University of North Carolina, Wilmington

Aaron Kozbelt
Brooklyn College and the Graduate Center of the City University of New York

We investigated spatial errors nonartists make when drawing a face and the relationships between such errors with measures of perceptual constancies. Participants completed an observation-based free-hand drawing of a face, plus shape and size constancy tasks. Drawings were objectively measured with respect to errors in reproducing spatial relations among facial features as well as subjectively assessed using independent judges’ Likert scale-based holistic accuracy ratings. Results revealed systematic (rather than random) errors in the spatial relations between facial features. Further, although holistic accuracy ratings were negatively correlated with shape and size constancy errors, only some objectively measured spatial drawing errors were reliably correlated with the constancy measures. This suggests that holistic accuracy measurements may be too simplified for understanding the relationship between drawing accuracy and performance in nondrawing perceptual tasks, and that objective accuracy measures represent a useful complementary index of performance.

Keywords: face drawing, perceptual constancies, objective measurements

Empirical investigations into observational drawing behaviors have become a burgeoning area of study in cognitive psychology. Typically the aim of such investigations is to determine what psychological processes contribute to the widespread prevalence of errors in realistically reproducing an external model stimulus. Within these studies, participants are frequently asked to copy photographs of faces, most commonly in frontoparallel orientation (Brodie, Wyatt, & Waller, 2004; Cohen, 2005; Cohen & Bennett, 1997; Cohen & Earls, 2010; Cohen & Jones, 2008; Costa & Corazza, 2006; Freeman & Loschky, 2011; Hayes & Milne, 2011; Kozbelt, 2001; Kozbelt, Seidel, ElBassiouny, Mark, & Owen, 2010). As with the drawings of other categories of objects, these studies have demonstrated substantial individual variability in accuracy, with some reporting the unsurprising finding that trained artists draw faces more accurately than nonartists.

A major limitation of this research concerns how drawing accuracy has been measured. Potential errors in face drawing are heterogeneous: individuals could err in reproducing the shape or relative size of individual facial features, the spatial dimensions and relations between features, or other aspects of the face such as cues to individual identity or emotional expression, and such errors may be largely independent of one another. Thus, the accuracy of face drawings is a complex multidimensional variable. However, the method of measuring face drawing accuracy used in most past research typically not reflected this; virtually all relevant studies have used uni-dimensional holistic subjective accuracy ratings, where independent judges view the model stimulus and each face drawing together, and then provide a single Likert-scale accuracy rating—though occasionally judges make separate Likert-scale ratings on the accuracy of isolated facial features and spatial relations between features (see Cohen & Earls, 2010, Cohen & Jones, 2008, Experiment 4, for details). Such measurement methods allow assessment of how others holistically perceive the accuracy of a drawing, but do not inform which specific aspects of a face are drawn with a high versus low degree of accuracy.

The near-exclusive use of subjective accuracy ratings in the research literature has resulted in major gaps in our knowledge about face drawing performance. Most notably, the nature of errors most individuals make when drawing faces is simply not empirically well-established. For instance, are some kinds of drawing errors more frequent (or serious) than others? Are the errors random or systematic? How do various types of drawing errors relate to subjective accuracy ratings or to performance on other perceptual tasks? Such basic questions remain unanswered, particularly among adult nonartists drawing from observation. How-
ever, a few studies using objective measures of face drawings by adult artists or children provide some clues. For instance, two studies of adult artists drawing from observation (Costa & Corazza, 2006; Hayes & Milne, 2011) have revealed some systematic biases, such as the eyes being drawn too round and too close together, the mouth being drawn too round and too close to the chin, and the shape of the head being drawn too narrow. Another study (McManus, Chamberlain, Christopherson, Prince-Mobbs, Robinson, & Stelk, 2012) assessed the types of errors young children and nonartist adults make when drawing self-portraits from memory, finding systematic biases to draw the eyes too far up the length of the head and the shape of the head too round. Besides its basic scientific interest, developing objective measures of face drawing accuracy and extending earlier findings to determine their robustness would have obvious relevance to art education: if art educators are explicitly aware of the stereotyped patterns of errors most nonartists make when drawing a face, instruction could be effectively targeted to correct for these biases.

The use of objective measures of face drawing accuracy would also inform key theoretical issues bearing on the psychology of realistic drawing skill, such as the oft-discussed relation between drawing accuracy and aspects of perceptual processing (see Cohen & Bennett, 1997; Cohen & Jones, 2008; Ostrofsky, Kozbelt, & Seidel, 2012) For instance, are particular aspects of face drawing accuracy meaningfully associated with performance on perceptual tasks assessing phenomena such as shape or size constancy? Most relevant here is a study reported by Cohen and Jones (2008), who found that subjective face drawing accuracy ratings were predicted by the degree to which individuals experience the shape constancy effect.1 Because individuals who experienced more errors in perceiving the stimulus in the shape constancy task tended to produce drawings that were later rated as less accurate than those who experienced less perceptual errors of shape, these findings were interpreted to be consistent with the misperception theory of drawing accuracy (Cohen & Bennett, 1997; Cohen & Earls, 2010). This theory proposes that observational drawing errors are largely caused by the fact that the visual system routinely makes complex computations on visual input, which often results in transformed perceptual representations of the retinal pattern of stimulation. Such computations aim to infer the actual structure of the objects that are being perceived rather than consciously reproducing a particular pattern of retinal stimulation. In this view, drawing errors arise because individuals attempt to draw the transformed perceptual product rather than the veridical pattern of light stimulating the retina. It is presently unclear which elements of face drawing accuracy are related to basic perceptual processes because the subjective drawing accuracy ratings generally used to assess drawing accuracy do not provide information about the specific types of errors present in a drawing. As such, it is unclear whether perceptual processes relate to all or only some elements of face drawing accuracy. This is unfortunate, as identifying specific errors in face drawing would provide a richer understanding of the relationship between perception and drawing performance.

The present study aims to begin the process of filling in these gaps in our knowledge relating to drawing performance. We developed a novel system of quantifying errors present in observation-based face drawings produced by a sample of non-artists. We focused on measuring errors of drawing different aspects of the spatial dimensions and relations between the features constituting a face. We specifically focused on measuring the spatial accuracy of face drawings for numerous reasons. First, although there are other important aspects of face drawing ability besides the relative spatial positioning of features (e.g., accurately reproducing the features themselves, such as the shape of an eye), objective measurements of spatial accuracy are relatively easy to operationally define and measure, in contrast to the accuracy of individual features. Second, the ability to perceptually recognize faces relies strongly on the processing information about the relative spatial positioning of features (Rotshtein, Geng, Driver, & Dolan, 2007; Tanaka & Sengco, 1993), suggesting that accuracy judgments of face drawings might likewise rely on the basic spatial relationships between facial features. Finally, shape constancy errors have been shown to have a stronger relationship with perceived spatial drawing accuracy than with perceived featural or overall drawing accuracy (Cohen & Jones, 2008, Experiment 4). Thus, in the extant published research on face drawing, the reproduction of spatial relationships between features appears to be related to individual differences in perceptual processing more strongly than other aspects of face drawing.

In this study we aim to answer three questions. First, are nonartists’ spatial drawing errors random, or are there systematic biases? There are at least two reasons to expect some systematic biases in observational drawings of faces derived from review of the psychology and art education literature. As mentioned previously, McManus and colleagues (2012) found that there are numerous systematic biases present in self-portraits produced from memory that are present throughout childhood development and that persist into adulthood (e.g., head drawn too round and eyes drawn too far up the length of the head). As to whether such systematic spatial drawing biases in nonartists’ face drawings are specific to memory-based drawings or whether such biases generalize to observation-based drawings of a directly perceivable face is open to question. Second, upon review of the art education literature, one repeatedly finds explicit schematic guidelines as to how to draw the relative spatial proportions of facial features correctly (e.g., Edwards, 2012; Hamm, 1963; Hogarth, 2002; Kraavanger, 2005; Okabayashi, 2009). Sometimes, such instruction is presented alongside warnings to avoid generating specific spatial errors, such as drawing the shape of the head as too circular (Hamm, 1963) and drawing the eyes too far up the length of the head (Edwards, 2012; Okabayashi, 2009). Such points suggest that some systematic errors are commonly made by novice drawers,
although such anecdotal speculations have thus far not been subjected to strong empirical evaluation.

The second question this study aims to answer is whether subjective and objective measures of drawing accuracy are related to one another. Despite the common use of subjective ratings for assessing drawing accuracy, we lack a strong understanding of what types of drawing errors influence accuracy judgments. Although this study by no means intends to be an exhaustive investigation of the types of face drawing errors that influence subjective accuracy ratings, we do anticipate that many objectively measured errors in the drawing of spatial relationships between facial features will be negatively correlated with subjective accuracy ratings. As mentioned above, a subjective accuracy rating is a judgment of how well a drawing represents a recognizable depiction of a model stimulus. Individuals appear highly sensitive to changes in the vertical positioning of the eyes and mouth and the horizontal distance between the eyes for both novel and familiar faces (Haig, 1984; Hosie, Ellis, & Haig, 1988). Therefore, we expect that subjective accuracy ratings of drawings should also be sensitive to the accuracy in which the spatial relations between facial features are reproduced. What is more open to question is the strength of the relationship between subjective accuracy ratings and the objectively measured errors in drawing various spatial relationships between facial features.

The final question this study aims to evaluate is whether objectively measured spatial errors in face drawings are predicted by individuals’ experience of perceptual constancy errors. Despite findings that subjectively rated drawing accuracy is negatively correlated with shape (Cohen & Jones, 2008) and size constancy (Ostrofsky, Kozbelt, & Seidel, 2012), it is presently unclear as to what specific types of spatial errors produced while drawing a face are predicted by these perceptual constancy effects, if at all. To our knowledge, previous research has not investigated the link between face recognition and the degree to which individuals experience shape and size constancy effects, and thus, this aspect of the study is exploratory in nature. However, if errors in drawing the spatial relationships between facial features are caused by perceptual transformations of the stimuli (as the misperception theory proposes), we would expect that individuals who experience perceptual constancies to a strong degree would also tend to produce more spatial drawing errors than individuals who experience perceptual constancies to a comparatively weaker degree (Cohen & Jones, 2008). However, open to question at this point is the strength of such relationships and whether the accuracy in drawing all or only a select subset of spatial relationships between facial features are predicted by perceptual constancies.

**Method**

**Participants**

Forty-eight individuals (35 females, 13 males) participated in the drawing and perceptual constancy tasks, $M (SD) = 20.3 (1.5)$ years. Participants were recruited from the Brooklyn College Psychology Undergraduate Subject Pool and were provided partial course credit as compensation. No participants reported any formal training in drawing.

Twenty-four individuals with no formal training in drawing (18 females, 6 males) participated as independent judges to rate the accuracy of the drawings produced by the participants described above, $M (SD) = 23.4 (5.9)$. These participants were recruited and compensated in the same way as the individuals who participated in the drawing and perceptual constancy tasks.

**Materials and Procedures**

Participants were administered a free-hand drawing task, a shape constancy task, and a size constancy task, in that order.

**Free-hand drawing task.** Participants were asked to create a free-hand drawing of a gray-scale photograph of a woman’s face, measuring approximately $6.8 \times 10.1$” and printed on an $8.5 \times 11$” sheet of white paper in portrait orientation (see Figure 1). Participants were provided with an $8.5 \times 11$” sheet of white paper and a sharpened pencil with eraser to create a drawing and were instructed to draw the photograph of the face as accurately as possible. They were told that they could use any drawing techniques they wished (except for tracing) and that their drawings would later be measured for accuracy. They were given a 15 min time limit and were warned when they had 5 min and 1 min remaining.

**Shape constancy task.** We used a modified version of the shape constancy task described in Cohen and Jones (2008). Participants were shown individual computerized presentations of images of target shapes, and were asked to indicate the shape they perceived by pointing to a single option on a printed response sheet with 21 shape options (see below for details). Two conditions were tested. In the depth cue condition, participants were shown individual stimulus presentations of four photographs of a rectangular window embedded in a brick wall seen at different viewing points, which cause the projective shape of the window to deviate from rectangularity to varying degrees. These viewing points deviated from the frontoparallel view at approximately 26, 52, 65, and 78 degrees. In the nondepth cue condition, participants were shown individual presentations of quadrilaterals composed of four black lines presented on a uniform white background. The shape of the four quadrilaterals perfectly matched the shape of the four windows used in the depth cue condition.

Participants used a response sheet to indicate their choice of the shape that matched each stimulus. The sheet consisted of 21 quadrilaterals composed of black lines printed on an $8.5 \times 11$” sheet of white paper in the portrait orientation. Four quadrilaterals perfectly matched the shape of the stimuli presented in the depth and nondepth cue conditions; an additional 17 quadrilaterals were also shown that were morphed in shape relative to the four target stimuli and which collectively subtended deviations from the frontoparallel view from 0 degrees (a perfect rectangle) to $\approx 84$ degrees, and changed shape in $\approx 4$ degree intervals. The shapes were

---

2 The use of a relatively large sample of nonartists to serve as judges of drawing accuracy has been strongly justified by Cohen (2005). Further, artists and nonartists have relatively high levels of agreement with each other as to what constitutes accurate versus inaccurate drawings (Kozbelt, Seidel, ElBassiouny, Mark, & Owen, 2010; McManus, Chamberlain, Loo, Rankin, Riley, & Brunswick, 2010). Thus, we feel that the results we obtained through the subjective accuracy ratings would not be radically different if we had decided to sample expert artists for the judgment task rather than nonartists.

3 See Cohen and Jones (2008) for stimuli and the response sheet. Note also that in the present study, no numbers were printed on the response sheet.
a shape that was more rectangular than the target, consistent with the classic shape constancy effect whereby the perceived shape of the stimulus is biased toward its known rectangular shape.

Size constancy task. We administered the size constancy task used in Ostrofsky, Kozbelt, and Seidel (2012). Participants saw two circles on the computer screen. The upper circle was always the target, and participants were instructed to use arrow keys on the computer keyboard to manipulate the size of the lower circle to match the size of the target. Participants were explicitly instructed to focus on matching the actual size of the circles—that is, if they were measured on the computer screen—rather than their interpretation of their size.

Two conditions were tested. In the depth cue condition, the circles were shaded to suggest spherical forms and were presented against a textured, converging perspective background to give the appearance that the upper target sphere was more distant than the lower manipulated sphere. In the nondepth cue condition, both circles were shown in a uniform shade of gray matching the overall value of the spheres in the depth condition. The background likewise maintained the same contrast of light and dark and included a similar texture as the depth condition; however, no depth cues were present.

Each condition was tested as a separate block: 50 trials in the depth cue condition and 25 trials in the nondepth cue condition, with condition order counterbalanced across participants. To facilitate analyses, in each condition, the target was always one of five standard sizes (156, 208, 260, 212, and 364 pixels in diameter)—10 trials each for depth condition and five trials each for nondepth condition—within each block, presented in a random order). On each trial, error scores were computed based on the diameter of the two circles using the number of pixels as the unit of measure—specifically, by dividing the diameter of the manipulated circle by the diameter of the target circle. A value of 1 indicates a perfect match in size between the two circles. An error greater than 1 indicates that the participants manipulated the size of the lower circle as larger than the target size of the upper circle, indicating that they perceived the size of the target upper circle to be larger than it actually was (reflecting the size constancy effect).

Subjective measure of drawing accuracy in the free-hand drawing task. We used two methods of measuring accuracy of the face drawings. First, we used the traditional method of subjective accuracy ratings. Here, 24 independent judges viewed the model photograph of the face and each drawing individually and provided a single 20-point Likert-scale response indicating their assessment of the accuracy of each drawing, with 20 representing the highest accuracy. Each judge was instructed to provide this...
numerical rating with respect only to the realistic accuracy of the drawing and not rate the drawings based on any other criteria, like aesthetic or creative factors. To control for any idiosyncratic biases in the judges’ use of the 20 point scale, we transformed each judges’ set of ratings into z scores. Interjudge agreement was high (Cronbach’s α = .972), so z scores were averaged across judges to create a single subjective accuracy rating score for each drawing.

Objective measures of drawing accuracy in the free-hand drawing task. Next, we made 12 spatial measurements (in centimeters) of the model face photograph and each drawing (see Figure 2 for a representation of the 12 different spatial measurements. A through L). We measured: (a) the length of the head from the top of the head (including hair) to the bottom of the chin, (b) the width of the face (with landmark points being at the point of the image where it appeared that the upper part of the ear connected to the side of the face), (c) the vertical distance from the top of the head to the middle of the eye-line (if the eye-line was not perfectly horizontal, the vertical distance between the top of the face and the midpoint between the two eyes was measured), (d) the distance between the two outer corner of the eyes, (e) the diagonal distance between the outer corner of the left eye (from the observer’s perspective) and the center of the bottom of the lower lip, (f) the diagonal distance between the outer corner of the right eye (from the observer’s perspective) and the center of the bottom of the lower lip, (g) the width of the eyes (the width of both eyes were measured and averaged to create one width measurement), (h) the interocular distance between the two inner corners of the eyes, (i) the width of the nose, (j) the horizontal distance between the outer corner of the left eye and the left side of the face (from the observer’s perspective), (k) the horizontal distance between the outer corner of the right eye and the right side of the face (from the observer’s perspective), and (l) the vertical distance between the center of the bottom of the lower lip and the bottom of the chin. We then calculated 13 ratios that quantified most of the spatial relations that Hamm (1963) observed in the model, I/B ratio: \( t(45) = -3.80, p < .01 \), (c) the difference between the outer corners of the left and right eyes’ diagonal distance to the bottom of the lower lip (E/F ratio), \( \beta = -0.497, t = -4.87, p < .01 \), (d) the distance between the left eye and the left side of the face (J/B ratio), \( \beta = -0.289, t = -2.95, p < .01 \), and (e) the difference in width between the left and right eyes (IG(L) – G(R))/G(ratio), \( \beta = -0.479, t = -4.44, p < .001 \). The drawing errors of the remaining eight spatial relationships were not reliably different from 1 (all \( p > .05 \)).

Relationship Between Objective and Subjective Measures of Drawing Accuracy

Next, we wished to determine whether the objective measures of spatial drawing accuracy were related to the independent judges’ subjective ratings of drawing accuracy. We recalculated the objective spatial drawing errors as the absolute difference between the spatial relation ratio value of the drawing and the model for each of the 13 spatial relation ratios. Then, we conducted a multiple regression analysis, aiming to predict subjective accuracy ratings by the absolute errors of the 13 spatial relation ratios. This analysis resulted in a significant model, \( F(13, 32) = 8.98, p < .01 \), adjusted \( R^2 = .785 \), indicating that the collection of the objective spatial drawing errors, as a whole, are capable of predicting how independent judges perceived the accuracy of the drawings.

Table 3 displays more specific information pertaining to how the drawing errors of each individual spatial relationship are related to the subjective accuracy ratings. Five spatial drawing errors significantly predicted subjective accuracy ratings in the negative direction (larger errors predicted lower accuracy ratings) at the .01 α level. They were: (a) the vertical position of the eyes on the length of the head (C/A ratio), \( \beta = -0.499, t = -3.52, p < .01 \) (b) the vertical position of the mouth on the length of the head (L/A ratio), \( \beta = -0.329, t = -3.32, p < .01 \) (c) the difference between the outer corners of the left and right eyes’ diagonal distance to the bottom of the lower lip (E/F ratio), \( \beta = -0.497, t = -4.87, p < .01 \), (d) the distance between the left eye and the left side of the face (J/B ratio), \( \beta = -0.289, t = -2.95, p < .01 \), and (e) the difference in width between the left and right eyes (IG(L) – G(R))/G(ratio), \( \beta = -0.479, t = -4.44, p < .001 \). The drawing errors of the remaining eight spatial relationships were not reliably predictive of the subjective accuracy ratings.

Shape Constancy

Average errors in the shape constancy task are displayed in Figure 3a. We wished to determine whether participants reli-

6 Only drawings that included all facial features (both eyes, nose, and mouth) were included in the analysis. This resulted in having to discard data from two participants.
ably experienced the shape constancy effect. To determine this, we conducted a 2 (Stimulus Condition: Depth vs. Non-Depth cue conditions) × 4 (Target Shape: 26 vs. 52 vs. 65 vs. 78 degree frontal view deviation) repeated measures ANOVA testing for effects on shape matching errors. Because of a violation of the assumption of sphericity, we corrected the degrees of freedom using the Huynh-Feldt procedure to reduce potential inflation. We observed a reliable main effect of Target Shape, $F(2.03, 95.57) = 36.06, p < .001$, partial $\eta^2 = .43$, indicating differences in the size of error across target shapes. We also found a reliable main effect of Stimulus Condition, $F(1, 47) = 77.28, p < .001$, partial $\eta^2 = .62$, indicating that errors were larger in the depth cue condition relative to the nondepth cue condition. Analysis of condition means indicated that participants perceived the shape of the window to be more rectangular than the projective shape, reflecting participants’ experience of shape constancy. We also observed a reliable interaction between stimulus condition and target shape, $F(3, 141) = 8.95, p < .001$, partial $\eta^2 = .16$. This interaction was explored by quasi-$F$ tests, comparing errors between the depth and nondepth conditions.
Table 2  
M (SD) Values of the Spatial Relation Ratios and Error Ratios

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Spatial Relation</th>
<th>Error Ratio</th>
<th>% of Participants Erring in the</th>
<th>t(45)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B/A</td>
<td>.671 (.090)</td>
<td>1.111 (.151)</td>
<td>83</td>
<td>5.33***</td>
</tr>
<tr>
<td>C/A</td>
<td>.440 (.062)</td>
<td>.835 (.118)</td>
<td>96</td>
<td>9.47***</td>
</tr>
<tr>
<td>E/D</td>
<td>.952 (.093)</td>
<td>.969 (.097)</td>
<td>67</td>
<td>2.14*</td>
</tr>
<tr>
<td>F/D</td>
<td>.935 (.086)</td>
<td>.972 (.090)</td>
<td>72</td>
<td>2.07</td>
</tr>
<tr>
<td>E/F</td>
<td>.998 (.060)</td>
<td>.998 (.060)</td>
<td>50</td>
<td>0.25</td>
</tr>
<tr>
<td>G/B</td>
<td>.230 (.036)</td>
<td>1.009 (.156)</td>
<td>57</td>
<td>0.38</td>
</tr>
<tr>
<td>H/B</td>
<td>.284 (.052)</td>
<td>1.198 (.220)</td>
<td>87</td>
<td>6.11***</td>
</tr>
<tr>
<td>J/B</td>
<td>.118 (.040)</td>
<td>.840 (.285)</td>
<td>74</td>
<td>3.80***</td>
</tr>
<tr>
<td>K/B</td>
<td>.131 (.051)</td>
<td>.830 (.323)</td>
<td>67</td>
<td>3.57***</td>
</tr>
<tr>
<td>I/B</td>
<td>.246 (.040)</td>
<td>.904 (.149)</td>
<td>74</td>
<td>4.39***</td>
</tr>
<tr>
<td>L/A</td>
<td>.161 (.037)</td>
<td>1.227 (.285)</td>
<td>76</td>
<td>5.42***</td>
</tr>
<tr>
<td>(GL-GR)/G</td>
<td>.014 (.099)</td>
<td>.014 (.099)</td>
<td>47</td>
<td>0.96</td>
</tr>
<tr>
<td>(J-K)/B</td>
<td>-.013 (.060)</td>
<td>.745 (3.441)</td>
<td>50</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Note. Mean Direction of Error Ratio: >1 for Ratios B/A, G/B, H/B, and L/A; <1 for all others. Error Ratio = Drawing Ratio/Model Ratio.  
*Mean ratio error is reliably different from 1 at the .01 α level.  **Mean ratio error is reliably different from 1 at the .001 α level.

Table 3  
Regression Model Predicting Subjective Accuracy Ratings From Objective Spatial Drawing Errors

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Standardized coefficient</th>
<th>Unstandardized coefficient</th>
<th>SE</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>—</td>
<td>2.01</td>
<td>0.23</td>
<td>8.89***</td>
</tr>
<tr>
<td>B/A</td>
<td>.153</td>
<td>1.77</td>
<td>1.48</td>
<td>1.20</td>
</tr>
<tr>
<td>C/A</td>
<td>-.499</td>
<td>-6.79</td>
<td>1.93</td>
<td>-3.52**</td>
</tr>
<tr>
<td>E/D</td>
<td>.109</td>
<td>1.45</td>
<td>1.62</td>
<td>0.90</td>
</tr>
<tr>
<td>F/D</td>
<td>-.229</td>
<td>-3.13</td>
<td>1.82</td>
<td>-1.72</td>
</tr>
<tr>
<td>E/F</td>
<td>-.497</td>
<td>-10.60</td>
<td>2.18</td>
<td>-4.87***</td>
</tr>
<tr>
<td>G/B</td>
<td>-.005</td>
<td>-0.17</td>
<td>4.07</td>
<td>-0.04</td>
</tr>
<tr>
<td>H/B</td>
<td>.040</td>
<td>0.81</td>
<td>2.06</td>
<td>0.39</td>
</tr>
<tr>
<td>J/B</td>
<td>-.289</td>
<td>-8.93</td>
<td>3.03</td>
<td>-2.95**</td>
</tr>
<tr>
<td>K/B</td>
<td>-.040</td>
<td>-0.88</td>
<td>2.40</td>
<td>-0.37</td>
</tr>
<tr>
<td>I/B</td>
<td>.035</td>
<td>0.84</td>
<td>2.46</td>
<td>0.34</td>
</tr>
<tr>
<td>L/A</td>
<td>-.329</td>
<td>-9.33</td>
<td>2.81</td>
<td>-3.32***</td>
</tr>
<tr>
<td>(G/L-G(R))/G</td>
<td>-.479</td>
<td>-5.39</td>
<td>1.21</td>
<td>-4.44***</td>
</tr>
<tr>
<td>(J-K)/B</td>
<td>.104</td>
<td>2.07</td>
<td>2.13</td>
<td>0.97</td>
</tr>
</tbody>
</table>

*p < .05. ** p < .01. *** p < .001.

cue conditions at each target shape.7 Results indicated that errors in the depth cue condition were reliably larger than errors in the nondepth cue condition for each target shape (all p < .001). However, between-condition effects reliably differed across target shapes, being greatest for the 65 degree target shape and smallest for the 26 degree target shape.

Size Constancy

Average errors in the size constancy task are shown in Figure 3b. We wished to determine whether participants reliably experienced the size constancy effect. To determine this, we conducted a 2 (Stimulus Condition: Depth vs. Non-Depth cue conditions) × 5 (Target Size: 156 vs. 208 vs. 260 vs. 312 vs. 360 pixel diameter) repeated measures analysis of variance (ANOVA) testing for effects on size matching proportion errors. Because of a violation of the assumption of sphericity, we again corrected the degrees of freedom using the Huynh-Feldt procedure. We observed a reliable main effect of Target Size, F(2.75, 115.68) = 255.69, p < .001, partial η2 = .86, indicating differences in error size as a function of target size. We also found a reliable main effect of Stimulus Condition, F(1, 42) = 425.26, p < .001, partial η2 = .91, indicating that participants’ errors were reliably larger in the depth cue condition compared with the nondepth cue condition. Analysis of condition means indicated that participants manipulated the size of the bottom sphere to be larger than the target sphere (to compensate for the target sphere’s greater apparent distance), reflecting participants’ experience of size constancy. Additionally, a reliable interaction was observed, F(3.17, 133.15) = 164.62, p < .001, partial η2 = .80. This interaction was explored by conducting quasi-F tests, comparing the errors between the depth and nondepth cue condition for each target size.8 Results indicated that errors in the depth cue condition were reliably larger than in the nondepth cue condition (all p < .001). However, between-condition effects reliably differed across target sizes, being largest for the 156-pixel diameter target and progressively decreasing as the target increased in size.

Relationship Between Drawing Accuracy and Perceptual Constancies

Finally we wished to determine whether the magnitude of size and shape constancy effects were related to subjective and objective drawing accuracy measures. For each participant, we calculated the absolute value of the average shape and size matching errors separately for the depth and nondepth cue conditions of the task.

With respect to the subjective accuracy ratings, drawing accuracy was reliably correlated with errors made in the depth cue version of the shape matching task, r(44) = -.370, p < .05, but

7 Each quasi-F test comparison for the shape constancy task was evaluated with 1 between-condition df and 105.41 within-condition df.

8 Each quasi-F test comparison for the size constancy task was evaluated using 1 between condition df and 62.73 within-condition df.
not with errors made in the nondepth cue version, $r(44) = -.117, p = .44$. Similarly, with respect to the size matching task, drawing accuracy was reliably correlated with errors made in the depth cue version, $r(41) = -.488, p < .001$, but not with errors made in the nondepth cue version, $r(41) = -.089, p = .57$.

Finally, as a preliminary exploratory analysis, we assessed whether perceptual constancy effects were related to objectively measured spatial drawing errors. Correlations between the objective spatial drawing errors and the four perceptual task errors are displayed in Table 4. With respect to the objectively measured spatial drawing errors, errors made in the depth-cue version of the shape matching task were reliably correlated with the C/A ratio errors (vertical position of the eye line), $r(44) = .340, p < .05$, the E/F ratio errors (representing the difference in outer corner of eye-bottom of the lower lip distances between the left and right eyes), $r(44) = .309, p < .05$, the I/A ratio errors (the vertical position of the bottom lip), $r(44) = .306, p < .05$, and the I/B ratio errors (the width of the nose), $r(44) = .423, p < .01$. Errors made in the nondepth cue version of the shape matching task were reliably correlated with I/B ratio errors (interocular distance), $r(44) = .316, p < .05$, and curiously, were negatively correlated with the J/K ratio errors (the difference in distance between the outer-corner of the eyes and the side of face between the left and right sides), $r(44) = -.316, p < .05$. Additionally, errors made in depth cue version of the size matching task was reliably correlated with the E/F ratio errors, $r(41) = .314, p < .05$, and with the J/B ratio errors (the distance between the outer-corner of the left eye and the left side of the face), $r(41) = .424, p < .01$. Finally, errors made in the nondepth cue version of the size matching task was only reliably correlated with the B/A ratio errors (the shape of the face as measured by the face width-to-length ratio), $r(41) = .415, p < .01$. No other correlations were reliable at the .05 level.

All told, the pattern of mostly numerically positive (rather than negative) correlations displayed in Table 4 was reliable at the .05 level by the sign test on three of the four conditions: for size constancy depth, $p = .003$; for size constancy nondepth, $p = .097$; for shape constancy depth, $p = .038$; and for shape constancy nondepth, $p = .022$. This set of results suggests, albeit tentatively, that performance in the depth and nondepth conditions of constancy tasks may be associated with drawing skill, that more than just susceptibility to constancy effects (that are mainly evident in the depth conditions) may be required to explain individual differences in drawing performance—a point taken up in the Discussion.

**Discussion**

**Systematic Error Biases in the Spatial Relationships Between Drawn Facial Features**

The first major question we aimed to evaluate was whether errors nonartists make when drawing the spatial relationships between facial features from observation were systematically biased or random in nature. Our findings demonstrated that many errors in reproducing the relative spatial positioning of facial features were systematically biased. To summarize, we found that nonartists tended to draw the face as too circular, with the eyes too high and too far apart (and too close to the sides of the face), the nose too narrow, and the mouth too high. Such findings empirically support anecdotal testimonies of art instructors who propose that errors such as drawing the head too circular and the eyes too high up the length of the head are common production biases in nonartists’ face drawings (Edwards, 2012; Hamm, 1963; Okabayashi, 2009).

In the attempt to have nonartists draw the spatial relations between facial features more accurately, drawing instructors have traditionally provided explicit instructions relating to schematic rules that define the relative spatial positioning of features of the average face (Edwards, 2012; Hamm, 1963; Hogarth, 2002; Kraavanger, 2005; Okabayashi, 2009). Such a pedagogical strategy seems to be an effective method of improving nonartists’ drawing accuracy, as previous empirical research has demonstrated that providing nonartists with explicit drawing rules as how to draw common objects or features has the beneficial effect of increasing the accuracy to which they draw models from observation (e.g., Clare, 1983; Rand, 1973). However, these studies focused on the drawings of nonface objects produced by children. Future research may wish to investigate whether explicit instruction of schematic spatial rules of drawing faces such as those found in the “how-to” drawing manuals referenced above increases face drawing accuracy in adult nonartists.

The direction of the systematic patterns of error we observed raises questions relating to the generalizability of these systematic biases. Our task used an adult Caucasian female as the drawing model. Previous research has demonstrated sex- and race-based differences in the average relative spatial positioning of facial features. It is an open question whether the drawing biases we observed would generalize to drawings of male faces or faces of different races. Examining which types of spatial drawing biases of faces are universal versus specific to a particular faces is an open empirical question.

Our findings also raise questions that pertain to the sources of these drawing error biases. Although our data provide no conclusive information about the source of these drawing biases, one possible mechanism could be related to how attention is deployed during face perception. Consider, for example, our finding that 96% of nonartists in our sample drew the vertical position of the eye-line farther up the head than it was in the model, and, thus, attenuated the length of the forehead in their drawings. Previous research has revealed that, when perceiving a face, attention is deployed more frequently and for longer periods of time to the eyes than to the forehead region (Heisz & Shore, 2008; Nguyen, Issacowitz, & Rubin, 2009). If such an attentional bias is present while drawing, then the attenuation of the forehead and the up-shift of the vertical eye-position could be explained by participants ignoring spatial properties of the forehead and not fully attending to the task of accurately reproducing this region. Eye-tracking could be a valuable tool to inform this issue across facial features, as such interventions where participants are instructed to attend to particular spatial relationships, to see if this improves drawing accuracy. Previous research on the drawings of children have demonstrated that instructions simply to focus attention on particu-

---

9 For these analyses, $N = 13$, except for shape constancy depth, where $N = 12$ (because one correlation coefficient in that condition was zero).
ular properties of a model have a beneficial effect on observational drawing accuracy (Sutton & Rose, 1998), suggesting that the development of drawing skill requires, to some extent, the training of attention as to what features and spatial relationships are important to attend to for achieving the goal of creating an accurate depiction of a reproduced model (Ostrofsky & Kozbelt, 2011).

Another potentially fruitful approach to understanding drawing biases concerns the developmental trajectory of how the spatial relations of faces are drawn. As stated earlier in the article, McManus and colleagues (2012) assessed biases in the memory-based drawings of the spatial relations between facial features produced by children ranging in age from 3 to 10 years old in addition to adults. They found that very young children have strong biases to draw the head very round and to draw the eyes too far up the length of the head. These biases decreased in magnitude as the children got older, but were not found to be completely eliminated through development, as they were still observed in the drawings produced by the older children and adults, consistent with the biases we observed in our sample of adult nonartists. Thus, some of the systematic error biases we observed seem to be present very early in life. Future researchers may benefit from investigating potential reasons why these biases develop early in childhood, so that we can understand why they persist through development into adulthood.

### Relationship Between Objective and Subjective Measures of Face Drawing Accuracy

The second major question this study aimed to evaluate was whether subjective and objective measures of drawing accuracy were related to one another. The extent to which different types of objective drawing errors are related to subjective perceptions of drawing accuracy is currently not well-understood. Here, our data allowed a preliminary analysis of the extent to which different types of spatial errors in face drawings were associated with the degree to which the drawings were perceived as accurate renderings of the model face. We found that most of the objectively measured spatial drawing errors were negatively correlated with subjective ratings of drawing accuracy. These findings suggest that subjective judgments of face drawing accuracy are influenced to some degree by the precision to which the spatial relationships between facial features are reproduced in the drawings. Such findings are consistent with the face recognition literature, which has repeatedly demonstrated that face recognition processes are sensitive to the relative spatial positioning of facial features (e.g., Haig, 1984; Hosie, Ellis, & Haig, 1988; Rotshtein, Geng, Driver, & Dolan, 2007; Tanaka & Sengco, 1993).

It is clear, however, that subjective accuracy ratings are not solely influenced by how accurate the spatial relationships between facial features are reproduced: our data revealed only weak-to-moderate relationships between subjective and objective measures. One obvious aspect of face drawing that the objective measure does not take into account is how accurately individual facial features themselves are reproduced. Additionally, however, is the under-discussed issue of how accuracy judgments are affected by characteristics of the judges themselves. Different evaluative criteria may be adopted by different judges based on cultural and experiential variables. With respect to experiential variables, at least one study has demonstrated differences in evaluative judgments related to drawing expertise, as artists have been shown to evaluate the accuracy of limited-line tracings somewhat differently than nonartists (Kozbelt, Seidel, ElBassiouny, Mark, & Owen, 2010). It would be beneficial to the psychology of drawing for future research to investigate which objective aspects of drawing accuracy are related to subjective accuracy ratings of, say, artist versus nonartist judges.

### Relationship Between Drawing Accuracy and Perceptual Constancies

The last major question our study aimed to address was the relationship between drawing accuracy and individuals’ experi-

---

**Table 4**  
*Pearson Correlation Coefficients Between Spatial Drawing Errors and Shape and Size Matching Errors*

| Ratio         | B/A | C/A | E/D | F/D | E/F | G/B | H/B | J/B | K/B | I/B | L/A | (GL–GR)/G | (J–K)/B |
|---------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------------|----------|
| Size: Depth   | .09 | .20 | -.16| .04 | .31 | .17 | .23 | .42 | .24 | .24 | .17 | .15 | .20        |
| Size: Nondepth| .42 | .19 | -.04| .25 | .14 | .16 | .23 | .05 | .10 | -.11| .15 | -.13 | .09        |
| Shape: Depth  | .23 | .34 | -.07| .00 | .31 | -.03| .14 | .19 | .15 | .42 | .31 | .06 | .19        |
| Shape: Nondepth| .06 | .03 | .25 | .09 | -.13| .07 | .32 | .11 | .08 | .27 | .17 | .03 | -.32       |

*Note.* The values in the table are Pearson $r$ correlation coefficients. Size matching task analyses: $df = 41$. Shape matching task analyses: $df = 44$.  
* $p < .05.
ence of perceptual constancy effects relating to shape and size. As mentioned in the Introduction, the misperception theory of drawing accuracy posits that errors in drawing are caused by perceptual transformations that operate on the retinal image. One prediction derived from this proposition is that individuals capable of producing accurate drawings should be able to more accurately perceive the veridical properties of a visual stimulus relative to individuals who are not as capable of drawing accurately. This prediction has been supported in the past by studies that have reported negative correlations between subjectively judged drawing accuracy and the degree to which people experience shape constancy errors (Cohen & Jones, 2008, but see our Footnote 1) and size constancy errors (Ostrofsky, Kozbelt, & Seidel, 2012).

Our findings replicated these observations, as subjective ratings of drawing accuracy were negatively correlated with errors made in the depth-cue versions, but not in nondepth-cue versions, of the shape and size matching tasks. Such findings suggest that subjectively judged drawing accuracy appears to be related to perceptual constancy errors that are generated by concurrently processing depth cue information along with the shape and size information of the target stimuli.

Moving beyond subjective accuracy ratings, we also investigated how objectively measured spatial drawing errors might be related to errors made in the depth and nondepth cue conditions of the size and shape perceptual matching tasks. Because of the exploratory nature of this analysis, some caution is in order in interpreting the precise values of the observed correlation coefficients. However, some observations suggest something about the overall relationship between spatial drawing and perceptual encoding accuracy. First, along the lines of classic perceptual constancies, objective spatial drawing errors were generally related to errors in the depth cue conditions of the size and shape matching tasks. (If there was no relationship between perceptual constancy effects and objective spatial drawing errors, we would expect an even distribution of positive and negative correlation coefficients, but this was not what we observed: recall the sign test p values at the end of the Results section where the majority of spatial drawing errors were numerically positively correlated with errors in the depth conditions of the size and shape matching tasks.) This suggests that individuals who are more capable of overcoming perceptual transformations of visual input are able to more accurately depict visual information like the spatial relationships between facial features.

A second point, however, is that similar, though weaker, trends were also found in the nondepth conditions of the perceptual matching tasks, suggesting that the association between perceptual errors and drawing errors might hold even in the absence of contextual depth cues. This suggests that spatial drawing errors may be more generally related to perceptual encoding accuracy than in just overcoming perceptual constancies, and that additional aspects of attentional deployment and attentional capacity may need to be considered for a full explanation of skilled drawing (see Ostrofsky & Kozbelt, 2011; Ostrofsky, Kozbelt, & Seidel, 2012).

Finally, from a methodological perspective, these results suggest an important difference between how objective and subjective measures of drawing accuracy relate to some types of nondrawing performance measures. Particularly, the result that the nondepth cue conditions of the perceptual matching tasks were differentially related to objective and subjective measures of drawing accuracy suggest that holistic subjective measures of drawing accuracy may at times mask relationships that potentially exist between drawing and nondrawing performance variables. Because there are many studies that aim to determine how drawing accuracy is related to nondrawing performance measures (e.g., Cohen & Jones, 2008; Kozbelt, 2001; McManus, Chamberlain, Loo, Rankin, Riley, & Brunswick, 2010; McManus, Loo, Chamberlain, Riley, & Brunswick, 2011) our findings suggest that it may be beneficial to complement subjective and objective analyses of drawing performance to obtain a more complete understanding of how some nondrawing performance measures are related or unrelated to drawing accuracy.

More generally, it remains unclear why some of the positive correlations between particular spatial drawing errors and perceptual matching errors (in both the depth and nondepth cue conditions) were not statistically reliable. This could be related to issues of statistical power, small ranges of drawing error for particular spatial relationships, or an indication that particular spatial drawing errors are simply not actually related to perceptual matching performance. Thus, it is important to reemphasize the exploratory nature of these analyses. Our conclusions remain tentative, and they await further investigation to replicate and extend these preliminary findings.

**Conclusion**

The development of objective measures of drawing accuracy promises to enrich our understanding of the cognitive processes supporting the production of observational drawings. The challenge of future research will be to devise means of objectively measuring other aspects of drawing accuracy. To date, there have been cases of other drawing studies using other categories of objects as the model stimulus, such as plain angles and line drawings of bodies, that have successfully used objective measures of drawing accuracy to understand the psychological processes associated with drawing behavior (Carson & Allard, 2013; Tchalenko, 2009). We hope that a multimethod effort, as attempted here, will continue to illuminate this all-too-poorly understood topic.

**References**


Received July 12, 2013

Revision received May 14, 2014

Accepted June 19, 2014

**Correction to Beghetto (2014)**

In the article “Creative Mortification: An Initial Exploration,” by Ronald A. Beghetto (*Psychology of Aesthetics, Creativity, and the Arts*, 2014, Vol. 8, No. 3, pp. 266–276. http://dx.doi.org/10.1037/a0036618), there was an omission on p. 272. A footnote should have been included to the sentence in the first paragraph under the Coding Procedure subsection reading, “In total, nine codes were used in this study.” The footnote should have read, “Initially, 10 codes were used to classify the responses. One code, adjust expectations, was dropped because it did not perform well (i.e., low consistency and low frequency of use).”

http://dx.doi.org/10.1037/aca0000003