Observational Drawing Research Methods

Justin Ostrofsky

The Oxford Handbook of Empirical Aesthetics
Edited by Marcos Nadal and Oshin Vartanian

Subject: Psychology, Cognitive Psychology Online Publication Date: Aug 2020
DOI: 10.1093/oxfordhb/9780198824350.013.17

Abstract and Keywords

This chapter provides a critical review of methods used in psychological research focused on understanding observational drawing performance. First, observational drawing is defined and distinguished from other types of drawing behaviors. Then, the general questions investigated in observational drawing research are described, along with a review of the methods used to address them and a discussion of each method's general limitations. Next follows a description of basic drawing and tracing tasks administered in laboratory studies. Following this, subjective and objective measurement methods used to assess drawing performance (or accuracy) are reviewed in conjunction with a discussion of the strengths and weaknesses of each measurement method. The chapter concludes with a discussion of possible directions of future research, specifically by discussing (a) ways that the approach of measuring drawing performance can be improved, and (b) methods one can adopt to better understand the causal relationship between drawing skill and cognitive and perceptual processing abilities.

Keywords: Observational drawing, research methods, artists, perception, cognition

Introduction

The production of drawings is a behavior that has been performed for most of human history, as evident by the discovery of the paintings found in the Chauvet Cave in France and the El Castillo Cave in Spain, respectively dated to be approximately 32,000 and 40,000 years old (Clottes, 2003). Additionally, drawing is a behavior people perform throughout most of their lifespan, as most children as early as 2 years old (and in many cases even younger) draw (Kellogg, 1970), and many motivated adults continue to create drawings into old age. Furthermore, the behavior of drawing is a culturally valued behavior for both aesthetic and communicative purposes, as evident, for example, by the fact that many paintings have been sold for millions of dollars, by the existence of many kinds of draftsman professions, and by the strong presence of required drawing/painting instruction in the United States' elementary and high school education systems. Thus, drawing is a universal and important human behavior.

Although there are many different types of drawings that each have their own depictive and creative goals (e.g., schematic drawings, expressive drawings, memory- or imagination-based drawings, abstract drawings), this chapter exclusively focuses on the activity of *observational* drawing. Observational drawing is the behavior of creating a recognizable depiction of a specific model object or scene that is directly perceived by the individual while they create the drawing (unless otherwise stated, all future mentions of "drawing" in this chapter will specifically refer to observational drawing). There is a high degree of variability in the quality of drawings individuals are capable of producing, with most adults finding it difficult to produce high-quality drawings. Indeed, extensive training in drawing is often required before individuals are capable of producing high-quality drawings. Why is this the case? What prevents most adults without formal training from producing high-quality drawings? What is acquired during formal training and practice that leads to the development of strong drawing skill? What differences exist between skilled and unskilled individuals that contribute to this individual variability in drawing performance?

This behavior and set of questions have attracted scientific interest from cognitive and developmental psychologists alike. Cognitive psychologists have pursued this line of questioning from an information-processing perspective, as the activity of drawing is supported by multiple cognitive processes. Individuals utilize perceptual, attentional, decision-making, and memory processes to guide the production of such drawings. Many cognitive psychologists have attempted to understand how individual variability in such cognitive processes contributes to variability in the quality of drawings (Chamberlain & Wagemans, 2016). Developmental psychologists have addressed this set of questions by studying how the quality of and production strategies used to create drawings vary over the lifespan. Some have studied the stereotypical changes in drawing production that occur from early childhood into late childhood and adulthood (e.g., Freeman & Janikoun, 1972). Others have adopted an expertise approach by studying how deliberate training and practice affects, or is associated with, drawing performance and performance in nondrawing tasks that assess various types of cognitive processing ability (Kozbelt & Ostrofsky, 2018; Kozbelt & Seeley, 2007).

Studies of drawing adopt one of two general focuses: a process-oriented approach or a product-oriented approach. Research adopting the process-oriented approach mainly studies the sequence of actions one engages in when producing a drawing. Some process-oriented studies have investigated how variability in the sequence of mark-making affects or is related to the quality of the final drawing (e.g., Cohen, 2005; Tchalenko, 2009). Other studies adopting this approach have investigated how the experimental manipulation of cognitive processing affects the sequence of marks made when producing a drawing (e.g., Sommers, 1984; Vinter, 1999). However, since this approach represents a small minority of drawing research that has been recently conducted, it will not be the focus of this chapter.

Rather, this chapter mainly focuses on research that has adopted the product-oriented approach, which represents the majority of the research that has been conducted on this topic over the past 20 years and is largely concerned with assessing the quality and content of the final drawing. A major focus is assessing the visual accuracy of the final drawing, which is commonly measured in such studies. Cohen and Bennett (1997) defined a visually accurate drawing as "one that can be recognized as a particular object at a particular time and in a particular space, rendered with little addition of visual detail that cannot be seen in the object represented or with little deletion of visual detail [seen in the object represented]" (p. 609). Studies that have adopted this approach have aimed to understand how drawing accuracy/skill is (a) affected by manipulating various cognitive and perceptual factors (e.g., Cohen & Earls, 2010; Ostrofsky, Kozbelt, Tumminia & Cipriano, 2016), and/or (b) associated with performance in nondrawing tasks that measure some aspect of cognitive ability (e.g., Ostrofsky, Kozbelt & Seidel, 2012).

In this chapter, I will provide an overview of the various methods used in laboratory-based observational drawing research that adopts a product-oriented approach. First, I will describe the general questions and methodological designs used in such research. Next, I will summarize the various types of drawing tasks administered in studies focusing on this topic. Afterwards, I will summarize the various methods used to measure drawing skill/accuracy. Finally, I will end the chapter with a discussion of the challenges research in this area currently faces and suggestions as to the future directions research on this topic would benefit from by adopting.

General Questions and Methodological Strategies in Drawing Research

Multiple lines of general questions have been addressed in drawing research over the past 20 years. Below, I will identify each type of general question that has been studied and describe the methodological strategies that have been adopted to address them.

Relationship between drawing skill and performance in nondrawing tasks: Correlational studies

Since drawing is a complex behavior that is supported by perceptual and cognitive processing, researchers have been interested in determining which particular stages and components of perceptual/cognitive processing support skilled drawing. One way researchers have addressed this is by identifying a perceptual or cognitive process that is hypothesized to support drawing skill, and determining if individual variability in the ability to engage those processes is associated with variability in drawing performance. Examples of cognitive and perceptual processes that have been studied along these lines include:

- perceptual constancies pertaining to shape and size (Cohen & Jones, 2008; McManus, Loo, Chamberlain, Riley, & Brunswick, 2011; Ostrofsky et al., 2012; Ostrofsky, Cohen, & Kozbelt, 2014),
- various visual illusions (Chamberlain & Wagemans, 2015; Mitchell, Ropar, Ackroyd, & Rajendran, 2005; Ostrofsky, Kozbelt, & Cohen, 2015),
- local (in contrast to global) perceptual processing biases (Chamberlain, McManus, Riley, Rankin, & Brunswick, 2013; Drake, 2013; Drake & Winner, 2011),
- flexibility of visual attention (Chamberlain & Wagemans, 2015),
- efficiency of perceptually encoding the shape of an object (Perdreau & Cavanagh, 2014),
- integration of object-based visual information across eye movements (Perdreau & Cavanagh, 2013), and
- absolute and relative spatial positioning ability (Huang & Chen, 2017).

Correlational studies have been used to investigate such associations. In these studies, participants are required to create a drawing and complete at least one nondrawing task that has been designed to assess the perceptual or cognitive processing ability of interest. Statistically significant correlations between performance in the drawing and nondrawing tasks are the primary evidence used to support claims that drawing skill is associated with a particular perceptual or cognitive process.

Although much has been learned from such studies, this methodological strategy has important limitations one should be sensitive to when interpreting the correlational evidence. First, the observation of such statistically significant correlations is not solid evidence that drawing skill is directly related, in a causal manner, to the perceptual or cognitive process of interest. There is always the possibility that some unaccounted for variable exists that directly affects ability in drawing and perceptual/cognitive processing in parallel without drawing skill and perceptual/cognitive ability themselves being directly related. Second, even if ability in drawing and perceptual/cognitive processing were directly related, it is impossible to determine the direction of causality. Correlational evidence alone cannot be used to determine whether variability in perceptual/cognitive processing directly affects drawing performance, or vice versa. Thus, the safest conclusion one can draw from such studies is simply that a predictive association between drawing skill and perceptual/cognitive processing ability exists.

Differences in perceptual/cognitive processing ability between artists and nonartists: expertise studies

Expert-versus-novice differences in perception, cognition, and neural structure/function is a topic that has received a lot of attention from psychological researchers. For example, experts and novices in the domains of chess (Chase & Simon, 1973), action videogame playing (Green & Bavelier, 2003), musical notation reading (Wong & Gauthier, 2012), and navigation (Maguire et al., 2000) have been observed to differ with respect to a variety of perceptual, cognitive, and neural processes. Such studies are conducted for one of two general purposes. Some studies make such comparisons to provide clues pertaining to the perceptual, cognitive, or neural processes that support skilled performance. Other expert-versus-novice studies are performed to probe the plasticity of perceptual, cognitive, and neural systems that are associated with extensive experience in a specific domain.

Similar research has been conducted for the domain of drawing by comparing expert artists and novice nonartists (Kozbelt & Ostrofsky, 2018; Kozbelt & Seeley, 2007). Examples of processing abilities that have been compared between artists and nonartists include:

- the experience of perceptual constancies (Cohen & Jones, 2008; Ostrofsky et al., 2012; Perdreau & Cavanagh, 2011; McManus et al., 2011),
- susceptibility to various visual illusions (Chamberlain et al., 2019),
- perceptual grouping (Ostrofsky, Kozbelt, & Kurylo, 2013),
- perception of the size of angles (Carson & Allard, 2013),
- the flexibility of visual attention (Chamberlain & Wagemans, 2015),
- the ability to recognize objects found in degraded images (e.g., out-of-focus images, images of objects with segments deleted from the images) (Chamberlain et al., 2019; Kozbelt, 2001),
- visual memory ability (Perdreau & Cavanagh, 2014),
- the volume and activity of various brain regions (Chamberlain, McManus, Brunswick-et al., 2014; Schlegel et al., 2015), and
- face recognition (Devue & Barsics, 2016; Tree, Horry, Riley, & Wilmer, 2017; Zhou, Cheng, Zhang, & Wong, 2012).

Such studies adopt a quasi-experimental approach where expertise (artist versus nonartist) is the independent variable and performance on a nondrawing task assessing perceptual, cognitive, or neural processing is the dependent variable. Although many interesting expertise-based differences (and nondifferences) have been observed in such studies, this type of research is generally similar to the correlational studies described above with respect to its limitations in supporting causal claims. First, perceptual, cognitive, and neural differences between artists and nonartists do not necessarily indicate that such differences are directly related to differences in drawing ability between the groups. Second, even if such perceptual, cognitive, and neural differences were directly related in a causal manner to expertise-based differences in drawing skill, it is unclear whether the differences in perceptual, cognitive, and neural processes preceded, were followed by or were developed in parallel with drawing expertise.

Another limitation of this line of research concerns the nonstandardized approach pertaining to how participants are classified as an expert (or artist). Different studies and/or different research groups vary with respect to the criteria used to classify a participant as an expert artist. The following groups have been recruited to serve expert participants: (a) undergraduate art students (sometimes just majors and other times both majors and minors), (b) graduate art students, (c) community members recruited via fliers and online advertisements, and (d) professional artists (most of the time, the criteria used to classify someone as a professional artist are not provided). Some studies use only one of these four groups to serve as expert participants, and other studies have mixed individuals from two or more of these groups. Sometimes, the qualification of an individual would lead them to be categorized as an expert via one study's criteria and a novice via another study's criteria. For instance, Tchlalenko (2009) assigned professional artists to an expert group and assigned both undergraduate art students and nonartists to a novice group. In contrast, many other studies have assigned undergraduate art students to an expert group. Finally, at least one study assigned participants to expert and novice groups based on their performance on a drawing task administered during the course of the study, where a median split was used to categorize the half who drew better as experts and the half who drew more poorly as novices (Perdreau & Cavanagh, 2014).

The nonstandardized approach in categorizing participants as experts and novices creates potential problems when attempting to assess cross-study reliability of specific expertise-based differences. For instance, if there are discrepancies between multiple studies concerning an expertise-based difference on a specific cognitive/perceptual/neural variable, it would be difficult to conclude that the effect is unreliable if the two studies adopted different criteria to categorize participants as experts and novices. Two studies provided discrepant findings concerning expertise-based differences concerning face recognition abilities, as assessed by performance on the Cambridge Face Memory Test. Here, Devue and Barsics (2016) observed artists performing significantly better than nonartists in this task, whereas Tree and colleagues (2017) found no expertise-based difference. The criteria used to categorize participants as expert artists varied between the two studies. The expert group in the Devue and Barsics (2016) study were recruited from internet advertisements, had an average age of 26 years, and varied according to whether they were professionally trained or self-taught. In contrast, the members of the artist group in the Tree et al. (2017) study were professional portrait artists (teachers or those who work for commission) who had a mean age of 42 years. Such differences in the characteristics of the expert samples make it difficult to determine the source of the discrepancy in findings between these two studies. It could be that the expertise-related effect on face recognition abilities is unreliable, or it could be that these two artist groups are not homogeneous, and thus, not comparable for the purposes of assessing cross-study reliability.

Extending on this point, even when it appears that two studies adopted the same criteria to define expert participants, the characteristics of the experts may still vary between the two studies. For example, imagine two studies that define artists as undergraduate and/or graduate art students. If the students are recruited from different institutions, the level of expertise and skill in drawing may vary if one institution's admissions process involves different standards than the other. In such a case, the level of expertise between the two samples of artists may substantially vary. In light of these considerations, future expertise-based research in this area may benefit from developing a standardized method of categorizing individuals as experts. This would facilitate researchers' ability to compare the results of different studies with each other. However, even if such standardization is too challenging to develop, one should always be cautious in evaluating such expertise-based differences and pay close attention to the precise criteria used to define expertise.

Factors that influence drawing performance: experimental studies

Understanding methods that can lead to direct changes in drawing ability is of interest to both psychologists evaluating theories of drawing performance and art instructors who wish to identify and/or validate instructional methods used for improving students' drawing ability. In the correlational and expertise-based studies described above, researchers were primarily interested in how participants' current drawing ability is associated with performance in nondrawing tasks. In contrast, other studies have been concerned with investigating methods that could be used to directly improve and/or impair drawing performance. Examples of factors that have been assessed to determine if they affect drawing performance include:

- level of familiarity with the object category of which the model (or, the physical object being drawn) is a member (Glazek, 2012; Sheppard, Ropar, & Mitchell, 2005),
- the orientation of the model (e.g., upright versus upside-down) (Cohen & Earls, 2010; Day & Davidenko, 2018; Kozbelt, Seidel, ElBassiouny, Mark, & Owen, 2010; Ostrofsky, Kozbelt, Cohen, Conklin, & Thomson, 2016),
- the presence versus absence of three-dimensional depth cues in the model (Mitchell et al., 2005; Ostrofsky et al., 2015; Sheppard et al., 2005),
- the presence or absence of declarative knowledge pertaining to the canonical structure of an object (Ostrofsky, Kozbelt, Tumminia et al., 2016),
- the isolation of high vs. medium vs. low spatial frequencies in the model image (Freeman & Loschky, 2011),
- the manipulation of the frequency in which individuals shift their gaze between the model and drawing during production (Cohen, 2005), and
- how individuals interpret the identity of the model object being drawn (Ostrofsky, Nehl, & Mannion, 2017; Vinter, 1999).

Experiments have been conducted to investigate these effects. In such experiments, a measure of drawing performance is used as the dependent variable and the independent variable is the researchers' manipulation of the factor that is hypothesized to have an effect on drawing performance. Significant differences in the mean values of the drawing performance measure between the different levels of the independent variable are the evidence used to claim that the independent variable has an effect on drawing performance.

One major limitation of such studies pertains to the generalizability of the effects observed in the experiments. In most cases, the model being drawn in a given experiment represents just one object category. For instance, in the experiments referenced above that investigated the effects of model orientation on drawing performance, the model being drawn was a face, and the researchers manipulated whether the model-face was displayed upright or upside-down (Cohen & Earls, 2010; Day & Davidenko, 2018; Ostrofsky, Kozbelt, Cohen et al., 2016). Across these studies, the results indicated that drawing performance was either impaired or not affected when drawing upside-down faces (relative to upright faces) depending on the particular measure of drawing performance analyzed. In no case was it observed that drawing performance was improved via face inversion. However, this is not strong evidence to conclude that inverting a model has no benefits to drawing performance in general, as these experiments only looked at performance in drawing faces. Further experimentation is required to assess whether inverting models of other object categories has similar or dissimilar effects as those observed in face drawings. More generally, the larger set of studies referenced above have the same limitation, in that drawing performance is typically measured with respect to drawing models that represent only one or just a few object categories. Furthermore, it is rare that researchers attempt to replicate such experimental effects using models of different object categories beyond those used in the experiments that originally demonstrated the effect. Thus, until such replication studies are conducted, one must avoid overgeneralizing experimental effects as existing for drawing performance in general as opposed to drawing performance for a particular type of model object.

Types of Drawing Tasks Administered in Laboratory Research

For most drawing studies, drawing skill is based on an assessment of observational drawings produced by participants in the laboratory. This section focuses on describing the drawing tasks that have been administered and will highlight some of the common methodological features researchers have adopted.

Free-hand drawing tasks

Free-hand drawing tasks refer to participants being exposed to a model and being asked to draw a reproduction of it as it appears. In studies of observational drawing, such tasks often instruct participants to strive for accuracy above all else when reproducing a model via drawing. Commonly, instructions set the participants' goal to producing as accurate a copy of the model as the participant is capable of producing without necessarily striving for producing a highly creative or even aesthetically pleasing drawing.

Different types of models have been used in such drawing tasks. For instance, photographs are often used, especially those that depict a single object such as an octopus (Ostrofsky et al., 2014), a hand holding a pencil, a Lego-block configuration (Chamberlain et al., 2013), faces (Ostrofsky et al., 2014), and a generator (Cohen & Bennett, 1997).

Physically present three-dimensional objects are also used at times, most often in the form of a “still-life” collection of objects. In some studies, the researchers standardize the distance and angle from the objects the participants are seated at (Chamberlain et al., 2013) and other studies allow participants to freely inspect the collection of objects by moving their head and sitting at different positions (Carson, Millard, Quehl, & Danckert, 2014).

Studies employing photographs and physically present objects as models represent the most complex free-hand drawing tasks. In order to skillfully draw such models, participants must accurately depict a wide variety of visual information that includes shading, angles, relative line length, detailed appearance of local features, linear perspective, relative and absolute spatial positioning of different features, relative size of different features, etc. Thus, in order to create a high-quality drawing of such models, participants must exhibit skill in many aspects of drawing. In other studies, researchers have used simpler models that most commonly come in the form of simple line drawings. Examples of such simple models have included faces (Day & Davidenko, 2018; Tchalenko, 2009), full human bodies (Tchalenko, 2009), an eye, a wine glass, Chinese ideograms (Glazek, 2012), parallelograms (Mitchell et al., 2005), cylinders (Matthews & Adams, 2008), and abstract shapes not representing meaningful objects (Sheppard et al., 2005). One reason such simple models are used is to assess a very specific aspect of drawing skill. For instance, Chamberlain, McManus, Riley et al. (2014) asked participants to draw multiple irregular hexagon models (the “Cain’s House Task”) in order to narrowly assess skill in drawing angles and relative line lengths. Another reason such simple models are used is to eliminate some aspects of drawing that are needed to be utilized in order to produce a high-quality drawing. For instance, copying line drawings of faces, wine glasses, eyes, and tables eliminates the need for participants to accurately shade in order to draw a high-quality reproduction of the model.

In most cases, the use of a free-hand drawing task is intended to allow researchers to have an ecologically valid assessment of general drawing skill where drawings are produced under natural conditions. Participants typically produce drawings on paper using pencil, are allowed to erase and modify their drawing during the course of its production, are free to utilize whatever drawing techniques they know/desire to use (with the exception of tracing), and are not constrained with respect to the sequential production of the drawing. However, one constraint most free-hand drawing tasks impose on participants that might reduce the ecological validity of the drawing assessment involves the time limits participants are often given to complete their drawings. Depending on the complexity of the model being drawn, this time limit can range from seconds (e.g., 30-s time limit used in Schlegel et al., 2015) to minutes (e.g., 15-min time limit used by Ostrofsky et al., 2012) to an hour (e.g., Huang & Chen, 2017). Such time limits are typically used to standardize the drawing tasks across participants and, for convenience, to ensure that participants complete all of the study’s required tasks within the allotted time for the data collection session. However, it is open to question as to how well timed tasks result in drawings that reflect a participant’s maximum level of drawing skill. It is always possible that a participant would be able to produce a higher quality drawing if they were allotted more time to produce it or did not know from the outset that they had to complete the drawing within a specific time frame.

Tracing tasks

It is widely accepted that observational drawing is a complex behavior that is guided by multiple perceptual and cognitive processes. Cohen and Bennett (1997) identified four main processes that guide the production of a drawing: (a) perceptual encoding of the model; (b) representational decision-making as to what features from the model to emphasize, de-emphasize, or neglect in the drawing; (c) eye–hand motor coordination; and (d) evaluation of the quality of the emerging drawing and making necessary corrections when deviations between the model and drawing are perceived. Tracing tasks have generally been used by researchers in order to study, in an isolated fashion, the decision-making, and/or eye–hand motor coordination processes involved in drawing. Cohen and Bennett (1997, Experiment 1) asked participants to trace photographs. This task eliminated the need for participants to perceptually encode many aspects of the photograph that would have needed to be accurately encoded in a free-hand drawing task (e.g., one does not need to accurately encode size-proportions, relative spatial positioning of different features, angles, etc.). This task also substantially reduced the difficulty of the evaluation process, as the only evaluation participants needed to make was whether they missed any lines or if their drawn line deviated from the path of the printed line found in the photograph. Thus, the researchers isolated decision-making (i.e., deciding which lines to emphasize/de-emphasize and how thick to draw a specific line) and eye–hand coordination as the pertinent skills needed to produce a high-quality depiction. In their second experiment, the researchers asked participants to trace a tracing produced by another individual, which further isolated eye–hand motor coordination as the only pertinent skill needed to produce a high-quality tracing. Here, decision-making processes are eliminated, or at least are substantially reduced.

A variant tracing task that has been argued to more sensitively assess representational decision-making processes is the “limited-line tracing task” (Chamberlain et al., 2019; Kozbelt et al., 2010; Ostrofsky et al., 2012). A potential limitation of the traditional tracing method is that individuals are free to use an unlimited number of lines to trace the model. For studies that aim to compare representational decision-making differences between groups (e.g., artists versus nonartists), this freedom may mask any differences between individuals who make stronger versus weaker depictive decisions. The limited line tracing task controls the lines participants are allowed to use to trace the model, so that the number, thickness, and length of lines used for the tracing are standardized and equated across participants. Critically, the number of lines participants are allowed to use is fewer than what is required to trace the entire the image. This method forces participants to be more economical in their decision-making, forcing them to prioritize and decide which segments of the model are more or less important to depict. This method has been useful in establishing differences in decision-making quality between those who are more versus less skilled in drawing, as the quality of such limited line tracings have been found to be (a) rated significantly higher for those produced by expert artists than those produced by novices, and (b) positively correlated with the quality of free-hand drawings.

Measures of Drawing Performance

Once drawings or tracings are produced, performance is measured using a variety of quantitative assessment methods. There are two general categories of quantitative measures of drawing performance that have been used in such research: (a) subjective rating methods that aim to assess perceived quality, and (b) objective measurement methods that aim to measure specific deviations in appearance between the drawing and the model. In most cases, the subjective and objective measurement methods aim to quantify how accurately the drawings reproduce the visual appearance of the model. Each will be described in detail below.

Subjective rating methods

Subjective accuracy ratings are used in cases where researchers are interested in understanding "perceived accuracy," or how accurate a drawing reproduces the model as visually perceived by independent observers. Studies that measure drawing performance via subjective accuracy ratings typically recruit a sample of independent judges who are instructed to view the drawings and the model side-by-side and provide judgments (most often in the form of Likert-type ratings) pertaining to how accurate the drawing reproduced the visual appearance of the model. Often, judges are instructed to base their ratings specifically on how well the drawing reproduced the visual appearance of the model without considering how creative or aesthetically pleasing the drawings are. Despite the subjective nature of this type of assessment, most studies employing this method find high levels of inter-rater reliability, as evident by observed Cronbach alpha levels of .70 or above, often exceeding levels of .90 (e.g., Chamberlain McManus, Riley et al., 2014; Cohen & Earls, 2010; Cohen & Jones, 2008; Ostrofsky et al., 2012). High reliability levels have been found in studies that employed expert judges, novice judges and both, and within studies that have sampled a relatively small (e.g., $N = 3$, Ostrofsky et al., 2012) and large (e.g., $N = 51$, Cohen & Jones, 2008) number of judges. This suggests that subjective accuracy ratings are a reliable method for assessing perceived accuracy, especially for the purposes of comparing the perceived accuracy between different drawings.

In many studies, each judge provides a single rating per drawing that represents overall perceived accuracy (e.g., Chamberlain et al., 2019; Cohen, 2005; Freeman & Loschky, 2011; Glazek, 2012; Ostrofsky et al., 2012). In some of these studies, judges are given little criteria to base their judgments on other than being asked to rate how accurately the drawing reproduced the appearance of the model (e.g., Ostrofsky et al., 2012). In other studies, researchers provide judges with a detailed rubric to review that specifies some aspects of the drawing they should attend to when determining what rating value to assign to a drawing (e.g., Chamberlain et al., 2019; Huang & Chen, 2017). For example, Huang and Chen (2017) provided a rubric that included the accuracy in reproducing: (a) the relative spatial positioning of multiple objects, (b) form and shape, (c) shadows, and (d) overall realism. Presumably, one goal of providing such rubrics is to increase inter-rater reliability, as one may expect higher reliability if all of the judges are basing their ratings on the same criteria as opposed to each judge utilizing their own idiosyncratic criteria of drawing accuracy in the absence of such a rubric. However, the effect of providing a rubric on such reliability has not been explored, as no study has compared rubric- vs. non-rubric-based ratings for a single set of drawings. Furthermore, high reliability levels observed in studies employing non-rubric-based rating methods suggest that rubrics are not necessary to establish strong inter-rater reliability.

One criticism of measuring drawing performance using single ratings to represent overall perceived accuracy is that this type of measure does not capture the complex, multifaceted nature of drawing accuracy. A single drawing can be relatively accurate with respect to reproducing some aspects of a model but relatively inaccurate with respect to reproducing others. For instance, a drawing may be highly accurate with respect to reproducing the relative spatial positioning of features, but may be highly inaccurate with respect to reproducing shading gradients that are needed to convey depth or in drawing the detailed appearance of individual features. This complexity of drawing accuracy is masked when drawing performance is reduced to a single-value accuracy rating. This potentially creates problems of interpretation for the types of correlational and experimental studies described earlier. When significant correlations or experimental effects are observed, it is unclear which aspects of drawing accuracy are related to nondrawing task performance or which are affected by experimental manipulations when drawing accuracy is assessed via single-value accuracy ratings. The correlations or experimental effects could pertain to all or only some aspects of perceived drawing accuracy, and the use of single-value accuracy ratings makes it impossible to determine which is the case.

In order to assess perceived drawing accuracy in a more specific way, some studies have instructed judges to provide multiple ratings that each focus on a different aspect of drawing accuracy. For instance, two face-drawing studies reported by Cohen and colleagues asked judges to provide three ratings per drawing: (a) overall accuracy, (b) accuracy in drawing individual facial features, and (c) accuracy in drawing the relative spatial positioning of the features (Cohen & Earls, 2010; Cohen & Jones, 2008). As another example, Hayes and Milne (2011) instructed judges to rate the accuracy of face drawings according to 10 aspects, including face shape, eye spacing, eye size, nose length, nose width, distance between the nose and mouth, mouth width, lip fullness, distance between the mouth and chin, and chin size. However, even though such ratings can provide us with a more nuanced understanding of how accurate a drawing is perceived to be, the use of such ratings may still mask some of the complexities of drawing accuracy. For instance, although the studies by Cohen and colleagues cited above instructed judges to provide specific ratings concerning the accuracy of drawing the relative spatial positioning of features, there are many spatial relationships within a model that could be drawn with different levels of accuracy. Cohen and Earls (2010) found that drawings of upside-down faces are rated as less spatially accurate compared with upright face drawings, suggesting that face inversion impairs the ability to draw spatial relationships between features. But, when assessing specific spatial relationships in a later study, Ostrofsky, Kozbelt, Cohen et al. (2016) found that not all spatial relationships within a face are impaired by face inversion; inversion impaired drawing accuracy for one spatial relationship (the vertical distance between the eyes and mouth), but not others (e.g., the horizontal distance between the two eyes; the vertical distance between the nose and mouth). Thus, although subjective accuracy ratings are useful for broad-level assessments of perceived drawing accuracy (i.e., allowing one to empirically establish whether one drawing is, overall, more or less accurate than another), they may mask some nuanced, and potentially important, aspects of how well a drawing reproduced specific aspects of the visual appearance of a model.

Objective measurement methods

Although subjective accuracy ratings are useful for assessing how accurate a drawing is perceived to be by observers, they do not allow one to specifically assess how a drawing deviated in appearance from the model. Subjective accuracy ratings often fall short in specifying the aspects of a drawing that are more or less accurate, and they do not allow one to precisely quantify the magnitude of drawing error. Thus, some studies have assessed drawing performance using objective measurements of drawing error that precisely quantify specific deviations between a drawing and the model. Generally speaking, there are three categories of objective measurement methods that have been used in drawing research: (a) anthropometric measures, (b) landmark-based morphometric measures, and (c) feature counting measures.

Anthropometric measures refer to those that quantify spatial aspects of a drawing (e.g., size of a feature, distance between multiple features) using proportional variables. Although anthropometry was developed specifically for measures of the human body in non-drawing contexts, the basic method has been used to measure the accuracy of drawings based on a variety of model object categories, including faces (Costa & Corazza, 2006; Harrison, Jones, & Davies, 2017; Hayes & Milne, 2011; Ostrofsky et al., 2014), the human body (Tchalenko, 2009), cylinders (Matthews & Adams, 2008), parallelograms (Mitchell et al., 2005), and houses (Harrison et al., 2017). As one simple example, the width of an eye has been quantified as the eye width divided by the face width (Ostrofsky et al., 2014). As another example, the height of a house’s second story window has been quantified as the height of the window divided by the overall height of the house (Harrison et al., 2017). Such proportioned measures control for differences in the absolute size between a model and drawing (and between different drawings) in order to facilitate comparisons, which is useful as it is generally accepted that the quality in drawing spatial aspects of an image is normally assessed based on accuracy in reproducing relative proportions rather than absolute sizes. Once the drawings and the model have been measured using this method, drawing errors can be computed using a variety of quantitative variables (e.g., computing the difference between the drawing and model measures, computing ratios of the drawing and model measures, computing the difference between the drawing and model measures and dividing this difference by the model measure to express error as a proportion or percentage of deviation from the model measure).

Anthropometric methods are useful in that they allow researchers to define specific spatial aspects of a drawing that are being assessed for accuracy, and provide precise measures of error in terms of direction and magnitude (e.g., by how much an eye was drawn too wide or narrow). However, they are not suited to assessing a drawing's overall accuracy for two major reasons. First, they are not capable of assessing important nonspatial aspects of a drawing, such as the accurate depiction of shadows and the detailed visual appearance of local features. Second, assessments of drawing accuracy are restricted to the specific spatial relationships selected by the researcher. In consideration of statistical power, researchers should be mindful of the number of spatial relationships assessed, as increases in the number of spatial relationships measured decreases statistical power if researchers appropriately control for inflated Type I error rates due to multiple comparisons. In cases where the model being drawn is visually complex and contains a large number of spatial relationships, this restriction may lead to researchers narrowly focusing on only some aspects of spatial drawing accuracy, leading them to potentially neglect assessing accuracy of important spatial aspects of the drawing that impact perceived accuracy or that are of potential theoretical interest. Thus, this method is generally best used in cases where a researcher wants to analyze the accuracy of a very specific subset of spatial relationships found in the depicted object rather than analyzing how accurately the drawing reproduced all possible spatial relationships (e.g., Mitchell et al., 2005; Ostrofsky, Kozbelt, Tumminia et al., 2016).

If a researcher is more interested in objectively measuring overall spatial accuracy rather than measuring drawing error for specifically defined spatial relationships, landmark-based morphometric measures are a useful alternative that have been used in multiple drawing studies (Chamberlain McManus, Riley et al., 2014; Day & Davidenko, 2018; Hayes & Milne, 2011; Perdreau & Cavanagh, 2014; for a general, nondrawing specific review of this method, see Webster & Sheets, 2010). Assessing drawing accuracy using this method entails first defining a number of landmarks in the model image, where the number and location of landmarks used are determined by the researcher. These landmarks are set to be positioned at easily identifiable locations in the image that can later be located in the drawings of that model. For instance, Chamberlain McManus, Riley et al. (2014) used photographic models of a Lego-block configuration and a hand holding a pencil. For the Lego-block model, the landmarks were set at most of the corners visible in the configuration of blocks. For the model depicting a hand holding a pencil, landmarks were placed at every knuckle on the hand and the two endpoints of the pencil. Once drawings of the model are produced, they are digitized and researchers place points at all of the predefined landmarks. The goal of the analysis is to determine the degree to which the drawings deviated from the model with respect to the relative positioning of the landmarks. This is often accomplished via Procrustes analysis. Here, the model and drawing landmarks are mapped in two-dimensional space and Euclidean transformations of position (translation), size (scaling), orientation (rotation), and reflection are applied to the drawings in order to minimize, as much as possible, the deviation in the relative spatial positioning (or coordinates) of all the landmarks between the drawing and model. After these transformations are applied, the deviation in the position of each landmark between the drawing and the model is computed as the Euclidean distance between them. Once the Euclidean distance has been computed for all of the landmarks, the Procrustes Distance statistic is calculated as (stated in an oversimplified manner) the sum of the Euclidean distances for all the landmarks. Thus, the greater the value of the Procrustes Distance statistic, the greater the drawing deviated from the model with respect to the relative spatial positioning of all the landmarks (or, in other words, the greater the degree of spatial drawing error).

As alluded to above, this method is only useful for assessing overall accuracy in drawing the relative spatial positioning of features. Unlike anthropometric measures, landmark-based morphometric statistics are not capable of determining the ways that a drawing erred with respect to reproducing specific spatial relationships between features found in a model. Further, like anthropometric measures, this method is only useful for assessing spatial drawing accuracy and cannot be used to assess other important aspects of drawing accuracy, like shading and the detailed appearance of local features.

At this point, we have discussed two major methods that are useful for objectively assessing spatial drawing accuracy. Moving beyond this, feature counting measures are another objective method for assessing drawing performance. Here, the researcher identifies a number of important visual features found in the model, and counts how many of them were included versus excluded from the drawing. For example, line junctions, or vertices, are an important visual feature that support object recognition in general (Biederman, 1987) and are useful in drawing in order to convey depth. Drawing and tracing performance has been partially assessed in some studies by counting the number of line junctions found in a model that were included in a drawing/tracing (Biederman & Kim, 2008; Ostrofsky et al., 2012). As another example, Drake (2013) assessed drawing performance in children by counting the presence versus absence of a number of important drawing features, such as the use of occlusion, foreshadowing, and the presence of single lines to represent the edges of objects as opposed to representing the entire object itself.

Such methods are useful in studying attentional and representational decision-making processes in drawing, as feature counting methods provide insights about what features individuals attend to versus neglect and/or what features individuals decide to include versus exclude from a drawing. For instance, in the two studies cited above that assessed drawings/tracings for the presence versus absence of line junctions, it was found that expert artists are more likely to depict line junctions than novice nonartists. This highlights that experts and novices differ in what visual features of the model they attend to and decide to include in their drawings, and thus, may be part of the reason why they differ in their ability to produce high-quality drawings. Skill in drawing may be associated with a greater sensitivity pertaining to the visual cues, like line junctions, that support strong object recognition (which is the essential goal of creating an observational drawing).

Directions for Future Research

In concluding this chapter, this section will highlight some suggested directions for future research, focusing on issues pertaining to the measurement of drawing performance and methodological approaches useful in assessing the causal relationships between drawing skill and perceptual/cognitive processing ability.

Measures of drawing performance

As explained in the prior section, subjective and objective measures of drawing accuracy differ in that subjective measures assess perceived accuracy whereas objective measures assess how drawings actually deviated in appearance from the model. Usually, studies utilize either subjective or objective measures. This results in a study either assessing how accurate a drawing is perceived to be by others without understanding how a drawing actually deviated in appearance from the model, or the study assessing how a drawing deviated in appearance from a model without understanding whether such deviations impacted perceived accuracy. Future drawing research would benefit by using subjective and objective measures in conjunction, as doing so would provide a more complete understanding of drawing performance within a study and may provide clues as to what types of objective drawing errors are more or less associated with how accurate a drawing is perceived to be by others. Just because a drawing objectively deviated from a model in some aspect does not necessarily mean that the specific drawing error is related to how accurate the drawing was perceived to be by others. For instance, the few studies that have assessed the relationship between subjective and objective accuracy measures for a single set of drawings have demonstrated that not all types of objective drawing errors are predictive of perceived accuracy. For example, face-drawing studies have demonstrated that subjective accuracy ratings are predicted by objectively measured errors in depicting some, but not all, spatial relationships between features (Hayes & Milne, 2011; Ostrofsky et al., 2014). Further, Chamberlain McManus, Riley et al. (2014) found that subjective accuracy ratings were predicted by objective, morphometric measures of error for drawings of some types of models (e.g., abstract hexagons, photographs of Lego blocks and hands) but not for others (e.g., the painting titled *Suprematism with Eight Red Rectangles* created by Malevich).

Despite the common use of subjective accuracy ratings to measure drawing performance, the factors that influence perceived drawing accuracy are currently not well understood. However, the common observation of high levels of inter-rater reliability of subjective accuracy ratings suggests that there are predictable properties of drawings that individuals attend to and evaluate in a consistent fashion when judging how accurate they perceive the drawing to be. Further, there have been observations of differences in the average subjective accuracy ratings provided by expert artists and nonartists (e.g., Kozbelt et al., 2010), suggesting that different populations vary in the criteria used when judging how accurate a depiction appears to be. Future research can be aimed at systematically analyzing how different aspects of drawing are weighted during the evaluation of perceived accuracy. What general properties of a drawing are more or less attended to when perceived accuracy judgments are being formed? What aspects of a drawing are weighted heavily versus moderately versus weakly in the judgment of perceived accuracy? How does the weighting of different aspects of drawing in this evaluation process vary across different populations? Such questions can be addressed via two approaches. One approach that has been used involves assessing correlations between subjective accuracy ratings and objective error measurements (Hayes & Milne, 2011; Ostrofsky et al., 2014). However, an observed correlation between subjective accuracy ratings and a particular objective error measure does not guarantee that the objective error directly influenced subjective accuracy ratings (due to the general limitations of correlational evidence). Alternatively, experimental methods could serve as another approach, where researchers systematically manipulate the appearance of drawings and determine how they affect individuals' perceptions of drawing accuracy. Here, participants can judge the accuracy of well-controlled stimuli that systematically vary the appearance of one aspect of drawing while holding all other aspects constant. In this way, researchers can determine the isolated contribution that the manipulated aspect has on perceived accuracy. In one of the only, if not the only, studies adopting this approach, Biederman and Kim (2008) presented participants with two drawings of a horse that were identical in every regard except the presence or absence of a particular line junction. They found that participants judged that the drawing with the line junction was a better depiction than the drawing without it, indicating that the presence versus absence of line junctions is an aspect of drawing that influences perceived drawing accuracy. This approach can conceivably be adapted to study how perceived accuracy is influenced by many different aspects of drawing accuracy.

Methods for understanding causal relationships between drawing skill and cognition/perception

As mentioned in an earlier section, correlational and expertise-based drawing studies generally aim to determine the differences in nondrawing-related cognitive/perceptual processes between skilled and unskilled drawers. Although the overarching aim of such studies is often to understand how cognitive/perceptual processing differences either support skilled drawing or are developed as a consequence of skill acquisition, the limitations of correlational and quasi-experimental evidence prevent one from understanding the specific causal relationships that produced the observed correlations between drawing and perceptual/cognitive processing ability or expertise-based differences in perceptual/cognitive processing ability. In order to advance our understanding of the cognitive/perceptual processes that directly impact drawing performance or to understand the cognitive/perceptual changes that occur as a result of the acquisition of drawing expertise, researchers can adopt an experimental, longitudinal methodological approach. If researchers are interested in understanding how a particular cognitive/perceptual process impacts drawing performance, researchers can conduct experiments where participants are trained over a long period of time in order to develop stronger ability for a particular cognitive/perceptual process, and then determine if that training results in improvements (or impairments) in drawing performance (relative to a baseline measure). Currently, I am unaware of any study that has adopted this approach, and thus, this void presents a great opportunity for novel and theoretically significant research on this topic. Alternatively, if researchers are interested in understanding how cognitive/perceptual processing ability changes as a result of the acquisition of drawing skill, researchers can conduct an experiment where participants are trained over a long period of time to improve their drawing skill, and then determine how cognitive/perceptual ability changes as a consequence of drawing training (relative to a baseline measure). There have only been a few studies that have adopted this approach (Kozbelt et al., 2016; Schlegel et al., 2015; Tree et al., 2017), and thus, there is a clear need for more research along these lines as changes in only a small number of cognitive/perceptual abilities have been assessed.

Conclusion

Observational drawing is a topic of scientific interest to psychologists. As reviewed in the chapter, there are various methods that have been used to understand individual variability in drawing ability and in order to measure drawing performance in laboratory-based studies. Although much can be learned from each method, this chapter highlighted significant limitations each method has pertaining to interpretations one can validly draw from the results of a study or pertaining to the aspects of drawing performance a particular measurement method is capable of assessing. However, such limitations do not indicate that drawing is a topic that cannot be subjected to scientific inquiry. Rather, such limitations simply highlight the need of researchers to critically evaluate the methods used by a particular study in order to draw valid scientific conclusions and to avoid overgeneralizing results in such a way that they are not supported by the methodological features used by a particular study. This is by no means unique to the study of observational drawing, as the study of all topics in scientific research is based on methods that have their own unique limitations. Since the scientific study of adult observational drawing performance is a relatively young field, one can expect refinements in the methods used to study this topic to be developed in the future. The preceding section of this chapter provided suggestions on how some of these limitations can be improved upon in future research, and thus, one can be optimistic that research on observational drawing will continue to develop and to provide more insights that can explain the tremendous range of individual variability in drawing performance that is found in the population.

References

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

Biederman, I., & Kim, J. G. (2008). 17,000 years of depicting the junction of two smooth shapes. *Perception*, *37*, 161–164.

Carson, L. C., & Allard, A. (2013). Angle-drawing accuracy as an objective performance-based measure of drawing expertise. *Psychology of Aesthetics, Creativity, and the Arts*, *7*, 119–129.

Carson, L. C., Millard, M., Quehl, N., & Danckert, J. (2014). Polygon-based drawing accuracy analysis and positive/negative space. *Art & Perception*, *2*, 213–236.

Chamberlain, R., Drake, J. E., Kozbelt, A., Hickman, R., Siev, J., & Wagemans, J. (2019). Artists as experts in visual cognition: An update. *Psychology of Aesthetics, Creativity, and the Arts, 13*, 58–73

Chamberlain, R., McManus, I. C., Brunswick, N., Rankin, Q., Riley, H., & Kanai, R. (2014). Drawing on the right side of the brain: A voxel-based morphometry analysis of observational drawing. *NeuroImage*, *96*, 167–173.

Chamberlain, R., McManus, I. C., Riley, H., Rankin, Q., & Brunswick, N. (2013). Local processing enhancements associated with superior observational drawing are due to enhanced perceptual functioning, not weak central coherence. *Quarterly Journal of Experimental Psychology*, *66*, 1448–1466.

Chamberlain, R., McManus, I. C., Riley, H., Rankin, Q., & Brunswick, N. (2014). Cain's House Task revisited and revived: extending theory and methodology for quantifying drawing accuracy. *Psychology of Aesthetics, Creativity, and the Arts*, *8*, 152–167.

Chamberlain, R., & Wagemans, J. (2015). Visual arts training is linked to flexible attention to local and global levels of visual stimuli. *Acta Psychologica*, *161*, 185–197.

Chamberlain, R., & Wagemans, J. (2016). The genesis of errors in drawing. *Neuroscience & Biobehavioral Reviews*, *65*, 195–207.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Clottes, J. (2003). *Chauvet cave: The art of earliest times*. Salt Lake City, UT: University of Utah Press.

Cohen, D. J. (2005). Look little, look often: the influence of gaze frequency on drawing accuracy. *Perception and Psychophysics*, *67*, 997–1009.

Cohen, D. J., & Bennett, S. (1997). Why can't most people draw what they see? *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 609–621.

Cohen, D. J., & Earls, H. (2010). Inverting an image does not improve drawing accuracy. *Psychology of Aesthetics, Creativity, and the Arts*, *4*, *16*8–172.

Cohen, D. J., & Jones, H. E. (2008). How shape constancy relates to drawing accuracy. *Psychology of Aesthetics, Creativity, and the Arts*, *2*, 8–19.

Costa, M., & Corazza, L. (2006). Aesthetic phenomena as supernormal stimuli: The case of eye, lip, and lower-face size and roundness in artistic portraits. *Perception*, *35*, 229–246.

Day, J. A., & Davidenko, N. (2018). Physical and perceptual accuracy of upright and inverted face drawings. *Visual Cognition*, *26*, 89–99.

Devue, C., & Barsics, C. (2016). Outlining face processing skills of portrait artists: Perceptual experience with faces predicts performance. *Vision Research*, *127*, 92–103.

Drake, J. E. (2013). Is superior local processing in the visuospatial domain a function of drawing talent rather than ASD? *Psychology of Aesthetics, Creativity, and the Arts*, *7*, 203–209.

Drake, J. E., & Winner, E. (2011). Realistic drawing talent in typical adults is associated with the same kind of local processing bias found in individuals with ASD. *Journal of Autism and Developmental Disorders*, *41*, 1192–1201.

Freeman, N. H., & Janikoun, R. (1972). Intellectual realism in children's drawings of a familiar object with distinctive features. *Child Development*, *43*, 1116–1121.

Freeman, T. L., & Loschky, L.C. (2011). Low and high spatial frequencies are most useful for drawing. *Psychology of Aesthetics, Creativity, and the Arts*, 5, 269–278.

Glazek, K. (2012). Visual and motor processing in visual artists: Implications for cognitive and neural mechanisms. *Psychology of Aesthetics, Creativity and the Arts*, *6*, 155–167.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*, 534–537.

Harrison, N. R., Jones, J., & Davies, S. J. (2017). Systematic distortions in vertical placement of features in drawings of faces and houses. *i-Perception*, *8*, 204166951769105.

Hayes, S., & Milne, N. (2011). What's wrong with this picture? An experiment in quantifying accuracy in 2D portrait drawing. *Visual Communication*, *10*, 149–174.

Huang, S. T., & Chen, I. P. (2017). The relationship between 2D positioning ability and realistic drawing skill. *Art and Design Review*, *5*, 1–12.

Kellogg, R. (1970). *Analyzing children's art*. Mountain, CA: Mayfield Publishing Company.

Kozbelt, A. (2001). Artists as experts in visual cognition. *Visual Cognition*, *8*, 705–723.

Kozbelt, A., Chamberlain, R., & Drake, J. E. (2016). Learning to see by learning to draw: a longitudinal study of perceptual changes among artists-in-training. Presented at the Congress of the International Association of Empirical Aesthetics, Vienna, Austria.

Kozbelt, A., & Ostrofsky, J. (2018). Expertise in drawing. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance*, 2nd edition (pp. 576–596). Cambridge, England: Cambridge University Press.

Kozbelt, A., & Seeley, W. (2007). Integrating art historical, psychological, and neuroscientific explanations of artists' advantages in drawing and perception. *Psychology of Aesthetics, Creativity and the Arts*, *1*, 80–90.

Kozbelt, A., Seidel, A., ElBassiouny, A., Mark, Y., & Owen, D. R. (2010). Visual selection contributes to artists' advantages in realistic drawing. *Psychology of Aesthetics, Creativity, and the Arts*, *4*, 93–102.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 4398–4403.

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

McManus, I. C., Loo, P., Chamberlain, R., Riley, H., & Brunswick, N. (2011). Does shape constancy relate to drawing ability? Two failures to replicate. *Empirical Studies of the Arts*, *29*, 191–208.

Mitchell, P., Ropar, D., Ackroyd, K., & Rajendran, G. (2005). How perception impacts on drawings. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 996–1003.

Ostrofsky, J., Cohen, D. J., & Kozbelt, A. (2014). Objective versus subjective measures of face-drawing accuracy and their relations with perceptual constancies. *Psychology of Aesthetics, Creativity, and the Arts*, *8*, 486–497.

Ostrofsky, J., Kozbelt, A., & Cohen, D. J. (2015). Observational drawing biases are predicted by biases in perception: Empirical support of the misperception hypothesis of drawing accuracy with respect to two angle illusions. *The Quarterly Journal of Experimental Psychology*, *68*, 1007–1025.

Ostrofsky, J., Kozbelt, A., Cohen, D. J., Conklin, L., & Thomson, K. (2016). Face inversion impairs the ability to draw long-range, but not short-range, spatial relationships between features. *Empirical Studies of the Arts*, *34*, 221–233.

Ostrofsky, J., Kozbelt, A., & Kurylo, D. (2013). Perceptual grouping in artists and non-artists: A psychophysical comparison. *Empirical Studies of the Arts*, *31*, 131–143.

Ostrofsky, J., Kozbelt, A., & Seidel, A. (2012). Perceptual constancies and visual selection as predictors of realistic drawing skill. *Psychology of Aesthetics, Creativity, and the Arts*, *6*, 124–136.

Ostrofsky, J., Kozbelt, A., Tumminia, M., & Cipriano, M. (2016). Why do non-artists draw the eyes too far up the head? How vertical eye-drawing errors relate to schematic knowledge, pseudoneglect and context-based perceptual biases. *Psychology of Aesthetics, Creativity and the Arts*, *10*, 332–343.

Ostrofsky, J., Nehl, H., & Mannion, K. (2017). The effect of object interpretation on the appearance of drawings of ambiguous figures. *Psychology of Aesthetics, Creativity and the Arts*, *11*, 99–108.

Perdreau, F., & Cavanagh, P. (2011). Do artists see their retinas? *Frontiers in Human Neuroscience*, *5*, 1–10.

Perdreau, F., & Cavanagh, P. (2013). The artist's advantage: Better integration of object information across eye movements. *i-Perception*, *4*, 380–395.

Perdreau, F., & Cavanagh, P. (2014). Drawing skill is related to the efficiency of encoding object structure. *i-Perception*, *5*, 101–119.

Schlegel, A., Alexander, P., Foggelson, S. V., Li, X., Lu, Z., Kohler, P. J., Riley, E., Tse, P. U., & Meng, M. (2015). The artist emerges: Visual art learning alters neural structure and function. *NeuroImage*, *105*, 440–451.

Sheppard, R., Ropar, D., & Mitchell, P. (2005). The impact of meaning and dimensionality on the accuracy of children's copying. *Journal of Autism and Developmental Disorders*, *37*, 1913–1924.

Sommers, P. V. (1984). *Drawing and cognition: Descriptive and experimental studies of graphic production processes*. Cambridge, England: Cambridge University Press.

Tchalenko, J. (2009). Segmentation and accuracy in copying and drawing: Experts and beginners. *Vision Research*, *49*, 791–800.

Tree, J. J., Horry, R., Riley, H., & Wilmer, J. B. (2017). Are portrait artists superior face recognizers? Limited impact of adult experience on face recognition ability. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 667–676.

Vinter, A. (1999). How meaning modifies drawing behavior in children. *Child Development*, *70*, 33–49.

Webster, M., & Sheets, H. D. (2010). A practical introduction to landmark-based geometric morphometrics. *Quantitative Methods in Paleobiology*, *16*, 163–188.

Wong, Y. K., & Gauthier, I. (2012). Music-reading expertise alters visual spatial resolution for musical notation. *Psychonomic Bulletin & Review*, *19*, 594–600.

Zhou, G., Cheng, Z., Zhang, X., & Wong, A. C.-N. (2012). Smaller holistic processing of faces associated with face drawing experience. *Psychonomic Bulletin & Review*, *19*, 157–162.

Justin Ostrofsky

Department of Psychology, Stockton University