Psychology of Aesthetics, Creativity, and the Arts

Skill Transfer in Visual Arts Expertise
Bonnie L. Angelone, Richard W. Hass, and Marissa Cohen

CITATION
Skill Transfer in Visual Arts Expertise

Bonnie L. Angelone, Richard W. Hass, and Marissa Cohen
Rowan University

Previous research on expertise, both within and outside arts domains, points to moderately narrow transfer of abilities across domain boundaries. The current study tested just how far visual arts abilities transfer by administering a number of speeded visuospatial tasks along with nonvisual tasks to a group of visual arts majors (experts, \( n = 12 \)) and a group of control students (novices, \( n = 15 \)). Visual artists outperformed novices on some of the tasks, most notably on the Building Memory task, which requires both visuospatial location encoding and semantic encoding. Experts also showed superior visual skills on 2 tests that indicate flexibility of closure, which is described as a kind of mental imagery-based ability. Importantly, the experts and novices performed similarly on 2 tasks of logical and mathematical ability, which provides evidence that visual art expertise may not transfer beyond visual tasks. The clearest implication from our results is that of a superiority of quick and accurate visual encoding shown by visual artists. The limitations of the quasi-experimental nature of the study and the small sample size are discussed, along with recommendations for future research.

Keywords: artists, cognition, expertise, visual perception

There is a long history within psychology of studying experts and how they differ from their novice peers (e.g., Chase & Simon, 1973; de Groot, 1946/1978). There is also a recent flurry of studies focusing on the benefits of arts-based curricula, and the transfer of cognitive and noncognitive skills from arts education to non-arts domains (for a review see Winner, Goldstein, & Vincent-Lacrupt, 2013). Research on expertise usually casts experts in terms of reproducible superior performance on specific tasks, such as playing a game of chess or performing a piece of music. The prevailing view is that expert levels of performance are attained mainly through years of deliberate practice (for a review see Ericsson, Nandagopal, & Roring, 2009), but there is also evidence that practice is not the only route to success (Macnamara, Hambrick, & Oswald, 2014). Regardless of the development of their performance skills, experts tend to demonstrate a remarkable depth of knowledge and skill within the domain in which they work. Because of this, many studies of expertise (e.g., Chase & Simon, 1973) have examined whether the depth of knowledge transfers to other task domains. As such, the current study was designed to examine just how far, both within and outside of the visual domain, a group of university-level arts majors’ skills would transfer.

A growing number of studies have focused on the general cognitive underpinnings of expertise in visual arts (e.g., Glazek, 2012; Kozbelt, 2001; Ostrofsky, Kozbelt, & Seidel, 2012), but have left open the question of the limits of transfer from visual arts expertise to other visual domains (but see Winner, Hetland, Veenema, Sheridan, & Palmer, 2006). Studies of the effects of the arts on general cognitive abilities run the gamut from correlational and quasi-experimental to randomized controlled arts-based interventions, with mixed results regarding the precise mechanisms for transfer (Winner et al., 2013). The current study was conceived with a combination of foci; the visual underpinnings of artists, and the degree to which arts-education might lead to general cognitive gains. To move beyond recent studies that have focused almost exclusively on experts in drawing and illustration (e.g., Glazek, 2012; Ostrofsky et al., 2012), we drew our sample from a varied visual arts program, and used tasks that did not require drawing. That is, our study represents a synthesis of approaches best described as tests of transfer from arts education to general cognition (e.g., Winner et al., 2006, Ch. 2), and those best described as cognitive analyses of the process of drawing and illustration. We see this as an important step forward from previous work that also offers the chance to replicate some key findings from the psychology of art. Before describing our study in more detail, we will review some important points from the literature on expertise and then on cognition in the visual arts.

Cognitive Underpinnings of Expertise Across Domains

Several studies have examined performance on various tasks by experts and novices in a number of domains. The classic studies on chess assessed the ability to recall the arrangement of chess pieces after 5 s of viewing for players of varying levels of expertise (masters, class A players and beginners). Experts only outperformed novices when the presented chess pieces are in a mean-
Cognitive Studies of Artists

Artists show cognitive abilities that separate them from non-artists (e.g., Kozbelt, 2001). Visual artists excel at tasks that involve visual spatial ability and they are better at geometrical reasoning due to their enhanced aptitude to visualize and inspect what they see (Walker, Winner, Hetland, Simmons, & Goldsmith, 2011). They show better performance than nonartists in generating and transforming mental images (mental rotation tasks; Getzels & Csikszentmihalyi, 1976), even when scholastic ability is controlled for (Winner & Casey, 1992). In addition, visual artists score higher than nonartists on tests of visual memory (Hermelin & O’Conner, 1986), visual analysis, form recognition (Kozbelt & Seeley, 2007), and change detection (Rosenblatt & Winner, 1988).

In a seminal study, Kozbelt (2001) asked artists and nonartists to complete a set of perceptual tasks that require active visual analysis; some also required drawing and some did not. The nondrawing tasks included naming out-of-focus objects and partially drawn objects, a Gestalt completion task, and a simple embedded figures task. For the drawing tasks, participants had to accurately draw or copy simple and complex objects (these included 2-D and 3-D images). Art students outperformed nonartists on both types of tasks suggesting they possess a general visual cognitive advantage. The question remains; how far will this generalization go?

Some of the arts-based research focuses on transfer of knowledge and skills from arts education programs to more traditional competencies like reading and math scores. (for a review see Winner et al., 2013 and Winner et al., 2006). Some studies have attempted to parse out the various cognitive processes that may be impacted by arts training (e.g., Glazek, 2012; Ostrofsky, Kozbelt, & Seidel, 2012; Kozbelt, 2001). For example, Ostrofsky and colleagues (2012) demonstrated that artists used top-down processing to suppress size constancy cues while drawing indicating that artists have more control over visual attention compared with nonartists. In a similar study, Glazek (2012) demonstrated that expert artists encode visual information with more economy than nonartists and spend more time sketching and less time encoding. Glazek’s results also showed that artists’ encoding abilities transferred to a recognition task involving unfamiliar Chinese ideogram stimuli. In short, artists seem to command their visual cognition in a superior manner compared with nonartists, and also need less time to encode familiar and unfamiliar visual information. Moreover, it remains important to explore the boundaries of transfer by specifically comparing expert and novice abilities that fall far from the expert’s domain. For example, a recent study showed that music training is associated with facilitation of non-native linguistic processing (Perfors & Ong, 2012), suggesting that arts training might transfer to specific but far domains of processing. On the other hand, results of classic studies of general memory in chess players (e.g., Chase & Simon, 1973; see also Ericsson & Charness, 1994) refute the notion that experts possess general cognitive advantages over their novice counterparts. In those studies, experts showed superior memory performance only when recalling stimuli that could be traced directly to their chess-playing experience. The goal of this study was to test the hypothesis that art expertise is associated with general visual processing gains and potentially other more general gains.

Current Hypotheses and Predictions

As previously described, our study represents a synthesis of prior methods, those designed to measure general cognitive ability, and those designed to measure perceptual skill and encoding. We compared the performances of artists and nonartists on a set of visuospatial tasks that varied in representational content, and in
encoding complexity. These tasks are actually speeded cognitive tests, developed in the 1960s and 1970s by psychometricians, and often used in current studies of cognitive ability (e.g., Silvia, Beaty, & Nusbaum, 2013). By utilizing a set of tests measuring various specific cognitive abilities, some of which tapped skills that favored artists, and some that did not, we controlled for familiarity of the experimental context—the two groups were not compared on artistic tasks, which would unfairly favor the artists. We also included two nonvisual tasks to determine whether artists were simply better at performing laboratory tasks, or whether their visual superiority only transferred to visual tasks. The tasks were drawn from a battery of cognitive tests (Kit of Factor-Referenced Cognitive Tests; Eckstrom, French, Harman, & Dermen, 1976) originally designed as indicators of several different cognitive abilities, many of which have been subsumed by the Cattell-Horn-Carroll (CHC) model of general intelligence (e.g., Carroll, 1993; McGrew, 2005). We were not, however, interested in the degree to which these tests were reliable indicators of the various factors within the CHC model. Rather, we were interested in tasks that provided both generic stimulus elements, and multiple task trials, and that did not require drawing. Additionally, these tests allowed for a more straightforward correction for guessing than would be allowed in cognitive reaction time (RT) experiments (cf. Link, 1982).

We chose seven tasks from the battery and administered them to participants in accordance with the procedures described in the test manual. There were five visuospatial tasks—Hidden Figures, Hidden Patterns, Copying, Building Memory, and Idential Pictures—and two nonvisual tasks—Diagramming Relations and Addition/Subtraction Correction. All of which are described in more detail in the Method section. The inclusion of the latter tasks allowed for the control of individual differences in test taking ability.

Our specific hypotheses regarding the differences between artists and nonartists stem from the possibilities of transfer of skill seen in other experts. Having expertise in the visual arts may lead to changes in those who have been trained. If these changes are similar to what is seen in chess masters, then we expect artists to only out-perform nonartists on tasks that are more specific to the visual arts. As such, we hypothesized that artists would most significantly out-perform nonartists on the Building Memory task. Our specific reasoning was that superior performance on the Building Memory task requires both visuospatial location encoding (where is each building?), and semantic encoding (what is each building?), the combination of which should be an important part of composing a visual artwork. We also expected that artists would outperform novices on the copying task as that task is most ecologically similar to drawing, though we did not specifically recruit drawing experts as a part of our sample (i.e., we were interested in art-expertise across the visual arts). The remaining visual tasks—Hidden Figures, Hidden Patterns, and Identical Pictures—seem not to rely on any art-specific skill, so if artists’ visual gains are domain general, similar to that of video gamers, these tasks should also favor artists. Finally, if the advantages seen in artists are fully domain general, we may even expect artists to outperform nonartists on the two nonvisual tasks, however no other studies have found that arts training leads to an improvement in verbal and math abilities (Winner et al., 2006).

**Method**

**Participants**

Twelve expert participants (10 female) were recruited from upper-level courses in the Rowan University Department of Art. The mean age of the expert participants was 21.9 years ($SD = 2.3$). Expert participants were compensated with either extra credit toward an art class (an alternative reading assignment was offered), or with candy provided by the experimenters. Participants were screened to ensure that they qualified as expert artists. All expert participants were currently in their junior or senior years of the art program at the school, and all expected to continue with careers in the arts following graduation. Sixty-seven percent of participants indicated that they engaged in daily artistic activities, whereas the rest estimated their activities to be at least twice a week. Sixty-seven percent of the artists had shown in galleries, both on and off-campus, whereas the rest were planning on showing soon. Finally, in contrast to some previous studies that examined visual artists specifically, the experts identified a range of preferred media including sculpting, painting, jewelry making, printmaking, drawing, illustration, and digital media.

Fifteen novice participants (five females) were recruited via a participation pool; all participants were enrolled in introductory psychology classes and participated in exchange for partial course credit. The mean age of novice participants was 19.3 years ($SD = 1.2$). Novices were also screened for artistic backgrounds, and only two indicated that they engaged in art activities more than once a month. Both of those participants had also taken arts courses, but had not declared an art major. None of the novices reported attending private arts instruction, showing art in a gallery, or any other professional arts activities.

**Materials**

Tests were administered using paper-and-pencil versions of subsets from the Kit of Factor-Referenced Cognitive Tests (hence: CT; Eckstrom et al., 1976): Hidden Figures, Hidden Patterns, Copying, Building Memory, Identical Pictures, Diagramming Relations, and Addition/Subtraction Correction. Tests were administered according to the guidelines given in the CT manual. Though we were not interested in substantiating the existence of the various factors as displayed in the CT manual, the subsets were drawn from a variety of factors described below.

An important component of these tasks is that they are not artistic tasks and, with the exception of the copying task, did not involve any drawing. Rather, these are best thought of as ability and speed tests. Each test is administered in two concurrent timed intervals, separated by a short break. We do our best to describe the visual aspects of each test, but cannot reprint figures because of copyright laws.

**Hidden figures.** The hidden figures test is a five-option multiple-choice test described by the CT manual as measuring

---

1 As a part of the instructions for each of the tasks on the Kit of Factor Referenced Cognitive tests, participants are explicitly informed that their scores for each task will be computed as the number correct minus a fraction of the number incorrect. In accordance with this procedure, participants were also told that they should avoid guessing.
flexibility of closure—“the ability to hold a given visual percept or configuration in mind so as to disembed [sic] it from other well defined perceptual material” (p. 19, Eckstrom et al., 1976). The test is administered in two 12-min intervals (18 items each), separated by a short break between the 18th and 19th item. Each item requires participants to choose from one of five options, the simpler shape that is embedded a complex reference figure that varies from item to item. For example, the complex reference figure might be a square with several different intersecting lines, and the options might be a series of different sized irregular polygons, only one of which can be traced within the target figure.

Hidden patterns. The hidden patterns test is a two-option force-choice test also described in the CT manual as measuring flexibility of closure. The test is administered in two 3-min intervals (200 items each), separated by a short break between the 200th and 201st item. Each item resembles a stick-figure, and requires participants to decided whether a simple reference pattern, comprised of two or three intersecting lines, is hidden within the stick figure. The differences between the hidden patterns task and the hidden figures task are threefold: (a) there is only one reference figure for all 400 items, (b) each item is the more complex of the two line-drawings, and (c) the reference patterns and items are far simpler than those used in the hidden figures task.

Copying. The copying test is the only one of the tests we administered that required drawing. The manual also includes this test among the tests measuring flexibility of closure. The test is administered in two 3-min intervals (32 items each), separated by a short break between the 32nd and 33rd pattern. Each item requires the participants to reproduce a simple line-drawing shown as connecting an array of dots. Correct responses are those in which the participant reproduced the figure exactly as it was shown on a blank array of dots.

Building memory. The building memory test is described by the CT manual as a measure of visual ability or “the ability to remember the configuration, location, and orientation of figural material” (p. 109, Eckstrom et al., 1976). It is a five-option multiple-choice test administered in the following way. First, participants are given 4 min to study a line drawing map consisting of streets and various labeled buildings. Following that, participants briefly view a masking page (hashed grating) before turning the page to view the same map with all of the structures missing. On this test map, sectors of the map are labeled with the letters A through E. Along the side of the map are 12 structures with the letters A through E listed underneath. Participants must correctly determine whether each structure (e.g., a line drawn church) belongs to an A-, B-, C-, D-, or E-labeled section of the map. Participants are given a 4-min interval to complete the 12 items. Then, the entire procedure (study map for 4 min, masking, test map for 4 min) occurs a second time, with an entirely different map. Thus the entire test lasts for 16 min.

Identical pictures. The CT manual describes the identical pictures test as measuring perceptual speed, or “speed in comparing figures or symbols, scanning to find figures or symbols, or carrying out other very simple tasks involving visual perception” (p. 123, Eckstrom et al., 1976). The test is administered in two 90-s intervals (48 items each), separated by a short break between the 48th and 49th item. This is also a five-option multiple-choice test. On each item participants are shown a small, simple line drawing (e.g., a stick figure or a small circular face), printed next to four similar-looking drawings (i.e., missing a feature), and one identical copy of the drawing. Participants must correctly choose the identical drawing.

Diagramming relationships. This is the first of the two tasks that are not described as measuring visual perception. The CT manual describes this test as a test of logical reasoning—“the ability to reason from premise to conclusion, or to evaluate the correctness of a conclusion” (p. 141, Eckstrom et al., 1976). The test is administered in two 4-min intervals (15 items each), separated by a short break between the 15th and 16th item. Items 1 through 15 and 16 through 30 are listed on two separate pages. On top of each page are five Venn diagrams labeled A through E, each of which relays a different syllogistic relationship. For example, one of the diagrams depicts two small circles entirely contained within a larger circle. Each item is a list of three words (categories), and participants must identify which of the five Venn diagrams maps the conceptual relationship among the three words. As an example, item 1 lists the words “dogs, mice, animals,” which is correctly represented by the previously described diagram (two small circles entirely contained in a larger one), signifying that dogs and mice are both members of the larger category of animals.

Addition and subtraction correction. This is the second of the two tasks not measuring visual perception. The CT manual describes the addition and subtraction correction test as a test of number, or “the ability to perform basic arithmetic operations with speed and accuracy” (p. 115, Eckstrom et al., 1976). The test is administered in two 2-min intervals (60 items each), separated by a short break between items 60 and 61. Each item is comprised of an addition or subtraction statement—for example, 120 + 40 = 160. Participants simply mark each statement as correct or incorrect. An example of an incorrect statement would be 120 − 40 = 160.

Testing Procedure

Participants were welcomed to an on-campus research lab and gave informed consent to participate in the experiment. Participants were run in either individual sessions or in groups of two or three. In all cases, participants were seated in library-style separator desks to limit distraction and to discourage copying or cheating. Before each task, an experimenter read the instructions aloud to participants, explicitly asked whether the participants understood, and answered any questions. As previously described, as a part of the directions for each task, participants were informed of the time limit for each task, and encouraged to complete the items quickly and accurately. However, they were also informed that their scores would be adjusted for the presence of incorrect responses. All measures were administered in a fixed order as presented above. Depending on the specific time limits listed for each task, the researcher gave the participants either a 5-min, 1-min, or 30-s warning before time expired.

Upon completion of all of the tests, participants filled out a demographics questionnaire and were debriefed by the experimenter. The critical items on the were two self-report questions about one asking whether and for how long the participant received college-level training in the visual arts, and the other asking the participant to list the courses taken.
Results

Preliminary Analyses

We conducted several preliminary analyses to determine whether it was appropriate to compare experts’ and novices’ performances on each test scored as a whole, using a score that is corrected for guessing based on the total number of items (i.e., before and after the break). To do so, we compared the proportion of items attempted before and after the break using separate t tests for novices and experts. We also examined whether the two groups successfully answered similar proportions of items before and after the break. All data scoring and analysis was performed using the base functions included in the R Statistical Programming Language (R Core Team, 2014). Data are also available at the second author’s Open Science Framework account (https://osf.io/qe3tp/).

Table 1 lists the descriptive statistics for the proportion of attempted items and proportion of correctly attempted items before and after the brief break in each test for both groups. Most comparisons yielded no significant differences in attempts or correct answers before and after the break for both groups. The exceptions were as follows: novices attempted a greater number of items after the break on the Addition/Subtraction Correction test, \(t(12) = -2.21, p = .05, d = -0.57\), novices correctly answered a greater proportion of items after the break on the Diagramming Relations test, \(t(14) = 2.41, p = .03, d = 0.50\), and novices attempted a greater number of items after the break on the Addion/Subtraction Correction test, \(t(14) = 2.20, p = .05, d = 0.25\). Experts attempted a greater number of items before the break on the copying test, \(t(11) = 2.33, p = .04, d = 0.42\), correctly answered a greater proportion of items after the break on the Identical Pictures test, \(t(11) = 3.21, p = .008, d = 0.96\), and attempted fewer items after the break on the Addition/Subtraction Correction test, \(t(11) = -3.91, p = .002, d = -0.55\). With the exception of the experts’ performance on the Identical Pictures test before and after the break, none of the effect sizes were large, and there is little to suggest that either expertise or task demands are driving these effects. Therefore, we elected to perform the primary analysis on pooled performance across the break in each test.

Evaluating Expertise

Performance after adjustment for guessing. The CT manual stipulates that each of the tests, with the exception of the copying task, should be corrected for guessing using formula scoring from classical test theory:

\[
C = R - \left( \frac{W}{k} - 1 \right),
\]

where \(C\) is the corrected total, \(R\) is the number of correct responses, \(W\) is the number of incorrect responses, and \(k\) is the number of multiple-choice options. There is some controversy about whether the classical formula adequately corrects for guessing when participants in the cases in which a participant has some prior knowledge about the answer choices (e.g., Chiu & Camilli, 2013). However, that seems to be more of an issue in aptitude tests than cognitive tests. Additionally, the directions for each test explicitly state that performance will be adjusted for guessing, and this was made clear to participants. That is, participants were made aware that they may be penalized for incorrect responses, and that though they should move as quickly as possible through the items, they should refrain from guessing. Thus, we implemented the classical formula for correction for guessing to each participant’s score on each test, with the exception of the copying task.

The corrected scores, along with the number correct scores on the copying task (task contains no multiple choice options), were then converted to proportions by dividing each corrected total by the number of items spanning both halves of the test. As stated above, each test was administered in two consecutive halves, as stipulated by the manual, to reduce fatigue. However, the numbering of items is continuous across the midpoints indicating that the test should be considered a single unit. Thus, each participant received a single score on each of the seven tests.

Hypothesis testing. Table 2 lists the descriptive statistics and Pearson correlations among the corrected scores all of which were normally distributed. As can be seen, there were significant correlations among the Hidden Figures, Hidden Patterns, and Copying items after the break corrected answered a smaller proportion of items after the break on the Building Memory test, \(t(12) = -2.21, p = .05, d = -0.57\), novices correctly answered a greater proportion of items after the break on the Diagramming Relations test, \(t(14) = 2.41, p = .03, d = 0.50\), and novices attempted a greater number of items after the break on the Addion/Subtraction Correction test, \(t(14) = 2.20, p = .05, d = 0.25\). Experts attempted a greater number of items before the break on the copying test, \(t(11) = 2.33, p = .04, d = 0.42\), correctly answered a greater proportion of items after the break on the Identical Pictures test, \(t(11) = 3.21, p = .008, d = 0.96\), and attempted fewer items after the break on the Addition/Subtraction Correction test, \(t(11) = -3.91, p = .002, d = -0.55\). With the exception of the experts’ performance on the Identical Pictures test before and after the break, none of the effect sizes were large, and there is little to suggest that either expertise or task demands are driving these effects. Therefore, we elected to perform the primary analysis on pooled performance across the break in each test.

Evaluating Expertise

Performance after adjustment for guessing. The CT manual stipulates that each of the tests, with the exception of the copying task, should be corrected for guessing using formula scoring from classical test theory:

\[
C = R - \left( \frac{W}{k} - 1 \right),
\]

where \(C\) is the corrected total, \(R\) is the number of correct responses, \(W\) is the number of incorrect responses, and \(k\) is the number of multiple-choice options. There is some controversy about whether the classical formula adequately corrects for guessing when participants in the cases in which a participant has some prior knowledge about the answer choices (e.g., Chiu & Camilli, 2013). However, that seems to be more of an issue in aptitude tests than cognitive tests. Additionally, the directions for each test explicitly state that performance will be adjusted for guessing, and this was made clear to participants. That is, participants were made aware that they may be penalized for incorrect responses, and that though they should move as quickly as possible through the items, they should refrain from guessing. Thus, we implemented the classical formula for correction for guessing to each participant’s score on each test, with the exception of the copying task.

The corrected scores, along with the number correct scores on the copying task (task contains no multiple choice options), were then converted to proportions by dividing each corrected total by the number of items spanning both halves of the test. As stated above, each test was administered in two consecutive halves, as stipulated by the manual, to reduce fatigue. However, the numbering of items is continuous across the midpoints indicating that the test should be considered a single unit. Thus, each participant received a single score on each of the seven tests.

Hypothesis testing. Table 2 lists the descriptive statistics and Pearson correlations among the corrected scores all of which were normally distributed. As can be seen, there were significant correlations among the Hidden Figures, Hidden Patterns, and Copying items after the break corrected answered a smaller proportion of items after the break on the Building Memory test, \(t(12) = -2.21, p = .05, d = -0.57\), novices correctly answered a greater proportion of items after the break on the Diagramming Relations test, \(t(14) = 2.41, p = .03, d = 0.50\), and novices attempted a greater number of items after the break on the Addion/Subtraction Correction test, \(t(14) = 2.20, p = .05, d = 0.25\). Experts attempted a greater number of items before the break on the copying test, \(t(11) = 2.33, p = .04, d = 0.42\), correctly answered a greater proportion of items after the break on the Identical Pictures test, \(t(11) = 3.21, p = .008, d = 0.96\), and attempted fewer items after the break on the Addition/Subtraction Correction test, \(t(11) = -3.91, p = .002, d = -0.55\). With the exception of the experts’ performance on the Identical Pictures test before and after the break, none of the effect sizes were large, and there is little to suggest that either expertise or task demands are driving these effects. Therefore, we elected to perform the primary analysis on pooled performance across the break in each test.
Table 2
Descriptive Statistics and Intercorrelations Among the Seven Cognitive Tests

<table>
<thead>
<tr>
<th>Task</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Hidden Figures</td>
<td>.38</td>
<td>.24</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2. Hidden Patterns</td>
<td>.47</td>
<td>.11</td>
<td>.46</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3. Copying</td>
<td>.55</td>
<td>.16</td>
<td>.40</td>
<td>.62</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4. Building Memory</td>
<td>.60</td>
<td>.24</td>
<td>.29</td>
<td>.41</td>
<td>.29</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5. Identical Pictures</td>
<td>.84</td>
<td>.13</td>
<td>.12</td>
<td>.45</td>
<td>.10</td>
<td>.16</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>6. Diagramming Relations</td>
<td>.54</td>
<td>.22</td>
<td>.31</td>
<td>.44</td>
<td>.22</td>
<td>.41</td>
<td>.15</td>
<td>—</td>
</tr>
<tr>
<td>7. Addition/Subtraction</td>
<td>.39</td>
<td>.12</td>
<td>.16</td>
<td>.02</td>
<td>.21</td>
<td>.06</td>
<td>.04</td>
<td>.18</td>
</tr>
</tbody>
</table>

Note. All scores are proportions of items correct, adjusted for guessing, except those from the copying tasks (see Results for description).

"p < .05. **p < .01"

tasks, which should be expected given that the fact that the original CT manual lists them as indicators of a common factor of flexibility of closure (Eckstrom et al., 1976). Interestingly, Hidden Patterns scores also correlated highly with all of the other test scores, excluding Addition/Subtraction Correction. Finally, there was a significant correlation between Diagramming Relations and Identical Pictures scores.

As stated previously, our intent was not to confirm a measurement model of these factors, but to use each test as a separate indicator for comparison between participants with and without visual arts training. Instead, we chose to run seven separate analyses of covariance (ANCOVAs)—one for each of the seven tests—with expertise as a fixed factor (coded 1 = expert; 2 = novice) and participant age and gender as covariates. Though Table 2 shows that there are correlations among some of the tests, we felt that an independent examination of each test as the dependent variable in an ANCOVA would be more illustrative regarding our hypotheses. Indeed, we did not expect experts and novices to differ on Diagramming Relations and Addition/Subtraction Correction, so multivariate analyses seemed inappropriate.

As explained, each test score was corrected for guessing using the standard formula, and then converted to a proportion correct score. Overall ANCOVAs for Hidden Patterns, F(3, 23) = 4.74, p = .01, η² = .30; Copying, F(3, 23) = 7.25, p = .001, η² = .42; Building Memory, F(3, 23) = 4.53, p = .01, η² = .29; and Identical Pictures, F(3, 23) = 4.74, p = .01, η² = .30, were all significant, with medium-sized effects. Contrary to prediction, the Hidden Figures ANCOVA failed to reach significance, F(3, 23) = 1.47, p = .25. Table 3 lists the parameter estimates for each significant ANCOVA, with the sign of the coefficients indicating whether novices scored lower (negative coefficients) or higher (positive coefficients) than the experts. Experts (n = 12) significantly out-performed novices (n = 15), after controlling for participant age and gender, on the Hidden Patterns Test (b = −0.16, p = .007), Copying Test (b = −0.22, p = .004), Building Memory Test (b = −0.40, p = .002), and the Identical Pictures Test (b = −0.16, p = .01).

The mean performance (as proportion of correct responses after correcting for guessing) on each test across the two groups are shown on Figure 1, along with bars representing 95% confidence intervals for the means. The confidence intervals were calculated using the “summarySEwithin” function from the Rmisc package for R (Hope, 2013), which calculates confidence intervals for repeated measures data. Taking the confidence intervals into account, the largest difference in performance between experts and novices was on the Building Memory task, with experts correctly identifying 74% of the building placements they answered, compared with just 48% for the novices. Performance was relatively high for both groups on the Identical Pictures task (91% and 79% for the experts and novices, respectively), and lower for both groups on the Hidden patterns task (54% and 42% for the experts and novices, respectively). Finally, artists completed and correctly copied 63% of the patterns on the Copying test compared with 49% correct for the novices.

In support of the case that artists would show some transfer of skill, but not demonstrate transfer of skills beyond visual tasks, the experts and novices did not significantly differ in performance on either the Diagramming Relations test or the Addition/Subtraction tests, as the corresponding ANCOVAs failed to reach significance, F(3, 23) = 0.83, p = .49, and F(3, 23) = 1.32, p = .29, respectively.

Regarding the significant contributions of the age and gender covariates, gender differences emerged on the copying task (b = −0.18, p = .003), on which males (n = 12; M = .59; SD = .14) outperformed females (n = 15; M = .52; SD = .17). The age of participants did not significantly relate to any of the test scores.

**Attempted items.** Though the correction for guessing should, in principle, also adjust for individual differences in the number of attempted items (correct or incorrect), we conducted a final set of ANCOVAs to examine whether experts and novices differed in the number of attempted items, while again controlling for age and gender. As in the preliminary analyses, we considered the proportion of attempted items (items for which the participant provided either a correct or incorrect answer), but this time, pooled attempts across the break in each test. There were no significant main

Table 3
Parameter Estimates of Effects on Corrected Scores (Intercepts Omitted) and 95% Confidence Intervals for the Four Models With Significant Overall ANCOVA Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Estimate (b)</th>
<th>SE (t)</th>
<th>t value</th>
<th>p value</th>
<th>Lower CL</th>
<th>Upper CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hidden Patterns</td>
<td>−.16</td>
<td>.05</td>
<td>−2.29</td>
<td>.007</td>
<td>−.27</td>
<td>−.05</td>
</tr>
<tr>
<td>Gender</td>
<td>−.06</td>
<td>.04</td>
<td>−1.45</td>
<td>.160</td>
<td>−.15</td>
<td>.03</td>
</tr>
<tr>
<td>Age</td>
<td>6.27 × 10⁻⁴</td>
<td>.01</td>
<td>.06</td>
<td>.951</td>
<td>−.02</td>
<td>.02</td>
</tr>
<tr>
<td>Copying</td>
<td>−.22</td>
<td>.07</td>
<td>−3.24</td>
<td>.004</td>
<td>−.40</td>
<td>−.08</td>
</tr>
<tr>
<td>Gender</td>
<td>−.19</td>
<td>.06</td>
<td>−3.36</td>
<td>.003</td>
<td>−.30</td>
<td>−.08</td>
</tr>
<tr>
<td>Age</td>
<td>7.25 × 10⁻³</td>
<td>.01</td>
<td>.53</td>
<td>.604</td>
<td>−.02</td>
<td>.04</td>
</tr>
<tr>
<td>Building Memory</td>
<td>−.40</td>
<td>.11</td>
<td>−3.52</td>
<td>.002</td>
<td>−.64</td>
<td>−.17</td>
</tr>
<tr>
<td>Gender</td>
<td>−.14</td>
<td>.09</td>
<td>−1.49</td>
<td>.151</td>
<td>−.33</td>
<td>.05</td>
</tr>
<tr>
<td>Age</td>
<td>−.03</td>
<td>.02</td>
<td>−1.29</td>
<td>.211</td>
<td>−.08</td>
<td>.02</td>
</tr>
<tr>
<td>Identical Pictures</td>
<td>−.16</td>
<td>.06</td>
<td>−2.68</td>
<td>.013</td>
<td>−.28</td>
<td>−.04</td>
</tr>
<tr>
<td>Gender</td>
<td>.04</td>
<td>.05</td>
<td>.77</td>
<td>.452</td>
<td>.06</td>
<td>.14</td>
</tr>
<tr>
<td>Age</td>
<td>−.02</td>
<td>.01</td>
<td>−1.42</td>
<td>.170</td>
<td>−.04</td>
<td>.01</td>
</tr>
</tbody>
</table>

Note. The comparison group for the expertise variable was expert (visual-arts major), whereas the comparison group for the gender variable was male.
effects of expertise, age, or gender on the following tests: Hidden Figures, $F(3, 23) = 1.10, p = .91$; Copying, $F(3, 23) = 2.36, p = .10$; Building Memory, $F(3, 23) = 1.10, p = .37$; Diagramming Relations, $F(3, 23) = 1.62, p = .21$; and Addition/Subtraction Correction, $F(3, 23) = 2.62, p = .07$. Significant ANCOVA results were found on the Hidden Patterns task, $F(3, 23) = 7.22, p = .001, \omega^2 = .42$, with experts attempting more items than novices ($b = .14, p = .002$). The same pattern (experts > novices, $b = 0.15, p = .02$) was found on the Identical Pictures test, $F(3, 23) = 3.56, p = .03, \omega^2 = .23$.

Discussion

The central goal of this study was to examine whether, and to what extent, undergraduates who chose to become visual art majors show transfer of their visual skills to general visual tasks outside of the domain of art. Prior studies (e.g., Kozbelt, 2001; Ostrofsky, Kozbelt, & Seidel, 2012) suggested that artists do indeed show gains in visual perception and encoding, but that there may be a limit to how far skill transfer extends (Winner et al., 2006). Our results are consistent with the results of those prior studies: visual arts majors showed superior performance on 3 general visual tasks, and a task requiring copying of a visual pattern. Importantly, the effect of expertise (arts major) was significant after accounting for the age and gender of the participants. Further, the visual arts majors did not show superiority on a logical reasoning task or an arithmetic-reasoning task, showing that the artists did not simply “test” higher than the novice participants. We further examined whether the performance differences might be an artifact of a difference in the proportion of items attempted by experts versus novices. In doing so, we found that experts did attempt a greater number of items on the Hidden Patterns and Identical Pictures tests, two tests on which they also outperformed the novices. All participants performed quite well on those tests, but experts completed roughly 10% more items on each test compared with the novices (see Table 1).

We first discuss our interpretation of these results before identifying methodological limitations and suggestions for future work. As described below, we are confident that our results show a positive association between arts training and visual skills, but caution the reader not to infer causality at this point, given the quasi-experimental nature of the study. At the same time, our results are highly consistent with the literature reviewed in the introductory paragraphs, and we will interpret our findings as such.

Visual Advantages of Artists

The clearest implication from our results is that of a superiority of quick and accurate visual encoding shown by a sample of visual arts majors compared to non-arts majors. The Building Memory task, which showed the largest separation between the art majors and novices (relative to the standard errors of the means), relies on accurate encoding and accurate retrieval, as the primary goal is to quickly encode a set of building locations and then reproduce them.
after a delay. The performance advantages on this task shown by artists in our sample connects our results with those of Glazek (2012) who showed that artists generally need less time to encode visual information than nonartists. Our results extend this encoding advantage beyond encoding of features of objects alone, as the Building Memory task requires participants to encode features of the buildings, and as well as the locations of the buildings on a blank map. This means that in addition to encoding features, the artists in our study were more accurate in encoding, and then recalling the spatial location of the buildings. The latter observation relates our results to studies of expert gamers, whose experience analyzing visual scenes also transferred to tasks similar to the Building Memory task (Boot et al., 2008; Dye, Green, & Bavelier, 2009; Green & Bavelier, 2003, 2007). It should be noted that some of our artists had backgrounds in photography, graphic design, and other media in which, presumably, visual scene analysis would be key to constructing a well composed image.

Interestingly, the results stand in contrast with studies of chess masters whose superior encoding skills tend to be constrained to relevant chess positions (e.g., Chase & Simon, 1973). The Building Memory test is similar to the task used by Chase and Simon (see also de Groot, 1946/1978) to determine visual expertise in chess. In the chess experiments, participants were shown a chess board, with pieces in various positions, and then asked to reconstruct the position after the board was hidden from view. In those studies, chess experts tend to only perform better than novices when they are asked to remember positions that would be likely do occur during game play (Chase & Simon, 1973). The literature on chess expertise favors the explanation that chess experts' superior performances result from the need to "look ahead" in the sequence of possible moves of the game in order to select the best move at the time (for a review, see Ericsson & Charness, 1994). That is, chess experts are using visual imagery to access information relevant to the game of chess, and as such, would not be as accurate in recalling chess positions that are nonsensical. The fact that recognition of chess positions can be modeled by rule-based artificial intelligence systems (e.g., Feigenbaum & Simon, 1984), may mean that the visual encoding necessary for chess is only a means to accessing more abstract knowledge about the situation. In contrast, visual analysis in art seems to revolve around continued production and evaluation of compositional and featural changes to works in progress (e.g., Weisberg, 2004; Weisberg & Hass, 2007), something more adequately handled in cognitive simulations by associative neural networks (e.g., Fukushima, 1988). That is, accurate encoding and retrieval in visual artists need not be tethered to more abstract, rule-based cognitive structures. All that matters is that their encoding and retrieval skills are flexible in relation to content of the visual scene subject to analysis. Our art experts' performances on the Building Memory test are a good example of that flexible encoding ability: features and locations of objects in never-before-seen visual scenes are more easily encoded and retrieved by art experts than novices. Again, further research is required to confirm these suspicions, and we encourage other researchers to continue to probe these issues. Based on our evidence, we would expect that this ability should be evident across a wide range of tasks that call for encoding and retrieval of both features and locations of objects.

The performances of our participants on the Hidden Patterns task demonstrate that visual artists may also have enhanced abilities to quickly and accurately extract discrete visual features from more complex arrays. Yet, advantages shown by artists on the Hidden Figures test (see Figure 1) did not reach significance, which stands in contrast to Kozbelt's (2001) results. There are several potential reasons for this. First, a comparison of the tasks across studies revealed that Kozbelt's (2001) Embedded Figures task was much more similar to the Hidden Patterns task used here in that the overall figure was simple and participants were instructed to identify simple target features embedded in a larger object. In contrast, the Hidden Figures task used in the current study contained much more complex figures each of which had the possibility of containing one of five different irregular polygons. As shown in Figure 1, there were large differences between experts and novices on the Hidden Figures task, but also large individual differences within the groups. The Hidden Figures test is also much more difficult—anecdotally, and according to Figure 1—than the Hidden Patterns test, so the lack of the convergence of the findings from these two tests may be biased by the size and characteristics of our sample. At the same time, the direction of the significant difference on the Hidden Patterns test, and the nonsignificant trend on the Hidden Figures test is consistent with previous research suggesting that artists have enhanced perceptual abilities (e.g., Glazek, 2012; Kozbelt, 2001; Ostrofsky et al., 2012). Future studies should continue to examine the extent of generality artists' perceptual skills.

Finally, the results from the Copying task and the Identical Pictures task are less interesting. Although the former provides some information about encoding advantages for artists, we included it mainly as a near-transfer task, which we expected would favor artists. We should emphasize, again, the artists in our sample were not all drawing experts, and also note that the copying task is not a very challenging task. Still, our artists correctly copied 63% of the items on that test compared with the novices’ 49% (see Figure 1). In our opinion, this may be taken as a sort of “quasi-manipulation-check,” such that it confirms that our artists were indeed better at a task that seems to tap a skill that they would likely have been taught in the course of the training—correctly copying a pattern.

Limitations and Future Directions: Is This Truly General Transfer?

Though we have identified some alternative explanations for our results in the previous sections our conclusions about transfer are obviously limited by the quasi-experimental nature of the independent variable—visual arts expertise. This is a necessary limitation that did not limit prior studies from making similar conclusions regarding an association between visual arts experience and visual ability. We cannot, however, firmly establish that the visual artists were not superior to the control participants before commencing with arts.
training. We did take steps to maximize internal validity. For example, both samples were drawn from a student population, and though there were some age differences between the two, all of the results reported here are from ANCOVAs that controlled for age, and found no age-related performance effects. More importantly, our selection of tasks was carefully calculated to isolate different abilities that we hypothesized to favor artists from those that should not favor artists. To that end, we were right on target with our predictions: visual arts students significantly outperformed novices on 3 visual tasks, but were indistinguishable on general intellectual tasks. Still, we urge others to continue to investigate the intriguing association between arts training and visual expertise that we describe.

In terms of the “distance” of the transfer of abilities shown by our expert participants, all of the visual tests we included could be argued to be tapping general visual intelligence (McGrew, 2005). In addition, factor analytic studies suggest that the hidden figures and hidden patterns tests were indistinguishable from other tests more traditionally seen as measures of spatial ability (Spatial Relations Test and Block Design Test; Macleod, Jackson, & Palmer, 1986). However, we found only moderate correlations among different tasks, suggesting that each task represents a narrower indicator of visual skills. As previously noted, our motivation was not to provide evidence to support a general intellectual factor as the basis for artistic skill, but to examine the degree of transfer artists would show across a number of different tasks. Indeed, the pattern of results we observed seems to favor the idea that arts abilities transfer to other narrow visual tasks, but we cannot rule out a priori visual advantages in our artist participants as a causal factor. Again, we urge others to validate our claims with their own samples.

We were also limited by our solicitation of volunteers from art classes on campus, and the stipulation that they be art majors. The university at which the study took place is a large state college, with a relatively small art program. It is reasonable to consider that we may have found different results had we recruited from nearby art colleges. However, the issue of assembling a control group that was similar to the artist group on many demographic variables (academic background, geographic area of origin, socioeconomic status, ethnicity, etc.) would likely be harder when drawing samples from two very different kinds of colleges. We believe, though, that our sampling procedure adequately matched the artists and novices on several unmeasured variables.

Finally, some questions can be raised about the novelty of a study that seems to draw inspiration from a rich group of prior studies. Though part of our intent was to replicate previous work (e.g., Kozbelt, 2001), it was also to extend those findings to a set of tasks not previously examined before. We also compared artists and novices on two nonvisual tasks that had not been previously used by other researchers. We also note that given the weight of the replication crises that have rocked psychology in recent years (e.g., Open Science Collaboration, 2015), it is even more significant to replicate and extend the effect of visual arts expertise on cognitive abilities.

Concluding Remarks

In light of efforts to promote arts-based interventions for struggling school students, our results suggest that a history of visual-arts training is associated with gains in visual encoding, and to a lesser extent retrieval of visual information. However flexible these gains might be, we did not find evidence of a general gain in cognitive abilities as a result of visual arts training. To this end, our findings converge with those of previous studies of visual expertise in artists, as well as studies of chess experts.

References


Received November 24, 2014
Revision received December 10, 2015
Accepted December 31, 2015