Assessing Visual Arts Talents of Hong Kong Chinese Gifted Students: The Development of the Impossible Figures Task

David W. Chan
The Chinese University of Hong Kong

The Impossible Figures Task (IFT-28), which consisted of 28 possible and impossible figures assembled for brief exposure and recognition, was completed by 297 Hong Kong Chinese gifted students. Expert judges (N = 2) rated these students' drawing abilities based on their performance on 2 drawing tasks adapted from Clark's Drawing Abilities Test (Clark, 1989). The IFT-28 scores correlated substantially and significantly with expert ratings of drawing abilities, and with mental rotation test scores. IFT-28 also was found to be sensitive to age effect. The conjecture that IFT-28 scores could reflect drawing abilities and possibly indicate visual arts talents was generally supported even when the effects of age and general spatial ability were controlled for in partial correlation analyses. A brief version (IFT-9) based on retaining 9 figures that discriminated students with high drawing abilities from those with low drawing abilities was suggested as a possibly more sensitive measure for screening students with talents in visual arts.

In Hong Kong, recent education reform efforts that aim to promote students' whole-person development have recognized that the arts offer much to support the academic achievement of students (see Murfee, 1995; Ruppert, 2006; Yu, 2001), and highlighted the important contribution of arts education to students' aesthetic development (Curriculum Development Council, 2002). This recognition is in line with the notion that talents and abilities need to be promoted not only in predominantly academic areas but also in the arts domain. Thus, educators have become increasingly interested in the assessment and development of diverse talents in students, including musical and visual arts talents (see Winner & Martino, 2000, 2003). It also is noted that students may have multiple gifts and talents in several domains (mathematics and music) or in several arts areas (visual

---

David W. Chan is Professor in the Department of Educational Psychology at the Chinese University of Hong Kong.

Journal for the Education of the Gifted. Vol. 31, No. 3, 2008, pp. 240–260.

arts and dance), or may have specializations within one area (painting or sculpting). In this connection, Gardner's (1983, 1999) theory of multiple intelligences provides a useful framework for the identification of different talents (see Chan, 2001). For example, some music educators have come to embrace Gardner's concept of musical intelligence (see Haroutounian, 2002). In contrast, there is no strong support in the domain of visual arts, perhaps because there is no separate identified visual arts intelligence in Gardner's approach. Rather, it is said that each of the multiple intelligences can be directed toward artistic ends. Indeed, the assessment of artistic talents could be very challenging, as visual arts talents could manifest themselves in different visual arts media and in many different ways.

Among the different modes of visual arts expression, drawing has been suggested as the assessment of choice. Clark and Zimmerman (2004), for example, maintained that a drawing task provides a direct and the most appropriate way of identifying high ability in visual arts in students. They cautioned, however, that drawing tasks should not be used as a sole criterion for determining artistic talents and multiple criteria should be employed in identification. Specifically, drawing with pencils and crayons is often the most frequently exercised art activity in school and therefore the least intimidating to students in a testing situation. In addition, drawing tasks are the easiest to assign, administer, and evaluate. More importantly, it also is maintained that only students with visual arts talents will persist in drawing, and drawing abilities are evidence of skills and knowledge in the arts and the art domains (DiLeo, 1977).

Regarding standard drawing tasks, there are few available assessment instruments yielding scores indicative of superior art abilities (Clark & Zimmerman, 2004). Clark and Wilson (1991) have reviewed some of these tests and have found them unsuitable because of the dated nature of illustrations and scoring methods and because these old tests were not developed to be diagnostic of superior abilities in the arts. In view of the lack of suitable identification or assessment instruments, Clark (1989) developed Clark's Drawing Abilities Test (CDAT) for screening and identifying students talented in visual arts for admission to the Indiana University Summer Art Institute. In the test, there are four drawing tasks: Draw an interesting house as if you were looking at it from across the street; draw a person who is running very fast; make a drawing of you and your friends playing in a playground; and make a fantasy drawing from your imagination. These four CDAT tasks are grounded in previous research, and are considered fundamental to art abilities by visual arts teachers (Clark & Wilson, 1991). Specifically, the tasks call for the demonstration of very different abilities, skills, and expressive responses (Clark & Zimmerman, 2004). The house drawing task requires depicting perspective, textures, meaningful shapes and sizes, and recognizable details. The running-person drawing task requires portrayal of actions, as well as body proportions and recognizable details. The persons-in-playground drawing task requires portraying figures accurately, composing in receding space, and grouping figures in that space. The fantasy drawing task provides opportunities for participants to use their imaginations to portray what they wish, the things they know and can draw well. A set of criteria has been formulated for scoring these tasks in terms of originality, expressiveness, and creative solutions as well as drawing skills.

The CDAT has been shown to be valid, reliable, and highly effective as a standardized screening and identification instrument for artistically talented students (Clark & Zimmerman, 2004) and has been adapted for use in research with Chinese students (e.g., Ka, 1999). However, it also is understood that the assessment of a student's visual arts talents based on CDAT drawing tasks requires an expert's judgment on the performance of the student. A problem arises when no visual artist other than the art teacher is generally available in school to serve as an expert judge. It is therefore of interest to examine whether the assessment of visual arts talents could be made independent of expert judgments on drawing tasks.

From a broader perspective, perhaps visual-spatial ability or visual-spatial intelligence in Gardner's multiple intelligences framework come closest to what one would describe as artistic or visual arts talents. However, there are different types of artistic talents. Whereas one artist might have both outstanding two-dimensional and three-dimensional abilities, another might excel in only one area that does not focus on spatial ability. For example, an artist might be successful in decorative painting and drawing, but might not excel in sculpting and architectural rendering. Further, it also is understood that visual-spatial ability is of many different types, and the levels of these different types within individuals can vary widely across areas. Indeed, researchers have distinguished spatial perception, mental rotation, and spatial visualization; spatiotemporal ability; and the generation and maintenance of spatial images (see Halpern & LaMay, 2000; Linn & Peterson, 1985). Some researchers also have maintained that general spatial reasoning ability could be represented and assessed by measures on mental rotation (see Casey, Nuttall, Pezaris, & Benbow, 1995; Jackson, 2003). Evidently, given that there are different types of visual-spatial ability, one would expect that certain specific types of visual-spatial ability might have more relevance to visual arts talents than others.

In the search for specific types of visual-spatial ability indicative of visual arts talents, the research by Winner and her colleagues on the association of dyslexia with visual-spatial ability is noteworthy (Von Karolyi, 2001; Von Karolyi & Winner, 2004; Von Karolyi, Winner, Gray, & Sherman, 2003; Winner, French, Seliger, Ross, & Weber, 2001). In their studies, they compared the performance of individuals with and without dyslexia on a task in recognizing impossible figures. Impossible figures contain surface or edge violations that prevent them from existing as three-dimensional structures (see Carrasco & Seamon, 1996). Based on the observations that individuals with dyslexia report visual-spatial strengths, the reported elevated incidence of dyslexia in visual artists, and the findings that individuals with dyslexia were found to recognize impossible figures more rapidly, but no less accurately than those without dyslexia, Von Karolyi and Winner suggested that a visual-spatial task testing the speed of recognition of impossible figures might provide an indicator of visual arts talent. They also hypothesized that the recognition of impossible figures requires global visual-spatial processing ability. Accordingly, a respondent who scans an impossible figure part by part, but fails to integrate the parts, will be led to conclude that the figure is possible (Mottron & Belleville, 1993). Only by scanning globally, or by holistic inspection, will the respondent be able to recognize that the parts conflict and that the figure is therefore impossible (Schacter, 1992). Von Karolyi and Winner also suggested that such global visual-spatial abilities may be distinct from other kinds of visual-spatial abilities and may also underlie important real-world activities, including visual artistry. Thus, based on these considerations, an impossible figures task might offer a promising lead to the assessment and identification of visual arts talents in students, especially those with outstanding spatial abilities.

Along this line, this study explored the relationship between students' performance on recognizing impossible figures and their drawing abilities in a sample of Chinese gifted students in Hong Kong. Specifically, 30 possible and impossible figures were first assembled for students to identify. The relationship between students' accuracy scores in identifying impossible figures and their drawing abilities judged by two visual artists on their performance in two CDAT drawing tasks was examined in relation to age and general spatial ability. A brief impossible figures task retaining figures that significantly discriminated between students high in drawing abilities and those low in drawing abilities was developed as a measure for screening visual arts talents for future investigations.

Method

Participants

A total of 297 Chinese students, 181 primary (grades 3 to 6) and 116 secondary (grades 7 to 12) students, participated voluntarily in this study. These students (183 boys and 114 girls), aged 7 to 19 (M = 11.11, SD = 2.36), were nominated by their schools to participate in a variety of gifted programs provided at different times at the Chinese University of Hong Kong over a period of 3 months. In nominating students, schools were requested to recommend students who were judged to be either gifted intellectually (e.g., with a high IQ score), or academically (e.g., with outstanding performances in school subjects), or had demonstrated talents in other specific nonacademic areas such as music, art, and leadership. Because there were no generally accepted standard measures in Hong Kong schools and schools often did not have access to information on specific IQ scores of students, teachers making recommendations would use their own judgment based on their knowledge of their students, bearing in mind that students could be regarded as gifted in one or more domains

(Education Commission, 1990). Thus, this sample of participants could be regarded as relatively heterogeneous in terms of their giftedness or talents and represented students from a broad age range.

Measures and Tasks

The Drawing Tasks. The house and the running-person drawing tasks adapted from CDAT were used in this study. These two drawing tasks, drawing a person (not necessarily a running person) and a house, were common drawing exercises in the repertoire of drawing experience of Chinese students and were therefore chosen among the four CDAT drawing tasks. The CDAT that contains four drawing tasks has been used and tested with more than 5,000 upper elementary, middle school, and high school students in the United States and other countries and has been shown to be valid, reliable, and highly effective as a standardized screening and identification measure for artistically talented students (Clark & Zimmerman, 2004). Clark and Zimmerman also reported that scores on the CDAT drawings correlated significantly with rankings of student success in classes for artistically talented students in summer arts institutes by teachers who used forms designed by Clark. In this study, because the two enlisted expert judges had not been trained to use the standardized CDAT criteria for judgments, they were only requested to make global ratings on students' drawing skills based on each of the students' two drawings on a 3-point scale, 1 (*below average*), 2 (*average*), and 3 (*above average*). Thus, although the drawing tasks were CDAT drawing tasks, the judgments and scoring were not done as in the standardized procedure.

The Mental Rotation Test. The spatial test of Jackson's (2003) Multidimensional Aptitude Battery II, the Mental Rotation Test (MRT), was employed in this study to represent general spatial ability. The test assesses respondents' ability to visualize abstract objects in different positions in two-dimensional space and to be sensitive to critical differences among alternatives. More generally, it requires reasoning in the figural-spatial domain combined with visual and imaginal processes that need to be evoked quickly and automatically. MRT thus involves testing respondents' mental rotation ability in a two-dimensional space. Each of the 50 items of the test consists of a criterion figure presented to the left and five alternatives to match on the right. Respondents were given 10 minutes to complete the 50 items. Hence, an excessive degree of checking responses because of cautiousness will impair the speed of performance. Points were given for each correct response. The maximum score on MRT was 50 points. Age has been found to affect performance substantially (Jackson, 2003).

Jackson (2003) reported that the spatial test correlated (r = .44) with Block Design of the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981) in the original study of test development based on N = 145 and in the range of .74 to .97 (M = .88) in different studies. It correlated (r = .44) with Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998) in a study based on N = 103 originally published by Kranzler and Jensen (1991). Psychometric properties of these tests have been reported in different studies (Kranzler, 1991; Krieshok & Harrington, 1985; Wallbrown, Carmin, & Barnett, 1988, 1989), and summarized by Jackson.

The Impossible Figures Task. The Impossible Figures Task (IFT-28) consists of 30 figures (13 possible figures and 17 impossible figures) assembled for brief display on a screen (2 seconds per figure) for students. These figures were selected either from those used by Schacter, Cooper, and Delaney (1990) or from the Figures library on the Web site of Impossible World (n.d.). An initial version has been tested in pilot studies with Chinese children prior to this study. In completing the task, students were asked to respond "impossible" or "possible" to these figures. The first two items were used as examples of one impossible figure and one possible figure. IFT-28 yielded a summary score by aggregating the correct responses on the 28 figures for each student, reflecting the student's global visual-spatial ability.

Procedure

All 297 nominated students participated voluntarily with the consent of their parents in a larger research project of which this study was a part. These students, in groups of 30 to 50, were requested to complete self-report questionnaires that included demographic information and drawing activities. Of particular relevance to this study were their completion of the IFT-28 and the MRT and their drawing-lesson experience (responses of "yes" or "no" to whether they had taken drawing lessons outside regular classes). In addition to these self-report data, students also completed two 15-minute drawing tasks based on the CDAT tasks of drawing a house and a running person. The drawings of students were evaluated as *below average*, *average*, or *above average* by two expert judges who made their judgments independently. In making expert judgments, the judges were provided with the six-sample CDAT drawings on the house and on the running person from Clark and Zimmerman (2004). The two expert judges (one man and one woman) enlisted for this study are recognized visual artists in Hong Kong.

Results

Judges' Ratings on Students' Drawing Abilities

Each of the students' two drawings (the running person and the house) was first rated independently by the two expert judges on the three-point scale. However, one judge also gave half-points (e.g., 2.5). Correlations between the ratings of the two judges were positive and significant ($r = .38$ for house, $r = .49$ for running person, and $r = .50$ for the combined ratings, all ps < .001). It was decided that the four ratings on a student's two drawings by the two expert judges could be aggregated to yield a global score (4 to 11) reflecting the drawing ability of the student. Based on the global score, students were arbitrarily divided into three groups of low, medium, and high drawing abilities. Specifically, because there were many more students with scores of 4 and very few students with scores above 6, it was decided that these cutoff scores were appropriate, yielding a distribution of 56.6%, 29.6%, and 13.8% for low, medium, and high drawing-ability groups, respectively.

To explore whether students' membership in the three drawing-ability groups was associated with their gender, grade level, and drawing-lesson experience, separate cross-tabulations were conducted.

The results indicated that the associations of students' drawing-ability group membership with gender, $\chi^2(2, N = 297) = 5.17$, Cramer's $V = 0.13$, and with drawing-lesson experience, $\chi^2(2, N = 297) = 2.48$, Cramer's $V = 0.09$, were nonsignificant ($p > .05$), but its association with grade level (primary vs. secondary), $\chi^2(2, N = 297) = 18.72$, Cramer's $V = 0.25$, was significant ($p < .001$), suggesting that the drawing abilities of secondary students were rated higher than those of primary students by the judges who had no knowledge of the age or grade level of the students.

Students' Performance on the Impossible Figures Task

The responses of the students to each of the 28 figures in IFT-28 (omitting the two sample figures) were scored 1 (correct) or 0 (incorrect). The mean scores for the figures are shown for the total sample and separately for the low, medium, and high drawing-ability groups in Table 1. These mean scores also were the endorsement proportions (the proportions of students endorsing the correct responses). For example, for item 25 (Figure 25), 55% of the students (48%, 58%, and 76% of students in the low, medium, and high drawing-ability groups, respectively) recognized this figure as a possible figure. Mean comparisons also were made to examine whether each of the 28 figures could significantly discriminate the low drawing-ability group from the high drawing-ability group. Using the arbitrary cutoff that a figure that could significantly discriminate the low and high drawing-ability groups ($p < .05$) could be considered to be included as good discrimination figures, the results indicated that four possible figures (Figures 25, 16, 3, and 26) and five impossible figures (Figures 19, 5, 14, 27, and 11) could be so included (as shown in Table 1). As a 28-item scale, IFT-28 achieved moderate internal consistency (Cronbach's $\alpha = .63$), as would be expected from dichotomous items tapping different aspects that made figures impossible.

Drawing-Ability Group Differences, and Gender and Grade-Level Differences on IFT-28

To explore whether students in low, medium, and high drawing-ability groups did perform significantly differently on IFT-28, a one-way

Table 1

Mean Ratings of Students' Responses to the 28 Figures of the Impossible Figures Task

	Mean Rating (Endorsement Proportion)				Difference Between Low and High Group
	Total ($N = 297$)	Low ($n = 168$)	Medium ($n = 88$)	High ($n = 41$)	$F(1, 207)$
Possible Figures					
25	.55	.48	.58	.76	10.34**
16	.72	.67	.76	.88	7.31**
3	.91	.89	.91	1.00	5.18*
26	.76	.70	.83	.85	4.15*
20	.76	.71	.80	.85	3.36
17	.82	.79	.85	.88	1.59
24	.40	.41	.33	.51	1.38
18	.62	.58	.65	.68	1.36
30	.81	.79	.84	.83	0.29
9	.80	.78	.83	.80	0.12
4	.85	.85	.88	.83	0.06
7	.84	.85	.84	.83	0.06
Impossible Figures					
19	.70	.62	.77	.85	8.39**
5	.68	.62	.72	.83	6.64*
14	.32	.26	.39	.44	5.03*
27	.75	.68	.82	.85	4.71*
11	.84	.79	.90	.93	4.09*
12	.88	.85	.90	.95	2.95
21	.76	.74	.77	.85	2.42
29	.41	.39	.41	.51	2.14
10	.84	.82	.86	.90	1.79
23	.66	.65	.64	.76	1.71
22	.79	.74	.85	.83	1.31
15	.79	.79	.78	.83	0.29
13	.85	.86	.84	.83	0.20
8	.53	.50	.59	.54	0.18
6	.66	.65	.69	.66	0.01
28	.63	.63	.65	.63	0.01

Note. Items are arranged separately for possible and impossible figures and in descending order of differences between the low and high drawing ability groups. The good discrimination figures are those with significant group differences with $p < .05$ and effect size statistic partial $\eta^2 > .019$. *$p < .05$; **$p < .01$.

analysis of variance (ANOVA) was conducted. The results indicated that the drawing-ability group main effect was significant, F (2, 294) = 13.82, $p < .001$, partial $\eta^2 = .086$. Follow-up post hoc tests using Bonferroni procedure indicated that the low drawing-ability group (M = 19.08, SD = 3.52) scored significantly lower ($p < .01$) than the high drawing-ability group (M = 21.90, SD = 3.18) and the medium drawing-ability group (M = 20.67, SD = 3.45), but the difference between the high drawing-ability group and the medium drawing-ability group did not achieve significance ($p = .181$), suggesting that IFT-28 scores could distinguish students of low and high drawing abilities.

To explore whether there were gender or grade level (primary vs. secondary students) differences on students' IFT-28 scores, a 2 × 2 (Gender × Grade Level) ANOVA was performed. The results indicated that the grade-level main effect was significant, F (1, 293) = 72.72, $p < .001$, partial $\eta^2 = .199$, suggesting that secondary students performed significantly better than did primary students. The gender main effect and the gender/grade-level interaction effect however were not significant ($p > .05$).

Correlations Between Students' IFT-28 Performance and Judged Drawing Abilities

Table 2 shows the correlations between students' performance on IFT-28 and their drawing abilities rated by two expert judges on the two drawing tasks. Because students' experience in drawing lessons outside regular classroom and their general spatial reasoning ability might be related to their drawing abilities, the variables of drawing-lesson experience and MRT scores were included in the computation. Also included was age, as spatial tests were known to be sensitive to age. The results indicated that, in general, students' IFT-28 scores correlated significantly with the judges' ratings on students' drawing abilities, suggesting that IFT-28 could reflect drawing abilities. The low and nonsignificant correlations between IFT-28 and drawing-lesson experience suggested that training in drawing might not affect IFT-28 performance, and drawing-lesson experience also correlated minimally with expert ratings on students' drawing abilities (r = .02 to .07), although 70% of the students checked yes to whether they had drawing lessons outside school. The substantial correlations

Table 2

Correlation Between Students' Performance on Impossible Figures Task (IFT-28) and Students' Drawing Abilities, Drawing-Lesson Experience, Performance on Mental Rotation Test, and Age (*N* = 297)

		Correlation With		Partial Correlation With IFT-28 Controlling for		
	IFT-28	MRT	Age	MRT	Age	MRT and Age
Judge 1						
Rating of Person	16**	03	15**	17**	11	13*
Rating of House	14*	10	20***	11	06	06
Aggregated Ratings	18**	08	21***	16**	10	11
Judge 2						
Rating of Person	27***	16**	31***	23***	16**	16**
Rating of House	25***	18**	24***	19***	17**	15*
Aggregated Ratings	30***	20***	32***	25***	20***	18**
Composite Ratings by Two Judges	29***	17**	31***	24***	18**	17**
Drawing-Lesson Experience	-05	-11	-16**	-00	02	04
MRT	42***	-	42***	-	30***	-
Age	43***	42***	-	30***	-	-

Note. IFT-28 = 28-item Impossible Figures Task. Decimals are omitted. MRT = Mental Rotation Test. *p < .05; **p < .01; ***p < .001.

between IFT-28 and MRT, as well as age, suggested that IFT-28 could be related to general spatial reasoning and could be sensitive to age, with older students performing better than younger students.

It also was of interest to compare the above pattern of correlations replacing IFT-28 with MRT and age as shown in Table 2. The significant correlations between age and all included variables suggested that judges' ratings and MRT were sensitive to age, implying that older students tended to obtain higher ratings on their drawing abilities and higher scores on MRT. However, the negative correlation between age and drawing-lesson experience suggested that younger students were more likely to seek training in drawing lessons. As expected, MRT was sensitive to age, and MRT correlated significantly with one expert judge's (Judge 2) ratings, suggesting this judge might consider general spatial reasoning in the ratings of students' drawing abilities.

Although the correlations between IFT-28 and judged drawing abilities were generally substantial and significant, the significant correlations between age and judged drawing abilities and the significant correlation between age and IFT-28 raised the question whether age could account for the association between IFT-28 and judged drawing abilities. Similarly, the same question could be raised for MRT in place of age. Consequently, to test this conjecture, partial correlations controlling age, MRT, and age and MRT were computed (see Table 2). Sizable reductions in the magnitude of correlations or changes of significant correlations to nonsignificant correlations might imply that the correlation between IFT-28 and drawing abilities was spurious. The results, however, indicated that there were moderate attenuations in correlations, suggesting that although age and general spatial ability might account partly for the correlation between IFT-28 performance and drawing abilities, the correlation was real, and IFT-28 performance could indeed reflect students' drawing abilities and might be indicative of visual arts talents in students.

The Brief IFT-9

Although IFT-28 could be employed as an initial screening measure for good drawing abilities or visual arts talents, its sensitivity as a measure could be improved. One approach was to develop a brief measure retaining only the discriminating figures of IFT-28 based on the data in Table 1. To test this conjecture, the 9-item IFT-9 was developed by retaining only the nine figures that significantly discriminated students with high drawing abilities from students with low drawing abilities. As a 9-item scale, IFT-9 achieved moderate and slightly lower internal consistency (Cronbach's $\alpha = .52$) than the full 28-item scale, but by adjusting for length using the Spearman-Brown formula, the estimated value of .77 would suggest that the brief scale was more internally consistent. The figure most discrepant from other figures was Figure 14 (Roger Penrose's ascending and descending staircases), which was also the most difficult item, as reflected in the lowest endorsement frequency; only 32% of the students scored correctly on Figure 14. This brief IFT-9 scale correlated .79 with the full IFT-28 version.

To check for the comparability and perhaps effectiveness of the brief IFT-9 with the full IFT-28 version, group differences analyses, correlation, and partial correlation analyses conducted with the IFT-28 were repeated for IFT-9. The ANOVA results on drawing-ability group differences indicated that the drawing-ability group main effect was significant, $F\,(2, 294) = 21.47$, $p < .001$, partial $\eta^2 = .127$. Follow-up post hoc tests using Bonferroni procedure indicated that the low drawing-ability group ($M = 5.71$, $SD = 1.77$) scored significantly lower ($p < .001$) than the high drawing-ability group ($M = 7.39$, $SD = 1.32$) and the medium drawing-ability group ($M = 6.67$, $SD = 1.58$), but the difference between the high drawing-ability group and the medium drawing-ability group only approached significance ($p = .067$), suggesting that IFT-9 scores could distinguish students of low and high drawing abilities. The ANOVA results on gender and grade-level differences indicated that the grade-level main effect was significant, $F\,(1, 293) = 33.30$, $p < .001$, partial $\eta^2 = .102$, suggesting that secondary students performed significantly better on IFT-9 than did primary students. The gender main effect and the gender/grade-level interaction effect, however, were not significant ($p > .05$). Thus, IFT-9 and IFT-28 yielded a similar pattern of results on group differences.

Table 3 summarizes the correlation and partial correlation analyses conducted for IFT-9. Although a similar pattern of results emerged, the reduction in correlations between IFT-9 and ratings of drawing abilities, controlling for age and MRT, was less sizable, suggesting that the brief measure could be more sensitive and could be used in further refinements in the development of impossible figures task in screening for visual arts talents.

Discussion

The assessment and identification of students with visual arts talents in the Hong Kong school setting presents a special challenge to teachers and educators. For one thing, although asking a student to

Table 3

Correlation Between Students' Performance on the Brief Impossible Figures Task (IFT-9) and Students' Drawing Abilities, Drawing-Lesson Experience, Performance on Mental Rotation Test, and Age (*N* = 297)

	Correlation With			Partial Correlation With IFT-9 Controlling for		
	IFT-9	MRT	Age	MRT	Age	MRT and Age
Judge 1						
Rating of Person	18**	03	15**	18**	14*	16**
Rating of House	23***	10	20***	21***	18**	18**
Aggregated Ratings	25***	08	21***	23***	20***	21***
Judge 2						
Rating of Person	28***	16**	31***	24***	20***	20***
Rating of House	29***	18**	24***	25***	24***	22***
Aggregated Ratings	34***	20***	32***	29***	26***	25***
Composite Ratings by 2 Judges	34***	17**	31***	30***	27***	27***
Drawing-Lesson Experience	-09	-11	-16**	-05	-04	-03
MRT	35***	-	42***	-	26***	-
Age	30***	42***	-	18***	-	-

Note. IFT-9 = 9-item Impossible Figures Task. Decimals are omitted. MRT = Mental Rotation Test. **p* < .05; ***p* < .01; ****p* < .001.

complete a drawing task is generally accepted as a simple and effective way to assess the student's drawing abilities, which might in turn reflect his or her visual arts talents, it is also recognized that an expert judge or visual artist needs to be involved in the assessment. In this connection, the use of a standardized measure such as CDAT, which has been found to be reliable and valid based on four drawing tasks, could be helpful. However, it should be noted that the two drawing tasks adapted for use in this study might favor rewarding representational or realistic drawings, thus disadvantaging students who do not draw figuratively or realistically (see Milbrath, 1998). The fourth task, which does not require realistic drawing ability and is judged mainly on inventiveness and creativity, was not used in the study.

Moreover, for the two CDAT drawing tasks employed in this study, the scoring did not follow the standardized procedure. Nonetheless, it is believed that the development of a visual-spatial task assessing visual arts talents of students independent of expert judgments would be of great value.

With this view, IFT-28 was developed in this study as a task testing students' global visual-spatial ability based on the recognition of impossible figures, following the pioneering work by Winner and her colleagues on the association of dyslexia with visual-spatial ability (see Von Karolyi & Winner, 2004) and using some of the possible and impossible figures in past studies (Schacter et al., 1990). The findings indicated that IFT-28 scores did discriminate students with high drawing abilities from students with low drawing abilities. Partial correlation analyses also indicated that the association between IFT-28 performance and drawing abilities was real, even though age and general spatial ability could account for part of the association. Thus, this study provided some initial support for the use of impossible figures in screening for visual arts talents.

To evaluate whether each of the figures used could discriminate students with high drawing abilities from those with low drawing abilities, global expert ratings elicited from two visual artists of the students' two CDAT drawing tasks were used to define the drawing-ability groups for comparison. The results indicated that nine figures could significantly discriminate the drawing-ability groups; it would be of interest to explore further why they were more discriminating. Future studies also might consider expanding the sampling of impossible figures, as well as employing more rigorous criteria for drawing-ability group comparison in selecting discriminative possible and impossible figures to be included in revised versions of the task. Thus, while the present IFT-9 could be suggested as a brief and sensitive measure, the need for cross-replication in the future development of IFT should be emphasized.

Admittedly, questions can be raised as to whether expert judgments on students' drawing abilities based on CDAT drawings could differ very much from one visual artist to another and whether using a different pair of experts could result in a different classification of students for drawing-ability group comparison in the choice of IFT figures. Because the two expert judges in this study have not been trained to make judgments based on the CDAT criteria, these questions could not be readily and adequately addressed. The decision that judges would only be asked to make very global ratings in terms of *below average*, *average*, and *above average* and the use of aggregated ratings from the two expert judges make it less likely that a radically different classification of students could result from a change of experts. Besides, the sample drawings with global ratings on drawing-ability levels from Clark and Zimmerman (2004) were made available to the present judges for comparison. Perhaps future studies could either enlist judges who are trained to use the CDAT criteria in making judgments or elicit specific implicit criteria from expert judges to make expert judgments less subjective and more consistent.

Certainly, there are many limitations in the present study. One limitation, among others, is the selection of the present sample of gifted students as participants of this study, as this student sample was nominated by teachers who could be biased in nominating only high academic achievers, even though teachers were urged to nominate students with talents in academic, as well as nonacademic, domains, including visual arts. In addition, it was understood that teacher judgment alone often is not a valid indicator of talent in many areas unless teachers are provided with specific criteria in judgment (see Hany, 1997). While it was hoped initially that enlisting gifted students with diverse talents that included visual arts talents might place fewer restrictions on the upper range of artistic abilities to be assessed, it was not known whether one could miss out on the lower end of the ability spectrum. Perhaps another limitation was the broad age range of respondents in this study. Because spatial tests such as MRT and the present IFT are highly sensitive to age differences, one has to decide whether future development of the revised IFT should include separate versions for children and for adolescents. Further, it was puzzling that training or drawing-lesson experience did not correlate with drawing-ability ratings. Perhaps respondents need to elaborate on their training beyond a simple "yes" or "no," as in the present study. Nonetheless, these and other problems need to be more carefully considered and addressed in future studies.

In summary, this study provided some initial data supporting the use of IFT and an opportunity to invalidate the claim that students' performance on IFT could reflect their drawing abilities. The substantial correlation observed between IFT and drawing abilities, even when the effects of age and general spatial ability were controlled, suggested that IFT could be further developed as a screening measure for visual arts talents. Future studies could be directed to further exploration of this connection to provide more insight into the avenues through which visual arts talents could be assessed and identified among students.

References

Carrasco, M., & Seamon, J. G. (1996). Priming impossible figures in the object decision test: The critical importance of perceived stimulus complexity. *Psychonomic Bulletin and Review, 3,* 344–351.

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31,* 697–705.

Chan, D. W. (2001). Assessing giftedness in Chinese secondary school students in Hong Kong: A multiple intelligences perspective. *High Ability Studies, 12,* 215–234.

Clark, G. (1989). Screening and identifying students talented in the visual arts: Clark's Drawing Abilities Test. *Gifted Child Quarterly, 33,* 98–105.

Clark, G., & Wilson, T. (1991). Screening and identifying gifted/talented students in the visual arts with Clark's Drawing Abilities Test. *Roeper Review, 13,* 92–97.

Clark, G., & Zimmerman, E. (2004). *Teaching talented art students: Principles and practices.* New York: Teachers College, Columbia University.

Curriculum Development Council. (2002). *Arts education key learning area curriculum guide (Primary 1–Secondary 3).* Hong Kong: Government Printer.

DiLeo, J. H. (1977). *Child development: Analysis and synthesis.* New York: Brunner/Mazel.

Education Commission. (1990). *Education Commission Report No. 4.* Hong Kong: Hong Kong Government.

Gardner, H. (1983). *Frames of mind.* New York: Basic Books.

Gardner, H. (1999). *Intelligence reframed.* New York: Basic Books.

Halpern, D. F., & LaMay, M. L. (2000). The smarter sex: A critical review of sex differences in intelligence. *Educational Psychology Review, 12,* 229–246.

Hany, E. A. (1997). Modeling teachers' judgment of giftedness: A methodological inquiry of biased judgment. *High Ability Studies, 8,* 159–178.

Haroutounian, J. (2002). *Kindling the spark: Recognizing and developing musical talent*. New York: Oxford University Press.

Impossible World. (n.d.). *Figures library*. Retrieved November 10, 2006, from http://im-possible.info

Jackson, D. N. (2003). *Multidimensional Aptitude Battery II manual.* Port Huron, MI: Sigma Assessment Systems.

Ka, L. (1999). The relationship between drawing ability and academic ability among primary one students in Hip Mei School [in Chinese]. In S. Wong (Ed.), *Art education: Research and horizon* (pp. 71–78). Hong Kong: Hong Kong Society for Education in Art.

Kranzler, J. H. (1991). The construct validity of the Multidimensional Aptitude Battery: A word of caution. *Journal of Clinical Psychology, 47,* 691–697.

Kranzler, J. H., & Jensen, A. R. (1991). The nature of psychometric g: Unitary process or a number of independent processes? *Intelligence, 15,* 397–422.

Krieshok, T. S., & Harrington, R. G. (1985). A review of the Multidimensional Aptitude Battery. *Journal of Counseling and Development, 64,* 87–89.

Linn, M., & Petersen, A. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56,* 1479–1498.

Milbrath, C. (1998). *Patterns of artistic development in children: Comparative studies of talent.* Cambridge, UK: Cambridge University Press.

Mottron, L., & Belleville, S. (1993). A study of perceptual analysis in a high-level autistic subject with exceptional graphic abilities. *Brain and Cognition, 23,* 279–309.

Murfee, E. (1995). *Eloquent evidence: Arts at the core of learning.* Washington, DC: National Assembly of State Arts Agencies.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scale.* Oxford, UK: Oxford Psychologists Press.

Ruppert, S. S. (2006). *Critical evidence: How the arts benefit student achievement*. Washington, DC: National Assembly of State Arts Agencies.

Schacter, D. L. (1992). Understanding implicit memory: A cognitive neuroscience approach. *American Psychologist, 47,* 559–569.

Schacter, D. L., Cooper, L. A., & Delaney, S. M. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General, 119,* 5–24.

Von Karolyi, C. (2001). Visual spatial strength in dyslexia: Rapid discrimination of impossible figures. *Journal of Learning Disabilities, 34,* 380–391.

Von Karolyi, C., & Winner, E. (2004). Dyslexia and visual spatial talents: Are they connected? In T. M. Newman & R. J. Sternberg (Eds.), *Students with both gifts and learning disabilities: Identification, assessment, and outcomes* (pp. 95–117). New York: Kluwer Academic/Plenum.

Von Karolyi, C., Winner, E., Gray, W., & Sherman, G. F. (2003). Dyslexia linked to talent: Global visual-spatial ability. *Brain and Language, 85,* 427–431.

Wallbrown, F. H., Carmin, C. N., & Barnett, R. W. (1988). Investigating the construct validity of the Multidimensional Aptitude Battery. *Psychological Reports, 62,* 871–878.

Wallbrown, F. H., Carmin, C. N., & Barnett, R. W. (1989). A further note on the construct validity of the Multidimensional Aptitude Battery. *Journal of Clinical Psychology, 45,* 429–433.

Wechsler, D. (1981). *Manual of the Wechsler Adult Intelligence Scale–Revised.* New York: Psychological Corporation.

Winner, E., French, L., Seliger, C., Ross, E., & Weber, C. (2001). Dyslexia and visual-spatial talents: Compensation vs. deficit model. *Brain and Language, 76,* 81–110.

Winner, E., & Martino, G. (2000). Giftedness in non-academic domains: The case of the visual arts and music. In K. A. Heller, F.

J. Mönks, R. J. Sternberg, & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2nd ed., pp. 95–110). New York: Elsevier.

Winner, E., & Martino, G. (2003). Artistic giftedness. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (3rd ed., pp. 335–349). Boston: Allyn & Bacon.

Yu, S. T. (Ed.). (2001). *The revolution of arts education* [in Chinese]. Hong Kong: Hong Kong Society for Education in Art.

Author Note

This study was supported in part by a direct grant for research from the Chinese University of Hong Kong.

Correspondence concerning this article should be addressed to David W. Chan, Department of Educational Psychology, Faculty of Education, the Chinese University of Hong Kong, Shatin, NT, Hong Kong. E-mail: davidchan@cuhk.edu.hk.