

Assessment of Potential Theater Arts Talent in Young People: The Development of a New Research-Based Assessment Process

Barry Oreck, Susan Baum, and Steven Owen

We report on the results of a new process for identifying potential theater talent in a diverse population of elementary school students.

INTRODUCTION

*God I hope I get it, I hope I get it
How many people does he need?
... how many boys, how many girls?
Step, kick, kick leap, kick touch.
Got it? ... going on. And. ...
from A Chorus Line*

The traditional audition can be a highly stressful, memorably discouraging experience. Yet, this is the model for many of the assessment processes used in the performing arts in education. Selection for magnet arts schools, advanced theater programs, gifted and talented programs and of course, school plays, are frequently conducted as one-time, on-stage presentations in which a scene, reading, song, or improvisation is performed in front of an audience of one or more assessors. Though efficient as a means to cast a play, many aspects of this process are incompatible with fair, authentic, and accurate assessment of student potential. The lack of formal drama instruction in most elementary schools in the United States (United State Department of Education, 1995) and the diversity in cultural backgrounds, socio economic status and English language ability challenge theater educators to develop more appropriate, research-based approaches to educational assessment.

The assessment of potential talent in theater can have many educational benefits beyond casting students for a play. The creative and expressive abilities and energy of many students go unappreciated in school. Students who are more skilled in oral than written language, those who thrive in collaborative learning situations, and individuals who immerse themselves deeply in stories, characters, and emotions, or who are physically expressive, may have little opportunity to demonstrate their abilities in school and may perform poorly on academic tests

Dr. Barry Oreck is a performer and arts educator in New York City. He is a consultant for schools and school districts in the arts and gifted education and teaches at Long Island University and the University of Connecticut. Dr. Susan Baum, professor of Education at the College of New Rochelle is a researcher and international consultant on talent development. After 30 years as an educational psychologist, Steve Owen is now a biostatistician at the University of Texas Health Science Center at San Antonio.

(Oreck, Baum, & Owen, in press). Systematic assessment can increase awareness—on the part of teachers, parents, and peers—of students' artistic abilities and interests, and this can have a powerful effect on personal motivation, self-esteem, academic expectations, and educational choices for all students, not just those with outstanding talent. Observation of students engaged in theater activities provides unique information and a perspective on student potential not available through other forms of verbal or written assessment used in school.

The performing and visual arts have been part of the official U.S. Department of Education definition of gifted and talented students since 1972 (Marland, 1972) but few students are included in gifted programs based on their artistic talent (Haroutounian, 2002; Richert, 1992). The overwhelming reliance on standardized test and IQ scores in gifted identification reflects a widespread attitude in American education that artistic talents are separate and distinct from other areas of intellectual ability. The lack of valid, equitable, research-based assessment processes in the performing arts hampers the inclusion of the arts in the assessment of the gifted and talented and contributes to the marginalized status of arts education in the United States.

In 1994, ArtsConnection, a New York City arts-in-education organization, received a grant from the U.S. Department of Education Jacob Javits Gifted and Talented Students Program (ArtsConnection, 1996) to create and field test a process to assess potential talent in theater. The Javits Act was established to support the development of new means for identifying gifted and talented students from economically disadvantaged, special education and bilingual populations that are traditionally severely under-represented in gifted programs. The Theater Talent Assessment Process (TTAP) was developed and tested by theater educators in New York City and was subsequently expanded to Ohio (Ohio Department of Education, 2000) where the assessment of performing and visual arts talent has been mandated by the state legislature. In this paper we will trace the development of the process and report on the research results in both the pilot and expansion studies.

BACKGROUND

The inconsistency of basic theater instruction in the United States (National Center for Education Statistics, 1998) makes it particularly difficult to distinguish innate talent from the effects of prior experience and instruction. The reliance on text and language in theater can also magnify cultural and language differences among students, further hampering equitable and objective assessment of potential theater talent. Dramatic talent in young people is easily confused with verbal skill. Parts in the school play usually go to students who are good readers and have loud voices. In this regard, theater talent may often appear to be highly correlated with other verbal areas of school performance. There are many highly creative and expressive students, however, with great ability, interest and motivation to be in theater, who do not read well or speak English fluently.

Another challenge for assessment is that theater talent encompasses such a wide range of capabilities and skills that go beyond simply being heard and understood by an audience. Acting, directing, playwriting, as well as technical theater, costuming, media and many other specialties involved in theater production, all require different abilities and draw people with a variety of interests. Aspects of theater talent—such as the playwright's ability to create characters and tell a

story; the director's skill of listening, finding new meanings, and encouraging others' work; and the improviser's skill of collaboration and focus—are more difficult to assess and require time and instruction.

A valid, equitable, performance-based process for the identification of talent requires a broad definition of talent and the means to elicit both obvious and more subtle behaviors. To do their best work, students need a supportive environment that allows them to take risks and help each other. They need time to warm up and to allow their ideas and feelings to emerge. To equalize language differences, students must engage in both verbal and non-verbal activities. Students need multiple experiences to help overcome their shyness or unfamiliarity with the art form. All students must be observed in a variety of situations, groupings, and roles. These conditions for assessment present challenges, to be sure, but also mirror good theater instruction and the basic tenets of authentic, performance-based assessments (Consortium of National Arts Education Organizations, 1994; Linn & Baker, 1996).

When ArtsConnection began to develop a systematic assessment process for elementary school students in theater arts there were few written and tested criteria or processes on which to draw, from either professional or educational theater. The published instruments used to evaluate theater performance such as the Detroit Public Schools Creative Process Scale (Byrnes & Parke, 1982) the South Carolina Guidelines for the Identification of Artistically Talented Students (Elam & Doughty, 1988) or the Connecticut Conn•Cept IV Program (Saunders & Schmidt, 1979) are normally used to pre-screen or select students for special programs or magnet schools and are highly sensitive to specific previous training. That is, they measure achievement, not aptitude. This is a particular validity problem for assessing children who have had limited opportunities for formal training. Measures of general creativity such as the Torrance Tests of Creative Thinking (Torrance, 1966) or perceptual motor ability tests such as the Purdue Perceptual Motor Survey (Roach & Kephart, 1966), or Basic Motor Ability Test (Arnheim & Sinclair, 1974) are used occasionally as screening devices and may measure certain components of artistic ability but cannot be considered valid measures of potential theater talent (Karnes, 1983; Richert, 1992).

Some schools use the recommendations of classroom teachers, parents, or arts teachers to screen or identify students for advanced theater instruction or for gifted and talented programs. There are a number of research-based teacher recommendation instruments for gifted programs but only a few include the arts, most notably, the Scales for Rating the Behavioral Characteristics of Superior Students (SRBCSS) (Renzulli et al., 1976) and the Gifted and Talented Evaluation Scale (GATES) (Gilliam, Carpenter, & Christensen, 1996). Unfortunately, classroom teachers' knowledge of their students' dramatic or general artistic behaviors or activities in or outside of school is often limited. Even theater teachers may lack the training or adequate knowledge of the students to make such judgments. Teacher nomination can be valuable as a way of raising teachers' awareness of their students' talents and interests but these screening forms are not intended to be reliable assessments of talent on their own.

Despite the many challenges to the validity of talent assessment in the performing arts, theater education provides a wealth of models and structures for authentic, performance-based, curriculum-embedded assessment. The active participation of students in individual and group processes, the opportunities for students to express ideas and feelings, and the open-ended, imaginative, prob-

lem-solving activities involved in theater education are ideal vehicles for assessment. With careful planning, structure, training, and data collection, assessments can be developed in theater that are authentic to the discipline, psychometrically sound, aligned with the curriculum standards, and equitable to all students.

METHOD

The Theater Talent Assessment Process (TTAP) has undergone two phases of testing, the first in one school in New York City ($N = 224$) (1994–97), in which only theater talent was assessed (ArtsConnection, 1996) and then in two schools in Cleveland ($N = 218$) and Hamilton, Ohio ($N = 304$) in which TAP assessments in dance, music, and theater were administered to the same students (2001–03).

Subjects

After an extensive application and interview process, a public elementary school in Queens, New York, was selected as the site for the initial test of TTAP. The school's diverse student population—with a range of English language abilities, cultural backgrounds, and socio-economic levels—made it an ideal site to test the new process. Eight fourth grade classes participated in the process over two years (5% African American, 47% Hispanic, 28% White, and 20% Asian). 5% of fourth grade received special education services and 25% were categorized as Limited English Proficiency.

Schools in the expansion study in Ohio, Project Start ID (Ohio Alliance for Arts Education, 2003; Pearson, 2003), were also selected under the Javits grant guidelines favoring a student population with low-socioeconomic status and an under-identified gifted population. The schools were neighborhood, non-arts magnet schools with typical (and in some instances extreme) limitations in space, arts specialists, and schedule flexibility. Students were assessed in grades three and four with an ethnic composition in Hamilton of approximately 84% White, 13% African American and 3% Hispanic, and in Cleveland 91% African American, 1% Hispanic, 1% White, 1% Asian and 6% Other. Approximately 10% of students were designated as Special Education and all of the classes in a grade participated in the assessment process. Students from small, multi-grade, special education classes were assessed with other regular classrooms. Human Subjects approval was received for the research from the participating school districts and releases were obtained from all teachers and parents who participated. Students were not identified by name in any aspect of the research.

Research Questions

The initial research focused on four major areas: content validation of the talent definitions and behavioral descriptors; discriminant and convergent validity comparing the results to other academic and affective measures and indicators; interrater reliability among the assessment panel (two artists and the classroom teacher); and the effectiveness of the process in selecting students who would be successful over time in a rigorous instructional program. The specific research questions were:

1. Do the talent criteria developed for the process constitute a complete and coherent definition of theater talent in the area of performance?

2. Is the process equitable and independent of other measured variables? (i.e., Do students identified through the process reflect the school population in terms of gender and ethnicity? To what extent are scores from TTAP correlated with other measures of school performance?)
3. Are the results of the process reliable? (i.e., Do the observers agree with each other and are their observations corroborated by independent experts? Are classroom teachers able to reliably assess their students as the theater experts do? How many sessions are needed to arrive at a consensus between raters?)
4. Are students selected through the process successful in an advanced instructional program?

Overview of the Process

TTAP was designed to systematically assess the potential theater talent of all students, to identify students who are ready for advanced instruction, and to provide empirical data for the designation of students as gifted and talented. A larger goal of the process is to raise appreciation and understanding of the artistic abilities and potentials of all students on the part of teachers, parents, and the students themselves. The talent assessments in both the New York and Ohio programs were followed by once or twice weekly classes during and after school taught by some of the professional teaching artists who conducted the assessment. The format of the assessment was based on ArtsConnection's successful talent assessment processes in dance and music (DTAP, MTAP) (ArtsConnection, 1993; Baum, Owen, & Oreck, 1996; Kay & Subotnik, 1994) in which intact classrooms participate in a series of four classes developed and facilitated by two trained teaching artists. The two instructors alternate leading activities and marking their observation forms, insuring that at least one of them is observing at all times. Three assessors—two arts instructors and a classroom teacher or specialist—rate all of the students on a written checklist. Final student scores are computed from the total of all assessors for all sessions. The top students from each classroom are invited to a fifth "callback session." The process focuses on talents and skills in three major areas—acting, directing, and playwriting. Design, technical theater, and other aspects of talent are not specifically assessed. Originally designed and tested for grade four and five the process has subsequently been adapted and used from grades 3 to 12.

Definition of talent. The talent definitions and behavioral descriptors used in the assessment were developed by a group of theater educators representing a range of theatrical styles and backgrounds including improvisational theater, playwriting, acting, and classical drama.¹ The panel eventually narrowed the general criteria to four categories—focus, imagination, cooperation, and physical awareness. In each category, 7–10 specific behavioral descriptors defined specific behaviors reflective of high level performance in acting, directing and playwriting. These characteristics and their behavioral descriptors are naturally interrelated and usually seen in combination. The purpose of articulating the behaviors in separate categories is to help the observers focus on specific aspects of talent as they observe students working in a range of theater activities over the five sessions. The criteria and descriptors are listed in Figure 1.

<p>Physical Awareness Responds with whole body Is in control of body parts Uses and perceives vocal qualities Can use voice flexibly Wants to be heard and understood Is aware of space Notices details Observes carefully Seems relaxed Is unembarrassed</p> <p>Focus/Commitment Gives energy Takes risks Participates fully Perseveres Focuses eyes on the imagined environment and other players Recalls instructions Can revise and improve own work</p>	<p>Imagination Offers ideas Comes up with original or unusual suggestions Finds multiple solutions Makes the situation "real" Solves problems Sees the whole picture Invents dramatic situations Has a sense of effective timing</p> <p>Collaboration Works with others Responds to the audience Accepts the "rules" of the exercise Listens to teachers and peers Takes direction and criticism well Gives helpful suggestions Takes a leadership role</p>
---	--

Figure 1. Talent identification criteria in theater arts.

This definition of talent is grounded in Renzulli's (1978) Three Ring Conception of Giftedness. Giftedness, in Renzulli's conception, is defined as a behavior rather than a permanent state of being and results from the interaction of three traits—above average (not necessarily prodigious) ability, creativity, and task commitment. Gifted behaviors can emerge in different combinations in each individual under certain circumstances, at certain times. In the arts, these three aspects of talent are highly related, merging personal qualities and demonstrated motivation with specific skills. The theater arts panel overwhelmingly concurred with this three-part definition as it relates to the development and recognition of the talented theater artist.

Assessment facilitators. Because few elementary schools have theater teachers and TTAP requires two trained facilitators, the most feasible way for schools and districts to find artists to administer theater assessment is in partnership with arts councils and local arts-in-education organizations such as ArtsConnection and Young Audiences. Organizations such as these can provide the training, supervision and support to many schools in a region to appropriately administer the process. The use of an instructional team offers the potential for diversity in artistic and teaching styles and in viewpoints. Whenever possible, the instructional team consists of a male and a female, representing different cultures and artistic styles. The arts facilitators participate in a four-day training process in which they familiarize themselves with the criteria and assessment framework, develop their own five-session version of the assessment curriculum, and field test some of the activities with students. At the end of the training and after a successful full pilot administration, facilitators are certified to administer TTAP (Oreck, 2002). The training for artists and classroom teachers has been funded through the Javits grants and by the Ohio Arts Council and directed in New York and Ohio by the developers of the TTAP.

Training classroom teachers. An important element of both the New York and Ohio Javits grants was to train classroom teachers to assess artistic talents. Observing the classes and discussing the students with the teaching artists helps teach-

ers recognize both the behaviors indicative of potential talent and the conditions and stimuli that enable those behaviors to emerge. Training for classroom teachers begins with a two-hour workshop with the artists to familiarize them with the process, talent criteria, curriculum, and rating system. Training continues over the course of the five sessions during 10-minute post-class discussions during which time the students work quietly or are escorted to another class or the library. This discussion allows assessors to compare observations and opinions, clarify vocabulary, and discuss relevant insights and information about students. Every student is mentioned by name after each session. The discussion is an essential part of staff development for classroom teachers and helps to focus the observers on students they may have missed and borderline cases that require more attention.

Assessment process curriculum. The curriculum and class structure for the process were initially developed by Laura Livingston and Michael Durkin of Freestyle Repertory Theatre, a New York City improvisational theater company. Other teaching artists, trained in the assessment methods and following the TTAP framework, have adapted and altered the specific exercises used to fit their own teaching style and technique. The framework is designed to allow each of the talent criteria and every student to be assessed in each session. Both verbal and non-verbal activities are included in every session and a range of individual, partner, and small group activities are used. In dance and music, individuals' abilities can often be seen and heard during whole group activities; in theater the students must be heard individually. Thus the structure of the activities and the groupings of students must be carefully and efficiently planned.

Each assessment session lasts for approximately 35–40 minutes and consists of two main activities—a warm-up exercise to encourage spontaneity and get the students working together, and a structured improvisation with a primary focus on one of the major talent criteria. For example, Laura Livingston's first session focuses on collaboration and consists of Gift-giving (a game in pairs in which students give each other imaginary gifts) and Experts (a game conducted as an interview in which the respondent speaks on a topic in which s/he is imagined to be an "expert"). The behavioral criteria for the assessment of this session are listed in Figure 2.²

Livingston's second session emphasizes physical awareness, the third session focuses on both imagination and physical awareness (especially vocal), and the fourth session stresses on focus and commitment. The fifth "callback" brings together the top students in each classroom to attempt to level the playing field between classes and increase the challenge level.

The teaching artists play an active role in most of the activities to help maintain fairness and encourage shy or inhibited students to participate. Because of the collaborative nature of theater, a student's performance can be highly affected by his or her partner and the involvement of the teaching artist helps engage students at their own comfort level, equalizing the assessment conditions as fully as possible.

The greatest challenge for artists in the training process is the need to minimize verbal instruction and structure the time so that every student can be assessed in every activity. This is why only experienced teaching artists have been considered for training.

Instrumentation

The primary instrument used in TTAP is the Observation Tally Sheet, filled out individually by the assessors for each session. The Tally Sheet lists the four

Behavioral criteria for Session 1

1ST SESSION - COLLABORATION

<p>Warm Up: GIFT-GIVING</p> <p>Collaboration: Do students accept the size, shape and weight of the mimed object a partner offers? Do students remember to thank the partner? Are they aware that they must make offers in order for their partner to work with them? Are they helpful and supportive of their partner? Do they accept their partner's offers?</p> <p>Physical Awareness: Do they play with the mimed gifts as well as they talk about them?</p> <p>Focus/Commitment: Do they participate fully? Do they recall the instructions? Do they maintain their focus on their partner and avoid distractions of others?</p> <p>Imagination: Do they always add to the gift they have been given? Do they act as though the gift were real?</p>
<p>EXPERTS—INTERVIEW</p> <p>Collaboration: Do they answer questions freely? Do they listen and build on other people's answers?</p> <p>Focus/Commitment: Are they willing to get involved in spite of the strangeness of the first day? Are they able to focus their attention back to the instructor when necessary?</p> <p>Physical Awareness: Are they comfortable moving; can they answer questions nonverbally? Do they understand how to take the space they need while allowing others room?</p> <p>Imagination: Do they respond with unusual or unexpected answers?</p>

Figure 2. Behavioral criteria for session 1.

items in the talent definition in a box for each student in the class. A summary of observations from all assessors is transferred after each class to an individual Student Talent Profile sheet for each student. Figure 3 shows one row of one Observation Tally Sheet.

Students are identified by large name tags and the Observation Tally Sheets are organized alphabetically on two pages—one for boys and one for girls—so that notations can be made quickly and efficiently. Students are rated on a simple notice/not notice scale. When an assessor notices one of the listed behaviors, a plus mark (+) is placed next to the most relevant item or items in that student's box on the Tally Sheet. Negative marks are not counted and plus marks may not be erased. Only one plus mark per item is counted for each assessor during each activity. A student's total score for each item is based on the number of assessors who noted the item during the class. Thus, for the two major activities in the class the maximum score a student can receive for a single assessor would be eight (2 activities × 4 items), and the total for all three assessors would be 24.

Additionally, each assessor provides an overall holistic rating (1–5) for each student in each session. This overall score allows the observers to consider their

TTAP Theater Observation Tally Sheet

5 = Definite Candidate 4 = Probable 3 = Possible
 2 = Unlikely 1 = Not a candidate

Physical Awareness			+
Focus/Commitment	+	+	
Collaboration	+		
Imagination			
ANNA			
3			

Physical Awareness			
Focus/Commitment			
Collaboration		+	
Imagination			+
BARBARA			
2			

Physical Awareness	+	+	
Focus/Commitment	+	+	
Collaboration	+		
Imagination			+
CORY			
5			

Physical Awareness	+		
Focus/Commitment			
Collaboration			+
Imagination			+
DENISE			
3			

Figure 3. One row of one assessor’s observation tally sheet.

intuitive impressions concerning the student’s readiness for advanced instruction. The item totals alone cannot account for students who were particularly outstanding in one or two areas but who were less consistent across all four items. The overall ratings are combined with the item scores and divided by the number of sessions attended (minimum 2) for a final total average score. Final scores are standardized by class to account for the different teacher assessors and class-to-class differences. Both class and grade rankings are used to set cutoff scores for advanced instruction (typically 15–20%) and, in the Ohio process, for official gifted designation (3–5% of students).

Psychometric Evaluation

Reliability evidence for the TTAP included interrater estimates and week-to-week stability estimates. Content and construct validity procedures included: item review by experts in theater and gifted education; factor analysis; convergent and discriminant evidence; and a contrasting group’s comparison of talent ratings.

For discriminant validity evidence, student scores were collected on the Metropolitan Achievement Tests math (MAT-Math) and the Degrees of Reading Power (DRP) and, in the expansion study, the Ohio Proficiency Tests in reading and math. In addition, teacher questionnaires and comment forms concerning student behavior in the classroom and performance in other academic subject areas were collected for all identified students.

RESULTS

Research Question One: Definitions of Talent

Content validity evidence. Content validity evidence for the instruments was obtained during the development phase. The four items arrived at by the panel of seven theater educators were further reviewed by additional theater experts, school district coordinators for arts and gifted education, and a psychometrician.

Construct validity evidence: factor analysis. To study the dimensionality of the talent definition, exploratory principal factor analyses were run for the observational ratings (items summed across the five sessions). The ratings delivered a single factor that explained 97% of the ratings covariation (alpha estimate = .86). Table 1 shows the loadings of each item with its factor.

Each set of observational ratings formed homogeneous, unidimensional com-

Table 1. Factor Loading for Theater Criteria

Factor	Loading
Imagination	.89
Focus	.78
Physical Awareness	.76
Cooperation	.68

posites with high loadings for all items. The simple structure in the data and large amount of covariation explained shows that very little unsystematic variation exists in these observational data.

Research Question Two: Construct Validity

Convergent and divergent evidence. It was hypothesized that that performance in the theater assessment process would be moderately correlated with success in other areas of school performance. The verbal skills, creativity, and problem solving that students display in theater should contribute to success in the classroom and on standardized tests. Correlations between reading and math scores and TTAP results shown in Table 2 were small but significant. Reading and math scores for identified students in the initial study ranged from the 2nd to the 99th percentile with 65% of students reading in the bottom two quartiles and 58% below grade level in math. In the expansion study the percentage of students achieving grade level proficiency in reading was significantly higher for the selected group (54.9%) than for the not-selected students (29.5) ($\chi^2 = 7.64, p < .01$). Percentage differences in math performance between selected (66% non-proficient) and not selected students (79%) was not found to be significant ($\chi^2 = 3.07, p > .05$).

Before the first session of TTAP teachers filled out a form to identify which of their students they perceived to be talented in theater. The Talent Searchlist asks teachers to identify students with theater talent as well as the other art forms and other general behaviors indicative of talent (Baum, 1990). The Talent Searchlist results were significantly correlated with final selection for theater with an average correlation of .42 ($r^2 = .17$ representing the overlap in variance between TAP scores and Searchlist ratings). While statistically significant, less than a fourth of eventually identified students were recognized by their teachers as potentially talented prior to the assessment process. The significant correlations are presented in Table 2.

Table 2. Significant Correlations between TAP and Academic Test Scores and Teacher Searchlist—Two Year Average

	Pilot Study		Expansion Study	
	<i>r</i>	<i>r</i> ²	<i>r</i>	<i>r</i> ²
Math-NCE	.18	.03	.17	.03
Reading-NCE	.24	.06	.27	.07
Teacher Searchlist	.40	.16	—	—

Table 3. Interrater Reliability Estimates for Initial and Expansion Studies

		Initial Study		Expansion Study	
		A	B	A	B
Teaching Artist	A				
Teaching Artist	B	.545		.718	
Classroom Teacher	C	.737	.716	.652	.736

In the initial study between 3 and 8% of selected students received special education services either in self-contained classrooms or "resource room" settings. In the expansion study, the Hamilton school, with a relatively large special education population (8–10%) had a 2–5% participation rate and the Cleveland school with only a few labeled students per grade had no identified students receiving special education services during the two years of the program. Student report cards, and teacher reports in both projects confirmed the range of academic levels represented among the selected students. Differences in the percentage representation of ethnic groups and genders between the selected group and their grade levels were non-significant (chi-sq., $p > .05$) in any of the test schools although extreme differences in ethnic groups sizes in each site makes the analysis somewhat unreliable.

Research Question Three: Interrater Reliability and Stability

Interrater reliability. In the initial research, mean interrater reliability estimates (intraclass correlation) among the three assessors ranged from .55 to .74. Reliability coefficients between artists and among the artists and teachers improved each session, reaching a peak by session four. Interrater reliability estimates improved across all of the assessment panels in the expansion study shown in table 3.

Alpha reliability estimates also improved in each subsequent administration of the process by the same group of assessors and between the original and expansion studies from .65 in the initial study to .87 in the expansion study.

Stability estimates for the instrument were calculated over three separate one-session intervals. Stability estimates ranged from .35 to .55. Because each session was designed to highlight different aspects of theater talent, stability estimates were not expected to be high. These data show some consistency over time, suggesting generalized talent characteristics, but also variation based on the specific demands of each session. Stability estimates are shown in table 4.

Number of sessions needed. Experiments conducted before the pilot testing phase involved seven assessment sessions. Interrater reliability reached a peak by the fourth session and over 98% of students who were eventually selected for advanced instruction had been identified by that point. Based on these results, the process was shortened to four sessions with a fifth session callback. Reliability estimates for the expansion study supported the four session structure. Interim rankings after two and three sessions suggest that further shortening of the process would result in as much as 30% mis-ranking of students.

Table 4. Stability Results for Expansion Study

Session 1–2	.480
Session 2–3	.352
Session 3–4	.550

Further construct validity evidence: blind ratings. Additional validity evidence was gathered by collecting new ratings on students a year after the original assessment process. A random sample of Identified, Waitlist and Not Identified students participated in a new talent assessment. It was conducted and rated by professional artists unfamiliar with the students, using activities that none of the students had previously done. The random selection aimed for a 30% non-proportional sampling (i.e., equal sample sizes). A Hotelling T^2 was used to compare Identified and Not Identified students on all ratings simultaneously. Univariate t -tests were used as a *post hoc* probe of the significant T^2 (Bonferroni alpha = .006). The t -tests show that Identified students dependably received higher talent ratings ($T^2 = 32.60, p = .004$). The univariate t -tests favored Identified students for each rated behavior.

Research Question Four: Construct Validity Evidence

Informal supplementary data. One of the most significant markers of the success of an assessment approach is its ability to predict future performance in a rigorous instructional program. In both studies, identified students in grades four through six participated in twice weekly classes (approximately 135 minutes) in two levels taught by professional teaching artists. The classes consisted of improvisation, playwriting, and performing of scripts. During two years of advanced instruction in both studies, an average of 75% of identified students made good to excellent progress on written semi-annual evaluations by arts instructors using the original identification criteria. Although the training was rigorous, only 10% of the students left the program for reasons other than leaving the school. Additional evidence of identified students' readiness for advanced training was seen in the high attendance rate during after-school classes, the amount of home practice reported, scholarships received to attend weekend and summer classes, and instructors' reports of students' on-task behavior during theater classes.

DISCUSSION

These studies provide evidence that potential theater talent can be validly, reliably, and equitably identified in diverse student populations with little or no prior formal instruction. The diversity of students identified and their success over two years in demanding instruction supports the validity of the process and the definition of talent.

Definition of Talent

The four general behavioral categories—physical awareness, imagination, collaboration and focus—were found to be highly correlated with each other.

Such high correlation suggests that fewer criteria could be used. The goal of the process, however, is to raise awareness and appreciation, on the part of both expert and non-expert observers, for the specific skills and behaviors indicative of potential talent. The categories and behavioral descriptors help to focus the observers' attention on common characteristics and provide the vocabulary with which artists and classroom teachers can discuss students. High correlations between the individual item totals and the overall ratings also support the congruity and completeness of the definition.

Facilitation

Capable facilitation is the key to the success of the process. In order for students to demonstrate artistic potential they must engage in an authentic artistic experience. A safe environment supports activities that allow students to take risks, use their imaginations, and solve challenging, open-ended problems. When students are deeply engaged, behaviors indicative of potential talent are apparent to all observers, including those without arts expertise.

The teacher-student interaction in TTAP differs in many respects from most educational assessments in which the facilitator's instructions and feedback is highly limited or scripted. The goal here is to predict how students will do in a more demanding instructional program, so the assessors must be able to see how students respond to direct feedback, listen and take direction and use suggestions. While this kind of teacher-student interaction is essential to create authentic, realistic conditions for assessment, it also poses a threat to the validity of the process. Artists must make every effort to maintain fairness and give all students comparable challenges. For example, prompts should be changed frequently so that the last students called on during an activity are challenged as fully as the first. Some students need clearer instructions or slightly more time to respond than others, but the teacher must carefully balance the urge to instruct with the need for equity. Activities that require extensive instruction generally are too complex or not appropriate for assessment purposes. Alternately, if all students can accomplish the task (the goal in an instructional situation) the assessment will not differentiate between students and will yield little useful information.

The most successful TTAP facilitators have been highly experienced teaching artists who can readily adapt their work to the assessment framework and conduct the assessment in a wide range of classrooms and school settings.

Activities

The improvisational theater exercises used in the process minimize dependence on reading skills but present other challenges to equitable assessment. Quick responses, humor, and verbal acuity can certainly be indicators of potential theater talent but can also overshadow less obvious but equally relevant characteristics. Some students need more time to think and react but offer thoughtful, meaningful contributions. It is important that the facilitators arrange the activities and manage the time to allow for a range of responses and that they encourage and support students who are less verbally facile, outgoing, and aggressive. The richer and more engaging the activities the easier it is to see all students at their best.

Scoring

It is essential that the process of assessment not undermine the students' artistic experience. Using the simple plus mark allows assessors to fill out their score sheets efficiently while maintaining focus on the students as the activities unfold. The two-part system includes both specific notices and an overall intuitive judgment of student potential and depends on the cumulative judgment of three people over the course of the entire process. This approach values the strength and consistency of the students' performance over time, in a wide range of activities that reflect actual practices in theater classes.

Reliability Results

Interrater reliability results improved in each subsequent administration of the process with the same artist/teacher teams. This is likely due both to improved facilitation skills and more effective facilitation on the part of the artists. The high level of agreement between the artists and the classroom teacher supports one of the central goals of TTAP—to increase teachers' awareness of and appreciation for the artistic abilities of their students.

It can be argued that the post-class discussion threatens the independence of the assessors' scores in subsequent sessions. The goal of the process, however, is to complete a thorough assessment for every student in the class; the discussion immediately alerts the assessors to students they have missed. The strongest and weakest students in the class are easy to spot. It is those in the middle, students who are less outgoing, or those who need a little extra time to understand instructions, who are most likely to be overlooked. Further, sharing this information helps assessors disclose their own biases and preferences by requiring them to articulate the specific behaviors they have noticed and to hear the observations of others. We consider incomplete or biased assessments to be stronger validity threats than the problem of assessor independence. Further, the discussion is an important component in training classroom teachers to understand the criteria, vocabulary, and approaches used by the artists, and it allows the teachers to share with the artists potentially relevant background information about students. Teachers often supply valuable information about students who may have particular language or learning difficulties, distractions or problems with other students, or who have been affected by an event outside of class.

Use of Assessment Information

Ultimately, the value of talent assessment and identification depends on the use of the information. New methods for identifying gifted behaviors, even if they are more valid and equitable than traditional methods, can have little impact on students if not supported by appropriate opportunities for talent development. One positive effect of talent assessment is to increase the motivation of parents and schools to provide both introductory and advanced levels of arts instruction. The knowledge that one's child or student is potentially talented can be a powerful incentive for parents and teachers to advocate for the arts in school, find classes outside of school, and make sacrifices to support the student's arts involvement.

For many students TTAP is their first formal theater instruction so it is especially important that it be a positive experience. While participation in TTAP can introduce students to content contained in state or national curriculum standards, the primary goal is talent assessment. Four classes (or five for callback students) cannot be considered adequate basic theater instruction. TTAP is designed to be used as part of an overall theater education program.

Classroom teachers stand to gain valuable perspectives on all of their students through TTAP. Regardless of their previous experience with theater, classroom teachers are almost universally positive about the process and are enthusiastic, insightful observers. TTAP provides an unusual opportunity for teachers to watch their students engaged in creative activities and to talk about the individuals in their class. Studies conducted through the Javits projects demonstrated that the talent assessment process can raise teachers' expectations of students, motivate changes in teaching strategies, and help students apply their talents to classroom work (ArtsConnection, 1996; Baum, Owen & Oreck, 1997; Oreck, 2004). These findings demonstrate the potential of artistic talent identification to motivate changes in teaching and learning and to raise appreciation for artistic abilities and use of artistic methods in the curriculum.

Limitations of the Studies

While the number of students in the sample is adequate for statistical purposes, having just three schools in the project makes generalizability of these data limited. The schools were chosen to be representative of their geographic areas in terms of size, student population, existing arts resources, and test scores but further research in a wider variety of schools will be needed to corroborate these results.

Improvisational theater games offer students opportunities to not only perform as actors but also to create stories and to work collaboratively as directors. The theater educators involved in developing this process hypothesize that characteristics shared by acting, directing, and playwriting are revealed through improvisation. To what extent these three areas of talent are assessed equally, however, is unknown. A study involving more varied opportunities for advanced instruction would need to be offered and students followed over a longer period of time to fully investigate the extent to which these talent areas overlap in the same individuals and if the process favors certain abilities over others. Further adaptations of the process could also pursue other areas of theater talent such as design, media, costumes, lighting, sound, etc.

The Future of TTAP

TTAP has continued to be used in its original pilot site since 1994. In addition, the process has been adapted for use in other elementary and junior high schools in New York City and Cleveland, Ohio and as a selection process for a magnet arts high school in Mississippi. In both New York and Ohio, the talent assessments were part of a larger program that provided advanced and introductory instruction to students, professional development for teachers, workshops for parents, and academic tutoring for students. Such comprehensive programming creates the ideal setting in which to conduct talent assessment. The instruments and materials package are available through ArtsConnection. Dr. Barry Oreck, developer of the process, along with an experienced TTAP facilitator, conducts the four-day training of artists.

The biggest challenge for schools and school districts in implementing TTAP is the need to recruit and train artist facilitators and plan for advanced instruction for identified students. Given current financial realities, schools and districts are unlikely to begin such a process without evidence that it will contribute to overall school improvement and the achievement of high standards in both the arts and other academic areas.

The action of the Ohio state legislature in revising the identification criteria for gifted and talented programs was taken, in part, in response to concerns about the lack of diversity and equity in gifted and talented programs and the need to become more systematic in all means of identification. Many other states, including Georgia, North Carolina, Kentucky, New York, and Oklahoma, have grappled with these complex issues in recent years and we hope that the existence of research-based assessment processes in the arts will encourage inclusion of the arts in gifted and talented identification in other states.

CONCLUSION

Theater programs that select students based exclusively on the traditional audition format may miss behaviors that emerge over time and may prematurely exclude many students with no prior experience in theater or limited English language proficiency. While not all students may be ready for rigorous advanced instruction, some talented behaviors emerge in virtually every student over the course of the structured assessment process. As a result, students can become aware of their own strengths and interests, teachers can observe and tap into the creative and artistic abilities of their students, and parents can become more motivated to pursue arts instruction for their children.

In the current educational climate of high stakes testing and accountability, arts educators need systematic, high-quality, research-based approaches to justify the essential place of the arts in the curriculum. Likewise gifted programs, long considered elitist because they rely solely on standardized test and IQ scores to identify students for special services, need to broaden their criteria and identification processes to increase diversity and equity and to recognize abilities other than those measured on written tests. Research on the nature of intelligence and talent highlight the potential of creative and artistic abilities to be the catalyst for success in school and beyond (Gardner, 1983; Renzulli, 1995; Sternberg, 1988). This project demonstrates a theater education program's potential to be such a catalyst.

NOTES

¹The working group involved in the creation of the talent criteria included Laura Livingston, Michael Durkin, Freestyle Repertory Theatre; Abigail Adams, People's Light and Theater Company; Maxine Maxwell, Theater Works USA; Terry Griese, Irondale Ensemble Theater; David Shookoff, Manhattan Theater Club; Bruce Taylor, Metropolitan Opera Guild; and Jane Remer, Arts Education consultant.

²The full five session plan is available as part of the TTAP materials package from ArtsConnection, 520 8th Avenue, Suite 321 New York, NY 10018 or oreckb@artsconnection.org.

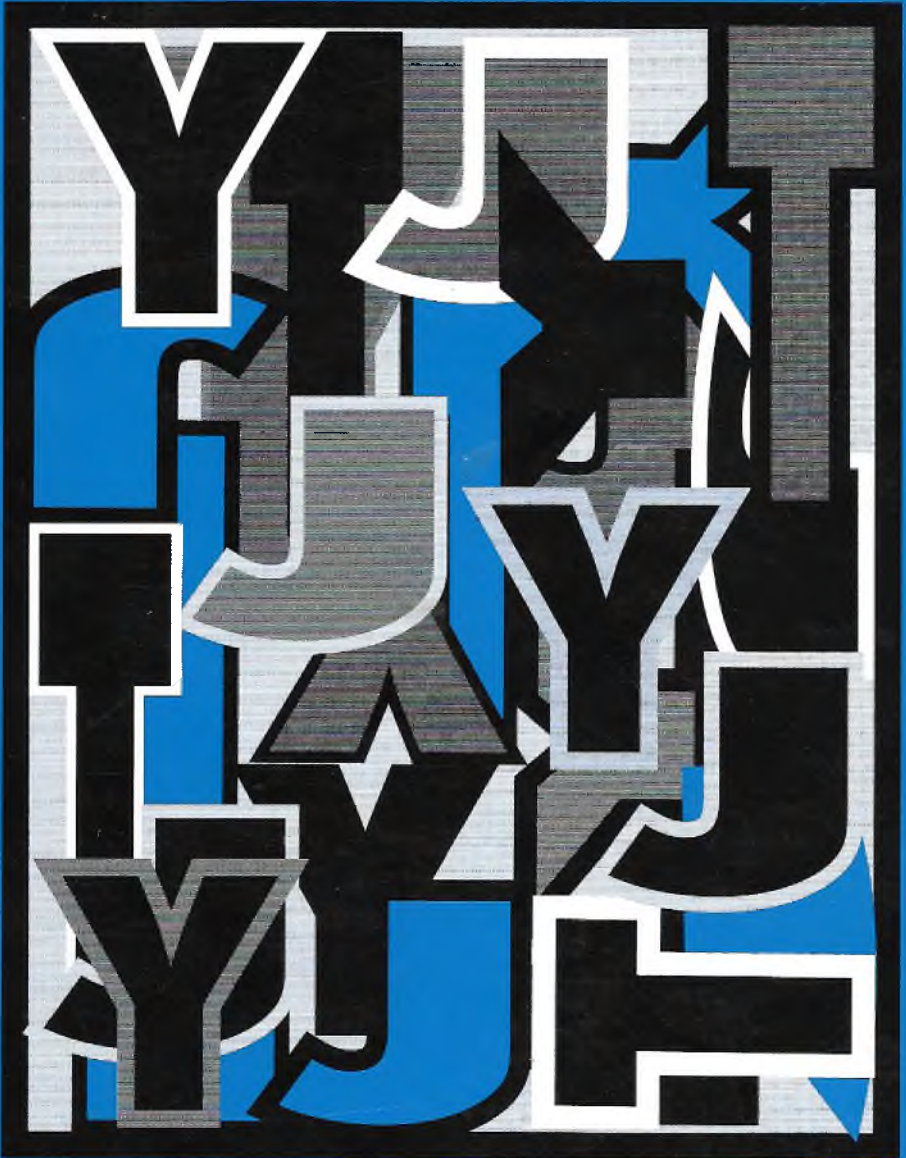
WORKS CITED

- Arnheim, D., & Sinclair, W. (1974). *Basic motor ability tests*. Long Beach, CA: Institute of Sensory Motor Development, California State University.
- ArtsConnection. (1993). *Talent beyond words: Report to the Jacob Javits gifted and talented students education program*. Washington, DC: United States Department of Education, Office of Education Research and Improvement, #R206A300148.
- ArtsConnection. (1996). *New horizons: Report to the Jacob Javits gifted and talented students education program*. Washington, DC: United States Department of Education, Office of Education Research and Improvement, #R206A30046.
- Baum, S. M. (1990). Recognizing talent in young children. In S. Baum, S. Reis, L. Maxfield (Eds.), *Nurturing the gifts and talents of primary grade students* (pp. 30–40). Storrs, CT: Creative Learning Press.
- Baum, S., Owen, S., & Oreck, B. (1996). Talent beyond words: Identification of potential talent in dance and music in elementary students. *Gifted Child Quarterly*, 40, 93–101.
- Baum, S., Owen, S., & Oreck, B. (1997). Transferring individual self-regulation processes from arts to academics. *Arts Education Policy Review*, 98, 32–39.
- Byrnes, P., & Parke, B. (1982). *Creative products scale: Detroit Public Schools*. Paper presented at the Annual International Convention of the Council for Exceptional Children, Baltimore, MD.
- Consortium of National Arts Education Organizations. (1994). *National standards for arts education*. Reston, VA: Music Educators National Conference.
- Elam, A., & Doughty, R. (1988). *Guidelines for the identification of artistically gifted and talented students* (Rev. ed.). Columbia, SC: South Carolina State Department of Education.
- Gardner, H. (1983). *Frames of mind*. New York: Basic Books.
- Gilliam, J., Carpenter, B., & Christensen, J. (1996). *Gifted and talented evaluation scales*. Austin, TX: Pro-Ed.
- Haroutounian, J. (2002). *Kindling the spark: Recognizing and developing musical talent*. New York: Oxford.
- Karnes, Merle B. (Ed). (1983). *The underserved: our young gifted children*. ERIC Clearinghouse, Council for Exceptional Children.
- Kay, S. I., & Subotnik, R. F. (1994). Talent beyond words: Unveiling spatial, expressive, kinesthetic and musical talent in young children. *Gifted Child Quarterly*, 38, 70–74.
- Linn, R. L., & Baker, E. L. (1996). Can performance-based student assessments be psychometrically sound? In J. B. Baron & D. P. Wolf (Eds.), *Performance-based student assessment: Challenges and possibilities* (95th Yearbook of the National Society of Education, part 1). Chicago: University of Chicago Press.
- Marland, S. (1972). *Education of the gifted and talented: Report to the congress of the United States by the commissioner of education*. Washington, DC: U.S. Office of Education.
- National Center for Education Statistics. (1998). *The NAEP 1997 arts report card*. U.S. Department of Education, NCES 1999-486.
- Ohio Alliance for Arts Education. (2003). *Project Start ID* (Interim report to the Jacob Javits gifted and talented students education program of the United States Department of Education, Office of Education Research and Improvement). Columbus: OH: Author.
- Ohio Department of Education. (2000). *Project START ID: Statewide arts talent identification and development project*. Columbus: Author.
- Oreck, B. A. (2002). *D/M/T TAP training manual*. New York: ArtsConnection.
- Oreck, B. A. (2004). The artistic and professional development of teachers: A study of teachers' attitudes toward and use of the arts in teaching. *Journal of Teacher Education*, 55(1), 55–69.
- Oreck, B., Baum, S., & Owen, S. (in press). Validity, reliability and equity issues in an Observational talent assessment process in the performing arts. *Journal for the Education of the Gifted* 27 (2).
- Pearson, L. (2003). Identifying and serving our gifted students: How professional development can help meet the challenges. *TRIAD: Journal of the Ohio Music Education Association*, 70(6), 37–41.

- Renzulli, J. (1978). What makes giftedness? Reexamining a definition. *Phi Delta Kappan*, 60, 180-184.
- Renzulli, J. S. (1995). *Schools for talent development*. Mansfield Center, CT: Creative Learning Press.
- Renzulli, J. S., Smith, L. H., White, A. J., Callahan, C. M., Hartman, R. K., & Westberg, K. L. (1976). *Scales for rating the behavioral characteristics of superior students*. Mansfield Center, CT: Creative Learning Press.
- Richert, S. E. (1992). *Equitable identification of students with gifted potential*. ERIC Document Reproduction Service, No ED366159.
- Roach, E., & Kephart, M. (1966). *Purdue perceptual motor survey*. Columbus, OH: Charles E. Merrill.
- Saunders, R. J., & Schmidt, L. (Eds.). (March 1979). *Con-Cept IV: Task force report on identifying the talented in the creative arts*. Hartford, CT: Connecticut State Department of Education.
- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Penguin.
- Torrance, E. (1966). *Torrance tests of creative thinking*. New York: Scholastic Testing.
- United States Department of Education. (1995). *Arts education in public elementary and secondary schools*. Washington, DC: National Center for Education Statistics. (Office of Educational Research and Improvement No. NCEES 95-082).

YTP YOUTH THEATRE JOURNAL

VOLUME EIGHTEEN ▲ 2004



PUBLISHED BY
AMERICAN ALLIANCE FOR THEATRE AND EDUCATION