

University of Arkansas, Fayetteville

ScholarWorks@UARK

Education Reform Faculty and Graduate
Students Publications

Education Reform

8-31-2017

The Play's the Thing: Experimentally Examining the Social and Cognitive Effects of School Field Trips to Live Theater Performances

Jay P. Greene

University of Arkansas, Fayetteville, jpg@uark.edu

Heidi Holmes Erickson

University of Arkansas

Angela R. Watson

University of Arkansas, Fayetteville

Molly I. Beck

University of Arkansas, Fayetteville

Follow this and additional works at: <https://scholarworks.uark.edu/edrepub>



Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Educational Leadership Commons](#), and the [Other Educational Administration and Supervision Commons](#)

Citation

Greene, J. P., Holmes Erickson, H., Watson, A. R., & Beck, M. I. (2017). The Play's the Thing: Experimentally Examining the Social and Cognitive Effects of School Field Trips to Live Theater Performances. *Education Reform Faculty and Graduate Students Publications*. Retrieved from <https://scholarworks.uark.edu/edrepub/7>

This Article is brought to you for free and open access by the Education Reform at ScholarWorks@UARK. It has been accepted for inclusion in Education Reform Faculty and Graduate Students Publications by an authorized administrator of ScholarWorks@UARK. For more information, please contact scholar@uark.edu.



UNIVERSITY OF
ARKANSAS

College of Education & Health Professions
Education Reform

WORKING PAPER SERIES

The Play's the Thing: Experimentally Examining the Social and Cognitive Effects of School Field Trips to Live Theater Performances

Jay P. Greene

Heidi H. Erickson

Angela R. Watson

Molly I. Beck

August 31, 2017

EDRE Working Paper 2017-13

The University of Arkansas, Department of Education Reform (EDRE) working paper series is intended to widely disseminate and make easily accessible the results of EDRE faculty and students' latest findings. The Working Papers in this series have not undergone peer review or been edited by the University of Arkansas. The working papers are widely available, to encourage discussion and input from the research community before publication in a formal, peer reviewed journal. Unless otherwise indicated, working papers can be cited without permission of the author so long as the source is clearly referred to as an EDRE working paper.

The Play's the Thing:
Experimentally Examining the Social and Cognitive Effects of School Field Trips to Live
Theater Performances

Authors

Jay P. Greene, University of Arkansas, Department of Education Reform,
201 Graduate Education Building, Fayetteville, AR, 72701
jpg@uark.edu, Fax, (479)575-3196, Phone (479)575-3172

Corresponding author

Heidi Holmes Erickson, hh018@uakr.edu

Angela R. Watson, arw007@uark.edu

Molly I. Beck, mibeck@uark.edu

Abstract:

Field trips to see theater performances are a long-standing educational practice, however, there is little systematic evidence demonstrating educational benefits. This article describes the results of five random assignment experiments spanning two years where school groups were assigned by lottery to attend a live theater performance, or for some groups, watch a movie-version of the same story. We find significant educational benefits from seeing live theater, including higher levels of tolerance, social perspective taking, and stronger command of the plot and vocabulary of those plays. Students randomly assigned to watch a movie did not experience these benefits. Our findings also suggest that theater field trips may cultivate the desire among students to frequent the theater in the future.

Keywords: arts education, drama based education, experimental design, informal learning, school field trips

The Play's the Thing:
Experimentally Examining the Social and Cognitive Effects of School Field Trips to Live
Theater Performances

Introduction

Student field trips to see live theater performances are a long-standing educational practice. Like many common school practices, however, there is little systematic evidence demonstrating educational benefits. Field trips to see plays continue mostly with the support of the wisdom of educators and a sensible deference to prior practice. With the rise of test-based accountability, however, many traditional school practices are under pressure, including school field trips to theaters and other cultural institutions (Rabkin & Hedberg, 2011; Gadsden, 2008). Cultural field trips face severe cutbacks if they cannot show improved performance on tested subjects or other important educational benefits.

This article describes the results of a series of five experiments conducted over two years in which school groups, from a variety of grades, were assigned by lottery to see live theater. We find significant educational benefits from taking students to a theater performance. In particular, students randomly assigned to see live theater demonstrate higher levels of tolerance and social perspective taking as well as stronger command of the plot and vocabulary of those plays. Our findings also suggest that theater field trips may cultivate the desire among students to frequent the theater in the future.

In addition, for two of the five experiments, a second treatment condition was added in which some students were randomly assigned to see a movie comparable to the play seen by other groups of students. Leaving school to see a movie did not produce the same benefits as viewing live theater. The evidence suggests that there are educational benefits to the traditional

practice of school field trips to see plays, and that those benefits are unlikely to be replicated by showing students movies instead.

Previous Research

While there has been little rigorous research that speaks directly to the effects of seeing live theater on students, there is a growing literature on related topics. For example, a recent, large-scale experiment found that a single school field trip to tour an art museum caused significant effects that could be observed nearly two months following the visit (Greene, Kisida, & Bowen, 2014). Students randomly assigned to receive the art museum tour were significantly more likely to be interested in visiting cultural institutions in the future and actually did so at a higher rate than students randomly assigned to the control group that had not toured the museum (Kisida, Greene, & Bowen, 2014). Students who toured the art museum also scored significantly higher on a measure of their critical thinking skills (Bowen, Greene, & Kisida, 2014). In addition, students who visited the art museum displayed higher levels of content knowledge, tolerance, and historical empathy as a result of their tour (Greene, Kisida, & Bowen, 2014). All of these benefits were more likely to be realized by students from more disadvantaged backgrounds, suggesting that advantaged parents can more easily substitute with their private efforts if schools fail to take students to cultural institutions. It appears that schools may play an essential role in providing equal access to cultural institutions and any benefits they produce.

There has also been some long-term analyses finding that exposure to cultural activities improves academic outcomes for students years afterwards. For example, Jægar and Møllegaard (2017) studied a large sample of monozygotic twins in Denmark to see if their cultural activity was related to later educational outcomes. By comparing outcomes among identical twins, the researchers control automatically for a large set of unobserved environmental and genetic

factors. The mothers of these sets of identical twins were asked about each child's cultural activity at the age of 12, including "How often child went to any type of museum" and "How often child went to the theater or a musical performance." The researchers then examined how well each child did according to teacher and independently proctored grades at the age of 15 as well as the rate at which they graduated from high school. The more cultural activity children engaged in, the higher were their independently proctored grades and the more likely they were to graduate from high school. Other long term studies find correlations between student involvement in the arts and later academic outcomes (Ruppert, 2006; Lacoë, Painter, & Williams, 2016), but most of these studies have difficulty establishing that this is a causal relationship.

Goldstein and Winner (2012) conducted a set of experiments to examine how students are affected by drama activities and find significant social-emotional benefits measured shortly after the intervention. There is also an extensive literature that examines how drama-based instructional techniques affect students. In a meta-analysis of that research, Lee, Patall, and Cawthon (2015) found that drama-based pedagogy can have significant academic effects as well as social-emotional outcomes. While this is a comprehensive review of that research, the authors acknowledged that the forty-seven studies they examined are quasi-experimental, not experimental, therefore some of the observed relationships may not be causal.

We have some evidence to suggest that students benefit from school visits to art museums, experience long-term academic gains from frequenting museums and the theater, and may learn from drama-based pedagogy and theater activities. But regarding the exact question addressed in this study – whether students benefit from school visits to see live theater – there is little direct evidence. The results of two of the five theater experiments contained in this article

were described in an earlier publication (Greene, Hitt, Kraybill, & Bogulski, 2015), but to our knowledge this is the first large-scale experiment to examine what students learn from seeing live theater.

Research Design

This study addresses the question of whether students benefit from school visits to see live theater. School groups were randomly chosen to receive free tickets to attend five live theater performances over a two year period.

Teachers applied for a chance to bring their class to one of five different theater performances. We then matched similar classes based on student demographics, typically this match occurred within schools. Matching occurred prior to randomization because students within the same or neighboring schools are thought to share similar observed and unobserved characteristics. Within the matched set, we randomly assigned one or more classes to receive tickets and one to serve as a control group, ensuring the treatment and control groups were similar. In total we conducted forty-seven lotteries, creating ninety-four treatment and control groups containing almost 1,500 students. Performances included *A Christmas Carol*, *Hamlet*, *Around the World in 80 Days*, and *Peter and the Starcatcher*, all performed by an award winning professional company, and *Twelfth Night* performed by university theater students.

It is important to note that these interventions did not include anything beyond the opportunity to see live theater. The treatment did not provide any additional training, materials for teachers, or supplemental activities for students. Any supplementary activities, such as reading the play or watching the movie in class, were assigned at the teachers' discretion and

could have occurred among treatment or control groups. This study is designed to identify the effects of seeing a play independent of any other school activities.

For the final two plays we were able to add a second treatment condition in which students would be randomly assigned to see a movie that was similar to what the theater treatment group was seeing. For the *Twelfth Night* experiment, applicant groups were randomly assigned to see the play, to see the 1996 film of the same Shakespeare story, or to serve in the control group which saw neither the play nor the movie. For the *Peter and the Starcatcher* experiment, applicant groups were randomly assigned to see the play, to see the 1991 film *Hook*, or to serve in the control group that saw neither. The addition of a movie treatment allowed us to test whether any effects of seeing a play were derived from the subject matter of the play or from the experience of seeing live theater. The fact that both play and movie treatment groups left school on the same bus, at the same time, and only differed in whether they walked into the University's Union to see the movie or into the theater to see the play allows us to examine whether any observed outcomes could be caused simply by leaving school on a field trip.

To collect outcome measures, we administered surveys to treatment and control students in their classrooms. On average, surveys were administered fifty-four days after the treatment group had seen the play. There were not differential participation rates among the treatment and control groups. We collected surveys from 77.6% of the students assigned to see a play, 76.0% for those assigned to see a movie, and 76.5% among control group students.

While the basic design— offering free tickets, matching similar applicant groups, and then conducting a lottery within matched sets of applicants—remained the same across all five plays, some important details did change over time. For example, after the first two plays, we changed the survey to add a scale designed to measure students' Social Perspective Taking (SPT)

(Gehlbach et al. 2008; Gehlbach, 2004) while dropping the Reading the Mind in the Eyes Test (Baron-Cohen et al. 2001). We believed that SPT would be a better way of capturing potential social-emotional effects of seeing live theater.

In addition to replacing one measure on the survey instrument after the first two plays, we also administered pre-treatment surveys to students after the first three plays. In prior administrations we lacked the resources to collect measures both before and after the intervention, so we relied on the lottery to give us equivalent treatment and control groups since we could not control for any pre-treatment differences. For the final two plays, however, we were able to administer surveys to all students both before and after the treatment occurred. This allowed us to check whether our treatment and control groups were similar on pre-treatment measures of the outcomes. Controlling for pre-treatment measures of the outcomes also improves the precision of our estimates of treatment effects.

Other outcomes measured in the survey remained unchanged across the five plays. We used a scale to measure Tolerance (Greene, Kisida, & Bowen, 2014). We suspected that by exposing students to a broader world through theater, they would increase their ability to understand other people's points of view (SPT) as well as gain greater acceptance of other people (Tolerance). We also suspected that seeing live theater would be an effective mechanism for conveying the plot and vocabulary of these plays, so we included measures of Content Knowledge. In addition, since past research suggested that visiting cultural institutions increased the desire to frequent those institutions in the future, we included measures of Theater Consumption in the survey. Lastly, we included measures of the desire to Participate in Theater because we thought it was possible that seeing live theater might inspire students to become more involved in theater by auditioning for plays, taking drama class, etc.

Implementation of Research Design

Examining the background characteristics of our treatment and control groups confirms that randomization was successful in helping ensure that we compared generally similar groups. There are no differences in background characteristics that are significant at $p < .05$ (see Table 1). There are two instances in which differences are significant at $p < .10$, but we might expect that by chance given that we are comparing three groups on ten different variables. On our pre-treatment measures of outcomes we observed no statistically significant differences between the control and treatment groups.

Students in our sample were just shy of their 15th birthday and in the middle of 9th grade on average. However, students varied in age, with some students as young as 4th grade and some as old as 12th grade. Approximately two-thirds of our sample identified as white, which reflects the broader community in which the experiments took place. About one-quarter of students had seen a play in the previous year, however this might have included school plays, church plays, and holiday shows. This probably reflects a relatively low level of previous exposure to theater and cultural activity, but it may be higher than the broader community given that some of the applicant groups were drama classes.

Two problems occurred during implementation of the research design that caused some applicant groups not to see the play despite being assigned to the play treatment. Severe winter weather forced the cancellation of a performance of *A Christmas Carol*. Additionally, during the *Around the World in 80 Days* experiment the theater made an error that caused the actors not to be available to perform when students arrived. To be very conservative, we report results for “intention to treat,” in which we count all students randomly assigned to the treatment group as if they received the treatment even if they failed to do so because of weather or a scheduling error.

To provide a more realistic estimate of the treatment effect, we also report the “impact on treated” derived from a two-stage model in which the first stage uses assignment to the treatment group as a predictor of whether students actually received the treatment.

Outcome Measures

The Tolerance scale consisted of seven items. Students were given four options, from strongly disagree to strongly agree, to respond to a series of statements designed to capture their general acceptance of other people and different opinions. The scale was adapted from Greene, Kisida, and Bowen (2014) and included statements, such as “People who disagree with my point of view bother me” or “I think people can have different opinions about the same thing.” The Cronbach’s alpha for this scale was .71, suggesting that there is an acceptable amount of internal consistency within this scale.

The Social Perspective Taking scale also consisted of seven items and was adapted from Gehlbach, Brinkworth, and Wang (2012). Students had five response options, from almost never to almost all of the time, to questions like “How often do you try to figure out what motivates others to behave as they do?” and “Overall, how often do you try to understand the point of view of other people?” The Cronbach’s alpha in our study for this scale was .85, indicating strong internal consistency in students’ responses.

Our measure of Content Knowledge consisted of six questions about the plot of each play and five questions about vocabulary drawn from the play. For example, for students in the *Hamlet* experiment we asked “What happens to Ophelia?” or asked them about the definition of “countenance” and provided four response options to each question. Given that the questions were different for each play, calculating a combined Cronbach’s alpha is not possible, but we are

confident that we captured meaningful variation in knowledge about each play's plot and vocabulary.

The Theater Consumption scale consisted of ten items adapted from Kisida, Greene, and Bowen (2014). Students had four response options to questions like "How interested are you in seeing live performances in a theater?" or statements like "I plan to see live theater performances when I am an adult." The Cronbach's alpha for this scale was .92.

The Theater Participation scale was also adapted from Greene, Kisida, and Bowen (2014) and consisted of four items. Students were asked questions like "How interested are you in taking a drama class?" and "If your school were having auditions for a new play, how interested would you be in trying to get a role in that play?" The Cronbach's alpha for this scale was .93.

In Table 1 all of these outcomes are expressed as the mean of a 0 to 3 or 0 to 4 scale, except for Content Knowledge, which is expressed as the percentage of questions answered correctly. For the purposes of the outcome analyses all scales were converted into z-scores with a mean of 0 and a standard deviation of 1. The reported results, therefore, are the effect sizes expressed as a percentage of a standard deviation.

(insert table 1 about here)

Analyses

Because the randomized controlled trial research design used here has the important feature of generating comparable treatment and control groups (which we confirmed with the analyses presented in Table 1), we can use a straightforward set of analytic techniques, designed for use in social experiments, to estimate the impact of a school field trip to see live theater on student

outcomes. In its simplest form, this technique can estimate mean differences using the following equation for outcome γ of student i in matched set m :

$$(1) \quad \gamma_{im} = \alpha + \beta_1 Play_{im} + \beta_2 Movie_{im} + \beta_3 Match_m + \varepsilon_{im}$$

where the binary variable $Play_{im}$ is equal to 1 if the student is in the treatment group that was randomly assigned to receive free tickets for a field trip to see one of the five plays and is equal to 0 otherwise. The binary variable $Movie_{im}$ is equal to 1 if the student is in the treatment group that was randomly assigned to take a field trip to see a movie and is equal to 0 otherwise.

Because the groups were created using a stratified randomization procedure within matched applicant sets, $Match_m$ is also included in the model as a vector of binary variables that have the statistical effect of estimating within, as opposed to across, matched groupings. Finally, ε_{im} is a stochastic error term clustered at the applicant group level to take into account the spatial correlation from students nested within applicant groups.

Proper randomization generates experimental groups that are comparable but not necessarily identical. The basic regression model can, therefore, be improved by adding controls for observable characteristics to increase the reliability of the estimated impact by accounting for minor differences and improving the precision of the overall statistical model. This yields the following equation:

$$(2) \quad \gamma_{im} = \alpha + \beta_1 Play_{im} + \beta_2 Movie_{im} + \beta_3 Match_m + \beta_4 Gender_{im} + \beta_5 Age_{im} + \beta_6 White_{im} + \varepsilon_{im}$$

where $Gender_{im}$ is a binary variable equal to 1 if the student is a female and 0 otherwise, Age_{im} is a continuous variable indicating precise age of student i at the time post-treatment surveys were administered, and $White_{im}$ is a binary variable equal to 0 if the student does not identify as

being white and is 1 otherwise. In this model, β_1 and β_2 are the parameters of interest and represents the effect of a class field trips for students in the play and movie treatment groups. Equation (2) is our preferred model and was used to produce the Intention to Treat results presented in Table 2.

Due to the non-compliance to treatment assignment during the first and third play experiments caused by weather and a scheduling error on the part of the theater, we are also interested in generating an Impact on Treated estimate. The Impact on Treated estimate describes what the effect would have been had all of the applicant groups actually seen the plays to which they were randomly assigned. The model used to generate that estimate is a two-stage least squares model in which the second stage is identical to Equation (2) except that $Play_{im}$ is derived from the first stage in which lottery assignment is used to predict treatment compliance.

Because we do not have strong theoretical expectations that different plays should produce different effects and because the sample size for each individual play is relatively small, we present in Tables 2 and 3 the results of all five plays combined. The outcomes for Social Perspective Taking in Table 2 only contain the results for students who saw the last three plays since SPT was not added to the survey instrument until that time. All of the other outcomes in Table 2 represent the results of students across all five plays.

For the last two plays we surveyed all students prior to the treatment and again after the intervention, which allows us to control for the pre-treatment measure of the outcome. For example, we can control for students' score on the Tolerance scale prior to the intervention when estimating the effect of the treatment on their Tolerance score collected after the intervention. The model we used to generate these results (as presented in Table 3) can be expressed as:

$$(3) \quad \gamma_{im} = \alpha + \beta_1 Play_{im} + \beta_2 Movie_{im} + \beta_3 Match_m + \beta_4 PreTest_{im} + \beta_5 Gender_{im} + \beta_6 Age_{im} + \beta_7 White_{im} + \varepsilon_{im}$$

This is identical to Equation (2) except that it adds $PreTest_{im}$, which is the pre-treatment measure of the outcome γ_{im} , including Tolerance, SPT, Content Knowledge, and Theater Consumption and Participation.

Results

As shown in Table 2, providing students with the opportunity to leave school on a field trip to see a live theater performance produced a number of significant effects. Students given the opportunity by lottery to see a play score .142 of a standard deviation higher on the Tolerance scale than if they were in the control group. If we adjusted for the non-compliance produced by bad weather and a scheduling error, the estimated effect of actually seeing a play on Tolerance increases to .190 of a standard deviation. Being assigned to see a movie instead of a play appears to have no effect on Tolerance.

(insert table 2 about here)

Social Perspective Taking increases by .169 of a standard deviation for students randomly assigned to go on a school field trip to see a play. Again, the effect is larger for those who actually saw the play. Being assigned to see a movie, however, has no effect on Social Perspective Taking. Student's Content Knowledge of the plot and vocabulary in these stories is also increased when students see the play. Watching a movie did not convey this Content Knowledge as effectively as seeing the live performance.

Taking a field trip to see a play may strengthen student interest in consuming theater in the future, but that effect was only significant at $p < .10$. Being assigned to the movie treatment

produced a similar estimated effect, but it was not statistically significant even at $p < .10$. Neither the play nor movie treatments had any detectable effect on student interest in participating in future theater activities.

While results for each play are not presented here, the point estimates are roughly consistent across all plays. That is, if the overall effect is positive and statistically significant, the individual play estimates are all positive and many are also statistically significant. The only interesting result from the analyses of each individual play has to do with the effect of the movie treatment on Content Knowledge. Seeing a movie of *Twelfth Night* increased understanding of the plot and vocabulary of that story relative to the control group by .045 of a standard deviation, but that effect is only statistically significant at $p < .10$ and is almost one-third as big as the play treatment effect. Seeing the movie *Hook* had no effect on the Content Knowledge for *Peter and the Starcatcher*, but that is not too surprising given that the stories are not as well aligned as are the movie and play of *Twelfth Night*.

When we control for pre-treatment measures of the outcome, as presented in Table 3, the results generally remain the same. The Tolerance and Content Knowledge effects from the play treatment remain approximately the same size and continue to be statistically significant. The result for Social Perspective Taking shrinks in magnitude and falls short of being statistically significant once we control for a pre-treatment measure of SPT. Similarly, the Theater Consumption effect, which was marginal in Table 2, also falls short of being statistically significant. Generally null effects of the movie treatment persist even after controlling for pre-test measures of the outcomes with the possible exception of Content Knowledge. When we control for prior Content Knowledge, seeing a movie may increase understanding of the plot and

vocabulary of the stories by .051 of a standard deviation, but that effect is only significant at $p < .10$.

(insert table 3 about here)

Descriptively, it is interesting to note that female students tend to score higher on the Tolerance, Social Perspective Taking, Theater Consumption, and Theater Participation measures, but these differences mostly dissipate when controlling for pre-treatment measures of those outcomes. Similarly, white students score higher on Tolerance and Content Knowledge outcomes but that entirely disappears when controlling for pre-treatment measures of those outcomes. Generally, we find little evidence that the play or movie treatments had differential effects on students by gender or race/ethnicity.

Discussion

The experimental evidence presented here clearly shows that students can benefit from school field trips to see live theater. The effects are most robust with respect to measures of Tolerance and Content Knowledge. Regardless of which play they see or whether we control for pre-treatment measures of outcomes or not, students experience an increase in Tolerance as well as greater understanding of the plot and vocabulary of stories if they see live theater.

There may also be a benefit from seeing live theater for student ability to engage in Social Perspective Taking, but that effect falls short of statistical significance when controlling for a pre-treatment measure of that outcome. We gain precision by controlling for pre-treatment measures but we also lose sample size because those pre-test measures were only collected for the last two plays. Therefore, it is unclear whether we should have greater confidence in the results in Table 2 or Table 3.

There is some indication that students randomly assigned to see live theater become more interested in frequenting the theater in the future, but that effect is only marginally significant and disappears when controlling for the pre-treatment measure of that outcome. Student interest in participating in theater does not seem to be affected at all by this experiment.

It is also important to emphasize that the movie treatment does not seem to have a robust effect on any of these outcomes. Going on a field trip to see live theater produces benefits that cannot be produced by watching a movie instead. And the fact that students who received the movie treatment also left school for a field trip suggests that the effects we have observed are caused by the experience of watching live theater and not simply caused by leaving school.

While this experiment demonstrates that live theater field trips cause an increase in Tolerance and perhaps in the related concept of Social Perspective Taking, it cannot tell us why these effects were produced. Our best explanation is that theater is a window for students to a broader world. Exposure to that broader world may increase their understanding and acceptance of that broader world, which is why we see increases in Tolerance and Social Perspective Taking. Plays may be more effective than movies in helping students understand and accept that broader world because we react differently to human beings acting out a story in front of us than to representations of human beings on a screen. The in-person experience may create greater emotional connections.

It is educationally significant and a bit surprising that watching a movie is not a particularly effective way of conveying content knowledge while watching a play is. Watching movies is an extremely common school practice, but it may produce little learning. Going to see a play, on the other hand, is less common, but appears much more effective. This is especially surprising given that many films may be higher quality productions than plays accessible to

school groups. But as we saw in our experiment, even seeing a university play of *Twelfth Night* taught students significantly more plot and vocabulary than seeing the film with award winning actors, like Helena Bonham Carter and Ben Kinglsey. The in-person experience, again, appears to trump the skill of the actors.

Of course, we were only able to observe effects seven to eight weeks after students saw the plays, so we do not know if these benefits endure over longer periods of time. And while we saw consistent results across multiple plays, produced by different theaters, and involving different school groups, all of our experiments occurred in one particular place of the country. So, we cannot know with confidence that these benefits would be produced for other students.

Nonetheless, this is the first major piece of experimental evidence on how viewing theater affects students and it shows significant benefits. Before we eliminate or further cut these cultural experiences from schools, we should attempt to replicate this experiment in other locations and with longer-term measures of outcomes.

References

- Baron-Cohen, S., Wheelwright, S., Hill, J., & Raste, Y. (2001). The “Reading the Mind in the Eyes” test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241-251.
- Bowen, D. H., Greene, J.P., & Kisida, B. (2014). Learning to think critically: A visual art experiment. *Educational Researcher*, 43(1), 37-44.
- Gadsden, V.L. (2008). The arts and education: Knowledge generation, pedagogy, and the discourse of learning. *Review of Research in Education*, 32(1), 29-61.
- Gehlbach, H. (2004). A new perspective on perspective taking: A multidimensional approach to conceptualizing an aptitude. *Educational Psychology Review*, 16(3), 207-235.
- Gehlbach, H., Brinkworth, M. E., & Wang, M. (2012). The social perspective taking process: What motivates individuals to take another’s perspective? *Teachers College Record*, 114, 1-29.
- Gehlbach, H., Brown S. W., Loannou, A., Boyer, M. A., Hudson, N., Niv-Solomon, A., Maneggia, D., & Janik, L. (2008). Increasing interest in social studies: Social perspective taking and self-efficacy in stimulating simulations. *Contemporary Educational Psychology*, 33(4), 894-914.
- Goldstein, T. R., & Winner E. (2012). Enhancing empathy and theory of mind. *Journal of Cognition and Development*, 13(1), 19-37.
- Greene, J. P., Hitt, C., Kraybill, A., & Bogulski, C. A. (2015). Learning from Live Theater. *Education Next*, winter.
- Greene, J.P., Kisida, B., & Bowen, D. (2014). The educational value of field trips. *Education Next*, 15(1), 78-86.
- Jægar, M. M., & Møllegarrd, S. (2017). Cultural capital, teacher bias, and educational success: New evidence from monozygotic twins. *Social Science Research*, 64, 130-144.
- Kisida, B., Greene J. P., & Bowen D. (2014). Creating cultural consumers: The dynamics of cultural capital acquisition. *Sociology of Education*, 87(4) 281-295.
- Lacoe, J., Painter, G.D., & Williams, D. (2016). Museums as classrooms: The academic and behavioral impacts of “school in the park.” Working Paper. Sol Price School of Public Policy, University of Southern California.

- Lee, B. K., Patall, E. A., Cawthon S. W., & Steingut, R. R. (2015). The effect of drama-based pedagogy on preK-16 outcomes: A meta-analysis of research from 1985-2012. *Review of Educational Research*, 85(1), 3-49.
- Rabkin, N. & Hedberg, E. C. (2011). *Arts education in America: What the declines mean for arts participation*. National Endowment for the Arts, Research Report #52.
- Ruppert, S. S. (2006). Critical evidence: How the ARTS benefit student achievement. National Assembly of State Arts Agencies. Retrieved from: <https://nasaa-arts.org/wp-content/uploads/2017/05/critical-evidence.pdf>

Table 1: Student Characteristics in Treatment and Control Groups

	Control Group	Play Treatment	Movie Treatment	Total N
	Mean	Mean	Mean	
	Full Sample			
Female	0.609	0.550*	0.596	1467
Age	14.751	14.717	15.175	1463
Grade	8.628	8.605	9.059	1485
White	0.672	0.660	0.681	1467
Seen Play In Last Year	0.263	0.241	0.304*	1467
	<i>Twelfth Night & Peter and the Starcatcher</i> Sample with Pre-Test			
Tolerance	2.986	2.936	2.959	485
Social Perspective				
Taking	2.557	2.451	2.471	485
Content Knowledge	0.368	0.394	0.354	485
Theater Consumption	2.229	2.252	2.323	485
Theater Participation	1.961	1.676	1.870	485

Note: The outcomes of interest (Tolerance, Social Perspective Taking, Comprehension, Vocabulary, Theater Consumption and Theater Participation) were only pretested with students who went to see *Twelfth Night* and *Peter and the Starcatcher*.
 * $p < .10$, ** $p < .05$, *** $p < .01$, two-tailed, relative to control group.

Table 1 Note:

Note: The outcomes of interest (Tolerance, Social Perspective Taking, Comprehension, Vocabulary, Theater Consumption and Theater Participation) were only pretested with students who went to see *Twelfth Night* and *Peter and the Starcatcher*.

* $p < .10$, ** $p < .05$, *** $p < .01$, two-tailed, relative to control group.

Table 2: Impacts of Play and Movie Treatment on Five Primary Student Outcomes

	Tolerance		Social Perspective Taking		Content Knowledge		Theater Consumption		Theater Participation	
	Intention to Treat	Impact on Treated	Intention to Treat	Impact on Treated	Intention to Treat	Impact on Treated	Intention to Treat	Impact on Treated	Intention to Treat	Impact on Treated
Play Treatment	0.142** (0.065)	0.190** (0.087)	0.169** (0.080)	0.222** (0.110)	0.101*** (0.016)	0.135*** (0.018)	0.127* (0.069)	0.170* (0.092)	-0.109 (0.082)	-0.147 (0.105)
Movie Treatment	-0.038 (0.095)	-0.009 (0.098)	-0.045 (0.087)	-0.014 (0.097)	-0.015 (0.027)	0.006 (0.026)	0.084 (0.104)	0.110 (0.108)	0.008 (0.120)	-0.015 (0.125)
Female	0.290*** (0.058)	0.289*** (0.057)	0.407*** (0.093)	0.406*** (0.090)	-0.016* (0.009)	-0.016* (0.009)	0.359*** (0.053)	0.358*** (0.052)	0.414*** (0.058)	0.415*** (0.056)
White	0.231*** (0.073)	0.235*** (0.071)	0.090 (0.071)	0.102 (0.071)	0.060*** (0.014)	0.063*** (0.013)	0.035 (0.062)	0.038 (0.061)	0.013 (0.061)	0.009 (0.059)
N	1441	1441	791	791	1442	1442	1442	1442	1442	1442

Note: Standard errors - clustered within matched sets - presented in parentheses. The effect for theater consumption also holds constant whether students saw a play in the previous year.

Coefficients for the effect of age, grade, and matched set are not shown. The movie treatment only occurred for *Twelfth Night* and *Peter and the Starcatcher*.

* p < .10, ** p < .05, *** p < .01, two-tailed.

Table 2 Note:

Note: Standard errors - clustered within matched sets - presented in parentheses. The effect for theater consumption also holds constant whether students saw a play in the previous year. Coefficients for the effect of age, grade, and matched set are not shown. The movie treatment only occurred for *Twelfth Night* and *Peter and the Starcatcher*.

* $p < .10$, ** $p < .05$, *** $p < .01$, two-tailed.

Table 3: Impacts of Play and Movie Treatment on Five Primary Student Outcomes Controlling for Baseline

	Tolerance	Social Perspective Taking	Content Knowledge	Theater Consumption	Theater Participation
	Intention to Treat	Intention to Treat	Intention to Treat	Intention to Treat	Intention to Treat
Play Treatment	0.182** (0.071)	0.056 (0.076)	0.154*** (0.025)	0.130 (0.081)	-0.040 (0.064)
Movie Treatment	0.074 (0.074)	-0.064 (0.064)	0.051* (0.028)	0.082 (0.066)	-0.005 (0.057)
Pre-Test	1.046*** (0.066)	0.681*** (0.051)	0.424*** (0.057)	0.977*** (0.045)	0.581*** (0.025)
Female	0.071 (0.081)	0.312*** (0.094)	-0.013 (0.015)	0.042 (0.076)	0.084 (0.073)
White	0.052 (0.099)	-0.005 (0.062)	0.040* (0.021)	0.039 (0.065)	-0.027 (0.073)
N	475	475	475	475	475

Note: Standard errors presented in parentheses. Coefficients for age and for the matched set parameters are not presented. Outcome measures for baseline were taken for students who saw *Twelfth Night* and *Peter and the Starcatcher* only.

* $p < .10$, ** $p < .05$, *** $p < .01$, two-tailed.

Table 3 Note:

Note: Standard errors presented in parentheses. Coefficients for age and for the matched set parameters are not presented. Outcome measures for baseline were taken for students who saw *Twelfth Night* and *Peter and the Starcatcher* only. * $p < .10$, ** $p < .05$,

*** $p < .01$, two-tailed.