I. Goals and Objectives

In collaboration with an interdisciplinary team of learning scientists, museum educators, learning technologists (Parsons’ Design and Technology program and Tufts’ Center for Engineering Education and Outreach), game designers (Learning Games Network), middle school science teachers, and evaluators (American Institutes for Research), the New York Hall of Science (NYSCI) proposes an Investing in Innovation development grant to address Absolute Priority 2—Promoting Science, Technology, Engineering, and Mathematics (STEM) Education. Under this Absolute Priority 2, we focus on the goal of increasing the number of individuals from groups traditionally underrepresented in STEM, including minorities, providing them with access to rigorous and engaging coursework in STEM that will prepare them for college and/or careers in STEM. We will do this by developing, implementing, and evaluating a high-quality digital learning environment called SciGames (Competitive Preference Priority 10) that can be used by teachers to bridge informal and formal science learning environments in such a way that both students’ science achievement and affect are improved for a diverse group of students.

Better integration of informal and formal science education is called for as part of the National Action Plan for Addressing the Critical Needs of the U.S. Science, Technology, Engineering, and Mathematics Education System (National Science Board, 2007) and is a prominent objective in the recently released PCAST report on K-12 STEM Education for America’s Future (PCAST, 2010). The PCAST report not only calls for better coordination of STEM education activities, but also suggests that, “every middle school and high school should have a partner in a STEM field, such as a research organization, college, university, museum, zoo, aquaria [sic], or company…” (p. 102).

Each SciGame uses a range of technologies including sensors and digital input devices that layer on top of existing playground equipment (e.g. a simple playground slide enhanced with photogate sensors) to support students playing games in an informal setting. The gaming is carefully designed to insure that science content is integrated into the parts of the game that are the most fun to play. In playing the game, students strive to make intuitive sense of qualitative
data patterns they see between their decisions and the score they receive. During game play, the data is automatically logged and later uploaded to a digital app to bridge back to the formal science classroom. Back in the classroom, students use the digital app to further inquire into the classwide dataset. The digital app ultimately supports students formalizing their understanding of the target science concepts. We describe below a SciGame we developed and piloted.

With SciGames, we will increase the number of individuals from groups traditionally underrepresented in STEM by accomplishing the following objectives: 1) Develop, implement, and test three different technology-enhanced playground games, each of which is designed to leverage play as the medium of scientific exploration and discovery for three distinct areas of middle school physics subject matter (energy interconversion and friction, force and linear motion, and force and rotational motion; 2) Develop, implement, and test three digital computing applications, each paired to one of the playground games, all designed to extend the inquiry that begins on the school playground back into the science classroom, where the science concepts are formalized; 3) Package both playground game technology and digital apps into portable SciGames kits that teachers can set up and use in their own school playgrounds and science classrooms; and 4) Develop, implement, and test a teacher professional development program for using the SciGames with fidelity.

In the next section, we review the research literature that has informed the development of the SciGames model. In addition, we present a SciGame pilot study that suggests the potential of this approach.

II. Statement of Need

Informal science environments have been shown to have positive impact on aspects of students’ science affect, including intrinsic motivation (National Research Council, 2009; Zuckerman, Porac, Lathin, Smith, and Deci, 1978) and engagement (Tisdal, 2004), which evoke longer-term attitudinal changes, like interest in science. However, while informal experiences have been shown to affect students’ thinking about various science domains (National Research
Council, 2009), many students arrive at alternative conceptions that need to be replaced with more normative understandings (Driver, Squires, Rushworth, & Wood-Robinson, 1997).

“Typically, exhibition evaluations include self reports from visitors that they have learned some content knowledge, usually small-scale, counterintuitive facts rather than large-scale abstractions or principles” (National Research Council, 2009). Only occasionally has a science museum exhibit experience been shown to challenge a common alternative conception held by visitors (National Research Council, 2009; Borun, Massey, Lutter, 1993). Overall, informal science environments are acknowledged to be less effective in building the kind of formalized science knowledge that is the goal of schooling, especially without the time, sequencing, and consistency necessary for learners to develop systematically deep conceptual understanding (Bevan, Dillon, Hein, Macdonald, Michalchik, et al., 2010; DeWitt and Storksdieck, 2008).

Inquiry-based lessons in formal science classrooms have been shown to help students address alternative conceptions and improve their understanding of, and ability to use, scientific principles (Kanter, 2010; Kanter & Schreck, 2006; Kolodner, Camp, Crismond, Fasse, Gray, et al., 2003; Krajcik, McNeill, & Reiser, 2008; Linn, Bell, & Davis, 2004; Marx, Blumenfeld, Krajcik, Fishman, Soloway, et al., 2004; Rivet & Krajcik, 2004; Schneider, 2002). However, even formal inquiry-based instruction has limited positive impact on minority students’ science affect: motivation, engagement, and interest (Hickey, Moore, and Pellegrino, 2001; Kanter, 2009; Kanter & Konstantopoulos, 2010). Formal science education engages only a small percentage of students and has been less successful for low-income and female students or students from ethnic or racial groups underrepresented in science and engineering careers (Atwater, Wiggins, & Gardner, 1995; Brickhouse, 1994; Kahle and Meece, 1994).

To retain more and more diverse students in the STEM pipeline, we need better ways of combining elements from both informal and formal science learning environments to support student improvement in both their science affect and their science learning. This gap in existing infrastructure that bridges informal and formal science learning environments will be addressed by the proposed project as a means by which to achieve Absolute Priority 2—Promoting
Science, Technology, Engineering, and Mathematics (STEM) Education, focusing on individuals from groups traditionally underrepresented in STEM, including minorities. We will provide a new kind of technology-enhanced instructional experience for these students that bridges formal science classrooms and informal school playgrounds, while placing engaging and rigorous science content at the center of the experience.

We will develop and test a new approach to improve students’ science affect and science learning. A SciGame uses a range of technologies including sensors and digital input devices to support a special kind of game play on the school playground. The gaming is carefully designed to insure that science content is integrated into the parts of the game that are most fun to play. For example, as described in our pilot below, a simple playground slide can be enhanced with photogate sensors to measure velocity. Goals, rules, and scoring are used to frame the game-like experience so that the fun of repeatedly sliding to win the game at the same time motivates a series of experiments exploring concepts such as frictional force. Each technology-enhanced playground game is paired to a computer application or “digital app” that runs on desktop, laptop, or portable computing device that simulates the playground game. This allows students to explore quantitative data logged automatically on the playground while also supporting them continuing to play the game to further explore and more deeply formalize the learning of the target science concepts. Gaming in science classrooms has shown the potential to impact affect, motivation in particular (Dede, Ketelhut, and Ruess, 2002; Honey and Hilton, 2011; Ketelhut, Dede, Clarke, and Nelson, 2006; Neulight, Kafai, Kao, Foley, and Galas, 2007; Tuzan, 2004).

We will develop experiences for middle school students that will sustain the positive impact on affect that is the hallmark of informal settings, amplify that impact through the use of technologies that support students’ engagement, deepen their learning through repeated game-like experimentation, and sustain this positive impact on affect into the classroom where science concepts can be formalized. The proposed project will have the positive impact of improving both student science achievement and science affect, especially necessary for students from ethnic or racial groups that are underrepresented in STEM careers.
A pilot study of a SciGame, conducted in January of 2011, provides empirical evidence of the promise of the approach. We leveraged a five-year, 1.5 million dollar initiative at NYSCI called SciPlay, the Center for Play, Science and Technology Learning, to conduct a pilot study of a SciGame. We used technology to develop a sliding game. In this game, students are challenged to deplete their energy as they go down the slide. Students set a personal target for how much they plan to deplete their energy. They are free to experiment with factors that affect their frictional force and thus their energy loss due to friction as they slide. Students choose their sliding mat material and the amount of weight to carry on their lap. Sensor technology, in the form of pairs of photogates, is added to the top and bottom of the slide and LabView software is used to calculate the students’ instantaneous velocity at the start and end of their slide (see Appendix J1). From these velocities (along with students’ weight, entered at the game’s start), students’ kinetic energy at the start and end of their slide is calculated. Students’ weight and height above the ground is used to calculate their potential energy at the start and end of their slide. On a monitor at the bottom of the slide, students are shown a real-time display of the difference between their total energy (the sum of their kinetic and potential energies) at the start and end of their slide (see Appendix J1). They get feedback on the monitor as to how close they came to their predictions. Before sliding again, students decide what to vary and why. The game supports students qualitatively inquiring into the variables that govern frictional force and energy loss due to friction. The better a student’s grasp of these qualitative relationships, the better they can score. Learning these physical science concepts is mandated by the middle school standards in the New York public schools and these same concepts are also in the middle school band of Core Idea PS2: Motion and Stability: Forces and Interactions in the newly-released NRC Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas, which will serve as the basis for a new national science standards effort.

Our pilot findings, presented next, support the idea that a SciGame like this can have a positive impact on improving student science achievement and affect. A total of nine middle
school students participated in the pilot. On average, students played the slide game for 20 minutes. A researcher facilitated students’ engagement with the activity.

**Impact on Affect.** To examine the effect of a SciGame on students’ affect, we assessed engagement with the activity. We expected to see student engagement expressed in three ways: behavioral engagement (spontaneous play and persistent play), emotional engagement (manifest joy), and cognitive engagement (students engaging in diverse experimentation). Students’ game play was video recorded, and each student was also provided with a digital recorder wristband to clearly capture their audio. Affective outcomes were evaluated through qualitative analysis of these observations, coding the video for the different types of engagement. The Children’s Playfulness Scale (CPS) (Barnett, 1991), a scale for the behavioral manifestation of playfulness, was used in constructing our observation rubric due to its high reliability across samples (Cronbach’s alpha coefficient = 0.88) and construct validity (Trevlas, Grammatikopoulos, Tsigilis, & Zachopoulou, 2003). Unlike the more commonly used self-report survey measures, observation tools such as this provide the means by which to assess engagement directly through students’ behaviors.

Students’ behavioral engagement was examined by observing whether they spontaneously engaged with the game without any extrinsic reward. On average, each student experimented eight times down the slide without any external encouragement. Students would rush to engage in subsequent runs, and would urge others to hurry up for their turn. None of the participating students stopped playing the game until the researcher ended the pilot session, another manifestation of behavioral engagement.

Students manifested emotional engagement through observed manifestations of enjoyment or prompted verbalizations. On average, each student exhibited seven instances of manifest joy during the pilot in the form of laughing, verbalizations (e.g., “Oh god, that was funny!”), singing, and even dancing.

Students manifested cognitive engagement by their spontaneous experimentation sliding down with the different mat materials to try to manipulate frictional force. We saw students
replicating or improving on other students’ strategies or developing their own methods of experimentation. On average, each student demonstrated four different unanticipated experiments, in addition to those we intended to support. For example, students experimented with sliding down the slide on a cart with wheels instead of the provided mats to try to manipulate the frictional force and ultimately the energy loss due to friction.

Admittedly this is a small sample, but the results are promising. Our analysis supports the theorized impact on affect as seen in emotional, behavioral, and cognitive engagement.

**Impact on Learning.** To measure the impact of a SciGame on learning, we looked at students’ prior knowledge and learning outcome data. Students’ prior knowledge on the topic of friction and energy was assessed with an open-ended question asking students to describe the concepts of potential and kinetic energy prior to engaging with the pilot activity (i.e., “Can you describe what these [potential and kinetic energy] mean? If you don’t know, just leave it blank.”). The purpose of the pre-assessment was to examine whether a student already understood these concepts. Students prompted verbalizations during the pilot were also analyzed for their conceptual understanding of the learning goals. These were used instead of a post-assessment in order to capture students’ new knowledge in a timely manner. The prompts for the learning outcome assessment included questions such as: What can you do to lose more energy? Why is that mat making you lose more energy?

Two out of nine students could articulate their understanding of kinetic and potential energy prior to their engagement with the SciGame. Three students had some prior knowledge of the target science but could not correctly define the concepts. The rest of the students had not heard of the concepts before (they left their responses blank). Despite different levels of prior knowledge, students showed a somewhat improved understanding of the science learning goals. Six out of nine students that participated in the pilot study could explain something about what reduces frictional force. The six students who demonstrated an understanding of the relationship between the choice of material and frictional force were further prompted with a question about the relationship between the frictional force and the energy loss. When prompted with the
question—will you lose more or less energy? Four out of these six students could verbalize a relationship between frictional force and energy loss.

While preliminary, these findings suggest that a technology-enhanced playground game can support learning while sustaining affect. These findings also emphasize the challenge in getting all students to the formal generalization of the target science concepts, and this is precisely why one needs to add additional technology to the SciGame. A digital app can be used to have students inquire further into the quantitative data logged during the playground game. The digital app will be designed to simulate the playground gaming context. (See Appendix J2 for a mock-up of what the digital app for the slide SciGame might look like.) To formalize the understanding of the target concepts, it has been shown to be necessary to add this kind of scaffolded reflection to an intrinsically motivating game (Habgood and Ainsworth, 2011; Honey and Hilton, 2011). With the addition of a digital app, we expect that a SciGame can boost affect in addition to having a positive impact on science achievement, which is itself reinforced by this boost in affect. With SciGames, we have a way to provide rigorous and engaging science coursework to students from underrepresented groups that has the potential to make a positive impact on learning and affect to attract into and retain these individuals in the pipeline to college and/or careers in STEM.

III. Design of the Development Project

The target population for this development project will be 8th grade science students and their teachers drawn from New York City public middle schools with a majority of students who are from racial and ethnic groups underrepresented in STEM careers. We will recruit students of 40 teachers from 9 schools that are all within five miles of NYSCI for the convenience of teachers being able to attend the design team meetings after school and to support our ability to conduct research in teachers’ playgrounds and classrooms. We will also target students from schools that have playgrounds maintained by the NYC Parks and Recreation Department. Given that NYC Parks is a partner in this work, targeting these schools will give us the opportunity to
implement our technology on these school playgrounds. The New York City Department of Education has helped us identify high-need middle schools from which we will recruit teachers and their students. Most schools have several 8th grade teachers. Appendix J3 summarizes the demographic data of the identified schools.

**Instrument Development and Methods of Analysis**

During Year 1, we will develop and test the instruments and analytic approaches necessary to design and develop the three SciGames. We will collect data to verify the intended use of the model components in both the school playground environment and classroom settings. We will also collect data to assess changes in students’ science affect and content learning. We will use this data as the basis for making iterative improvements to the SciGames. This data will support the cycles designing, testing, and redesigning the SciGames. The details of the instruments themselves are described under Project Evaluation, below.

**Design Cycles**

In the winter of Year 1, the design team, which includes five teachers from our 9 partner middle schools, research staff, and technology and game development partners, will review the piloted SciGame to elucidate its “a priori foundations” (Clements, 2007), specifically the goals for impact on affect and learning and the general approach by which to impact these goals. Also in this part of Year 1, the design team will create final versions of all existing elements of the SciGame for the playground slide game, as well as new elements such as the paired digital app, which was not previously piloted.

**Pilot Evaluation of SciGame Elements.** Pilot evaluation will begin in the spring of Year 1. We will explore how the five teachers and their students use the individual SciGame elements and if they use them to perform actions consistent with how they are intended to impact student affect and learning (Miles & Huberman, 1984, 1994). We will use descriptive checklist matrices to display data that are distinct indicators of students or teachers using the elements for their intended purposes. We will populate these matrices with observational data of students and teachers. For each element, we can draw conclusions by looking down the columns of the
checklist matrix to note patterns or themes and form a general idea of the extent to which each element was used as intended. We will verify conclusions by replicating each finding in another part of the data or looking for negative evidence within the data set (Creswell, 2008).

**Prototype Pilot Testing.** We will then assemble the individual elements into a completed prototype of the first SciGame that now includes the digital app and begin pilot testing in the spring. Questions to be explored include: Can students and teachers use the elements as a coherent whole? Do they use them to perform the desired actions? The nature of the data and analysis is similar to the above, but the focus is now expanded to include how the elements operate in an interrelated manner.

We will also collect more comprehensive data on how this first SciGame impacts student affect and learning, using the instruments and analytic approaches described below. The results of the prototype testing will be shared with the design teams, who will redesign during the summer. Also during the summer, we will develop and deliver teacher professional development (PD) for this first SciGame. We will design our PD around the five requirements laid out by Loucks-Horsley et al. (2003) to ensure that the experience produces “transformative” learning for teachers, promoting change in teachers’ deeply held beliefs, knowledge, and habits of practice. Many such changes are necessary for science teachers to use SciGames with fidelity. To create this kind of transformative learning experience for teachers, we will use a “practice-based” approach to teacher PD, supporting teacher learning in the context of teachers addressing issues encountered in practice (Ball & Cohen, 1999; Loucks-Horsley et al., 2003; Marx, Freeman, Krajcik, & Blumenfeld, 1998). We will support this by having teachers plan, then use one or more of the SciGames with a field trip group visiting NYSCI’s Science Playground, then reflecting, i.e. practice-based professional development as is advocated (Ball & Cohen, 1999; Loucks-Horsley et al., 2003; Marx, Freeman, Krajcik, & Blumenfeld, 1998).

**Field Testing.** During the fall, a new group of five teachers and their classrooms will pilot the first SciGame. These new teachers will be recruited from the same group of partner schools. We will make every attempt to engage teachers in pairs or trios from a given school in
order to avoid teachers feeling isolated field testing in their school. These new teachers will receive the PD training described above prior to field testing. The goal is to determine how and why the SciGame might work differently for these teachers and redesign it to work for a wider variety of users. Once again, the analysis will focus on students’ science affect and science learning, as well as the degree to which teachers are able to work confidently with the model. Field test findings will be used to make further refinements to this first SciGame. The redesign will be completed during the winter of Year 2.

These same design cycle phases will be repeated in Year 2 to develop both the second and third SciGames. We anticipate that the second SciGame will focus on a scooter cart game to learn about Newton’s Second Law of Motion as applied to linear motion. Students will select a target motion (e.g. low constant velocity, high constant acceleration, etc.) and then try to push their friend(s) on a scooter cart along a straight line on the playground to match the target motion. We anticipate using real-time motion tracking to provide an easy-to-interpret representation of students’ motion to compare to the target motion and the amount of force students are applying during their run. Students will experience first-hand the difference between velocity and acceleration and how force relates to acceleration. As they try to win the game they will have to induce the qualitative relationship between force, mass, and acceleration. We anticipate that the third SciGame will focus on a croquet-type game with a ball and mallet using similar real-time motion tracking as students try to hit the ball around a circle to learn about Newton’s Second Law of Motion as applied to rotational motion. In Year 2, we will engage new teachers from the same nine partner schools- five new teachers to design and prototype in the winter and spring, and after receiving PD in the summer, five additional teachers to field test in the fall. All teachers from Year 1 will again use the first SciGame in addition to being trained and then using the other two new SciGames.

During the winter of Year 3, we will complete all redesign work and arrive at final versions of the second and third SciGames. In the spring, we will take an initial pass at packaging the technology components of the three SciGames into easy-to-use kits. In the
Investing in Innovation – i3 – Development Grant                                New York Hall of Science

SciGames PI: Kanter

summer, we will provide PD and kits to all 20 remaining teachers from the nine partner schools. This new group of teachers will use all three SciGames in the fall, as will all teachers from Years 1 and 2. Also in Year 3, with the help of our partner the New York City Department of Education, we will reach out to 9 new schools for 40 new teachers who will serve as controls in a quasi-experimental impact study. This helps guard against confirmation bias by allowing us to determine to what extent the SciGames developed within an initial group of nine schools are portable to the different contexts of other schools (Shavelson, Phillips, Towne, & Feuer, 2003).

Assuming that we have seen indications of success of the SciGames in years prior, we will use Year 4 to conduct an experimental impact study. During Year 4, 20 of the 40 new teachers brought on during Year 3 will be randomly assigned to implement all SciGames and the other 20 will again serve as controls. During the spring and summer, we will provide PD and kits to 20 teachers who will implement SciGames in the fall. We will continue to support as many teachers from earlier years as wish to continue using the SciGames on their playgrounds and in their classrooms.

During Year 5, the 20 teachers who served as controls in the experimental study will be provided with PD and kits to insure they have the chance to use the SciGames with their students. Again, we will support as many teachers from earlier project years as wish to continue
using the SciGames. This last year will also be focused on writing up our findings. The advisory board will meet once annually to review progress. The figure above presents a summary of this project timeline and milestones. A thick line indicates a heavier emphasis on a given activity during this time. A thinner line indicates that some work on a given activity is taking place during this time.

Given that teachers continue in the project after their first year of involvement as outlined above, and assuming each teacher works with 3 classrooms of 25 students, we anticipate this work will reach approximately 12,000 students and 80 teachers by the end of the project. Thus, the estimate of the cost of the proposed work is approximately $288 per student annually. Given this estimate, it would cost $28.8 million, $72 million, and $144 million to reach 100,000, 250,000, and 500,000 students, respectively, although as a development grant it is not expected that we will reach such scaling targets. While economies of scale will make these costs much lower as we scale up the SciGames, $288/student is a reasonable cost given the potential of the proposed project to provide participating students a science learning experience that is both rigorous and engaging and thus has the potential to attract and retain students from underrepresented groups into the STEM pipeline. This cost is an investment in tripling the proportion of underrepresented minorities in science and engineering careers as is estimated to be necessary to meet the future needs of a globalized STEM-driven economy.

IV. Project Evaluation

Methods appropriate to size and scope of project: The evaluation of SciGames consists of two studies. The SciGames Implementation Study evaluates for formative purposes the extent to which 8th grade students and teachers implement SciGames as intended in playground and classroom settings, students learn the physics concepts to be learned through SciGames, and the professional development helps teachers implement SciGames. The Implementation Study will take place during Years 1 and 2, and will involve five different pairs of teachers each year. Each pair is located at one of the nine partner schools and is comprised of the teachers involved in the
Investing in Innovation – i3 – Development Grant

SciGames PI: Kanter

pilot and field testing, respectively. This study will provide timely and useful feedback for continuing development of SciGames and its professional development; The quasi-experimental and experimental Impact Study evaluates the impact of SciGames on 8th grade students’ science achievement and attitudes towards science, thereby providing evidence of the promise of the program for improving student outcomes. The impact study will occur during Years 3 and 4. In Year 3, the study will employ a quasi-experimental design to compare all 8th grade middle school teachers at the nine partner schools to similar teachers at nine matched comparison schools. In Year 4, an experimental study will randomly assign teachers from the Year 3 comparison schools to treatment or comparison conditions.

Key questions and proposed methods for addressing them: The Implementation Study evaluation questions are as follows: 1) Do students exhibit both cognitive and affective engagement with SciGames activities as intended? 2) To what extent do teachers and students implement the playground games and classroom inquiry activities as intended? What are facilitators and barriers to implementation? What are teacher opinions about the quality and effectiveness of the program? And 3) How effective is the professional development (PD) for preparing teachers to implement SciGames? How can the PD be improved?

The Impact Study evaluation questions are: 4) What is the impact of participating in SciGames’ formal and informal learning opportunities on students’ knowledge of science concepts and attitudes toward science? and 5) With what degree of fidelity do teachers implement the SciGames model?

Data sources: The implementation and impact studies will use data from five sources.

1. Performance Data. Data collected from both the playground games and the digital apps will provide several measures of program implementation (Questions 3 and 6, respectively). These data will indicate the extent to which the games and apps are used (i.e., number of instances of playground use, number of application simulations and number of days of use), and whether patterns of use reflect systematic inquiry (e.g., holding variables constant).
2. **Administrative data** from schools will provide descriptive data on student participants, such as GPA and demographics (e.g., gender, ethnicity), as covariates for the Impact Study (Question 5).

3. **Interviews, focus groups, and observations** will be conducted in Years 1 and 2 involving the five pairs of teachers in each year. These will be conducted during two site visits, each one day in duration, during the fall of each year. The evaluator will interview the participating teachers, conduct a student focus group, and observe instruction and project activities. During Year 1, the site visit will involve one evaluator, and during Year 2 (when there are three games implemented), it will involve two evaluators. The first visit, which will occur in the fall prior to any program implementation, will focus on teachers’ and students’ prior experiences with learning opportunities similar to SciGames. The second site visit will occur later in the fall and will focus on student engagement in the SciGames activities (Question 1), and teacher experiences with both program implementation (Question 3) and professional development (Question 4). Teacher interviews and student focus groups will be coded and categorized using an a priori framework as well as an inductive approach.

4. **Pre-post student survey** will be administered during the fall of each year, both prior to and upon completion of the last SciGame. The survey will be comprised of scales adapted from three previously developed surveys. Student attitudes towards science and academic-and career-related behaviors in science (Questions 2 and 5) will be measured using scales from the Test of Science-Related Attitudes (TOSRA; Fraser, 1982). On the post-survey only, the survey will measure student affective and cognitive engagement in SciGames activities (Question 1). We will adapt the engagement subscale of the Rochester School Assessment Package (RAPS-S; Tucker, Zayco, Herman, et al., 2002) that has been used to measure behavioral and emotional engagement. To measure cognitive engagement, we will use the active cognitive engagement sub-scale from a survey developed by Meece, Blumenfeld, and Hoyle (1988). Rasch analysis will be used in survey validation and the development, if appropriate, of construct level scores (Wright and Masters, 1982).
Investing in Innovation – i3 – Development Grant

5. **Teacher Surveys.** Teachers will complete two surveys. A web-based program implementation survey will be administered to all program teachers in the late fall of each year, asking them to indicate their level of implementation of each SciGame. Teachers will also describe barriers encountered and opinions of the games (Questions 3 and 6). In Years 3 and 4, comparison teachers will also complete a parallel implementation survey, except they will report on their curriculum and instructional practices teaching the science topic addressed by each of the three SciGames (Question 6). Rasch analysis will be used in survey validation and the development, if appropriate, of construct level scores.

6. **Assessment of Student Science Knowledge.** Students’ physical science content learning will be assessed by a written pre- and post-assessment (Questions 1 and 5). The assessment will consist of items from the Force Concept Inventory, developed to assess students’ basic concepts in Newtonian mechanics (Hestenes, Wells, & Swackhamer, 1992), adapted to address the New York State standards for eighth grade physics. The assessment will be administered prior to and after students’ use of the SciGames to assess their prior knowledge and learning outcomes. Here too, Rasch analysis will be used in survey validation and the development, if appropriate, of construct level scores. For a summary of the evaluation questions and their data sources, see Appendix J4.

**Implementation data and periodic progress assessment:** SciGames digital performance data, along with the teacher implementation survey, enable the collection of extensive data on implementation of the program components (Question 3). These data will be used to compare level of implementation across games, among components within games, and across teachers. To understand variations in implementation, these data are complemented by teacher and student reactions to the games (Questions 1 and 3), as reported through teacher interviews, student focus groups, as well as student and teacher surveys. Findings from the analysis of these data will be reported to and discussed with the SciGames team shortly after the conclusion of fall data collection each year. Additional feedback on whether the SciGames promote the anticipated
changes in science knowledge and attitudes will be provided in March following each program year. This feedback will be used in the iterative design process.

**Information for further development, replication, and testing:** Similarly, the SciGames digital data and teacher survey will depict implementation fidelity through a comparison to the anticipated use model. The teacher interviews will help to clarify the conditions (e.g., playground equipment, previous training) that are necessary or desirable for successful implementation. A comparison of participating and non-participating teacher survey responses in Years 3 and 4 will establish a contrast between these two groups. This information, when combined with other information generated by the SciGames program and its evaluation, will, we expect, suggest the benefits of SciGames and thus the need for further development, replication, and testing. SciGames and the evaluation already conducted will provide much of the infrastructure needed to support further implementation and study.

**Determining causal inference:** American Institutes for Research proposes a two-step approach to the evaluation of the impact of participation in SciGames on students’ science knowledge and attitudes. Step one (to be conducted in Year 3 of the grant) will provide initial estimates of impact near the end of the SciGames development period, using a quasi-experimental matched comparison group study to determine the impact on student outcomes for students in the 9 treatment schools in comparison to students in 9 matched schools not participating in SciGames. In Year 3, all 40 eighth grade science teachers in the nine schools will be implementing the three SciGames. To examine the preliminary, quasi-experimental impact of this program on the teachers’ students, we propose to recruit 9 comparison schools. These schools will be selected with the intent of creating a comparison group with similar contextual and demographic characteristics (both school- and student-level) to the treatment sample. Statistical analysis of the Year 3 quasi-experimental impact will use a hierarchical linear modeling approach with school random effects, and treatment “assigned” at the school level. See Appendix J5 for the format of the mixed model.
To determine the minimum detectable effect size, we conducted a power analysis based on the two-level design described above (Schoch et al., 2005). The power analysis assumptions included a two-tailed test, Type I error rate of .05, statistical power of .80, and a sample size of 18 schools, and 6,000 students (80 teachers with 75 students each in three sections). We generated a table of the minimum detectable effect size (MDE) for 9 scenarios (varying by intraclass correlation and R-squared, chosen to represent plausible ranges for these parameters). Appendix J6 summarizes the results of this analysis. Thus, the MDE estimates represent possible power estimates over a variety of circumstances. Assuming an intraclass correlation of .10, and an R-squared of .50 (given that the study will include a pre-measure of the outcome), the MDE for this quasi-experiment is estimated to be 0.299 (which seems plausible to detect given the proposed proximal nature of the outcome measures).

The second step (to be conducted in Year 4 of the grant) will use a multi-site cluster randomized trial to compare the impact of SciGames on students of 20 teachers randomly assigned to treatment within 9 additional schools, in comparison to the students of 20 teachers randomly assigned to the control condition. During the fourth year of the grant, we will implement a within-school (multi-site) cluster randomized trial to rigorously determine the impact of SciGames within a set of 9 new schools. The experiment will randomly assign teachers within schools to either teacher using SciGames, or not. Within the 9 schools, a total of 20 teachers will be assigned to treatment, and 20 teachers will be assigned to the control. It is assumed that each teacher will be teaching 3 sections of 25 students each. Statistical analysis of the Year 4 experimental impact will use a hierarchical linear modeling approach with school fixed effects and teacher random effects. See Appendix J7 for the format of the mixed model. For formative information, the model can also be adapted to estimate the treatment effect for each individual school to help determine which schools may be seeing greater or lesser impact of SciGames.

We again conducted a power analysis to determine the MDE for this experiment. The sample size used for the analysis was 40 teachers across 9 schools, with 75 students per teacher.
Appendix J8 summarizes the results of this analysis. Assuming an intraclass correlation of .10, and an R-squared of .50 (given that the study will include a pre-measure of the outcomes), the MDE for this quasi-experiment is estimated to be 0.210 (which seems like a plausible effect size to detect given the proposed proximal nature of the outcome measures).

A logic model that connects the SciGames intervention’s inputs to its anticipated outcomes through its processes and activities is provided in Appendix J9.

**Sufficient resources for an effective evaluation:** A budget of $344,745 over five years (see budget narrative), 10 percent of the total budget, is proposed for the evaluation. The resources are sufficient for the staff time, travel, and communications needed to conduct the evaluation, to participate in the Investing in Innovation community of practice, and to assist project staff in contributing to national evaluation activities.

**Capacity to evaluate innovative practices:** The research staff from American Institutes for Research (AIR) heading up the implementation and impact studies for this project, whose abridged bios can be found in the Management Plan, have expertise and experience that uniquely prepares them to evaluate the technology-enhanced SciGames, the development and testing of which is the focus of this proposal. The AIR team has led the implementation and impact studies of particular relevance for the proposed innovation and its evaluation: 1) an external evaluation of Iowa’s Every Learner Inquires program, a K-12 professional development effort that focused on promoting inquiry-based science instruction. AIR conducted a mixed-methods evaluation to examine program implementation and impact. The methodology of the evaluation included surveys, interviews, observations, and classroom logs; and 2) an external evaluation of the Cryptoclub: Cryptography and Mathematics Afterschool and Online. The evaluation is providing formative information about the program’s impact on student knowledge and interest, skill in mathematics, and interest in mathematics, using an assessment of math skills, a survey of student attitudes towards math, and post-professional development feedback surveys.

These studies exemplify the AIR team’s expertise in and capacity to evaluate technological educational innovations. During its 65 years, AIR has evaluated the
implementation and impact of such innovations designed to improve the educational performance, achievement, and attainment of at-risk students in thousands of schools using a wide range of designs, data collection methods, and analytic approaches. AIR has conducted evaluations for the U.S. Department of Education and many of the nation’s state and local education agencies. The evaluations serve both formative and summative purposes (i.e., they focus on program intention, design, and implementation, as well as effectiveness).

V. Capacity to Bring the Project to a Larger Scale

The New York Hall of Science reaches out to teachers through its existing portable lab kit rental and related professional development program. Renting out the portable SciGames kits will be incorporated into this ongoing program. The SciGames kits will also be integrated into the existing NYSCI program through which teachers bring approximately 13,500 students per year to the NYSCI’s 60,000 square foot Science Playground for field trips. We will approach the middle school teachers who book visits to have their students play SciGames on the Science Playground, and then use the SciGames digital apps back in the classroom. Lastly, our partner the City of New York Parks and Recreation Department has asked us to use the SciGames professional development to train playground associates who will integrate the use of the SciGames into the summer youth programs that they run citywide. This program reaches over 100,000 students each summer. By integrating SciGames into existing programmatic work of NYSCI and its partners, we will bring the project work to a larger scale.

VI. Management Plan

Partners and Key Personnel

The interdisciplinary team assembled, with expertise in science, science education, learning technologies, and game design, contains the essential expertise and previous experience required to design, develop, build, and test this novel SciGames approach to STEM learning that bridges across informal and formal settings.
New York Hall of Science (NYSCI): NYSCI is New York City’s hands-on, interactive science center serving nearly 500,000 visitors each year. There are over 100 languages spoken in a 5-mile radius of NYSCI, and visitors mirror the immigrant diversity of Queens and of the greater metropolitan region. It is precisely this audience that the New York Hall of Science serves most directly, engaging new generations of young people, building their skills and supporting their interests to futures in STEM. During the past two years, under the direction of Dr. Margaret Honey, NYSCI has been positioning itself to serve as an innovation partner with the formal education sector.

**David Kanter, Ph.D. (PI)** is the Director of the Center for Play, Science, and Technology Learning (SciPlay) at the New York Hall of Science. He has a Ph.D. in Biomedical Engineering, and was previously a Curriculum, Instruction, and Technology in Education and Biology faculty member at Temple University and a Learning Sciences and Biomedical Engineering faculty member at Northwestern University. Kanter was PI of a 6-year large-scale $1.25M NIH/NHLBI Minority K-12 Initiative for Teachers and Students and co-PI of a 5-year large-scale $923K NIH/NCRR Science Education Partnership Award. He has experience leading large science education development and research projects, of the size and scope of that proposed, through to presentation and publication of findings. Kanter was also an NSF Postdoc for Science, Math, Engineering, and Technology Education in the Learning Sciences.

**Margaret Honey, Ph.D. (Co-PI)** is the President & CEO of the New York Hall of Science. She is widely recognized for her work using digital technologies to support children’s learning in the STEM disciplines. She holds a Ph.D. in developmental psychology. Prior to joining the New York Hall of Science, she served as a Vice President of the Education Development Center (EDC) and Director of the Center for Children and Technology. Honey was the architect and overseer of numerous large-scale projects funded by organizations including the National Science Foundation and the Institute for Education Sciences. She recently edited the National Research Council report, Learning Science: Computer Games, Simulations, and Education (Honey and Hilton, 2011).
Tufts University Center for Engineering Education and Outreach (CEEO): CEEO is a leader in efforts to integrate engineering into K-12 education. CEEO’s goals are to increase student and teacher excitement for learning STEM, improve student and teacher skills so learning is more enjoyable in all subjects, increase the general public's technological literacy, and increase awareness of the importance of STEM.

Ethan Danahy, Ph.D. (Playground Game Design Consultant) has a Ph.D. in Electrical Engineering. He is Engineering Research Program Director at CEEO, where he designs, implements, and deploys educational technology tools. Danahy previously worked with Kanter on the SciGames pilot.

Parsons The New School for Design’s Design and Technology Program: Parsons The New School for Design’s Design and Technology program trains students in areas of study that include Interaction (mobile, games, web, and installation) and Physical Computing (programming code and chip-based applications).

David Carroll, MFA (Playground Game Design Director) is Associate Professor of Media Design and Director of the MFA Design and Technology program. He has an MFA in Design and Technology. His research is in digital media, especially for mobile devices, and in software and interaction design as social engagement.

Learning Games Network: Learning Games Network (LGN) is a non-profit organization established to spark innovation in the design and use of learning games. LGN is committed to the development and distribution of new games informed by research in the learning sciences across a complete range of subject areas.

Eric Klopfer, Ph.D. (Digital App Design Director) is Associate Professor and Director of the MIT Scheller Teacher Education Program and president of the Learning Games Network. His work on educational games and simulations combines development of new technologies with implementation and research in formal and informal settings. His most recent work includes Ubiquitous Games, science-based games for the mobile web. He is the author of Augmented Learning: Research and Design of Mobile Educational Games.
Scot Osterweil (Digital App Designer) is the Creative Director of the Education Arcade and a research director in the MIT Comparative Media Studies Program. He is a designer of award-winning educational games, working in both academic and commercial environments, and his work has focused on what is authentically playful in challenging academic subjects. He has designed games for computers, handheld devices, and multi-player on-line environments.

American Institutes for Research: American Institutes for Research (AIR) conducts rigorous quantitative and qualitative research that identifies effective programs and resources as well as provides the technical assistance to translate reliable evidence into effective action. AIR’s Center for STEM Education and Innovation focuses on using technology, evidence, information, and innovative practice to support continuous improvement and accountability in STEM education.

Jonathan Margolin, Ph.D. (Evaluation Study Director and Implementation Study Designer) obtained his doctorate in social psychology and is a senior researcher at AIR. Margolin directed a statewide initiative to promote inquiry-based science in Iowa. Margolin designed the evaluation to ensure that the findings provided actionable answers to the project questions. Margolin also directs the statewide evaluation of the Title II-D Enhancing Education Through Technology in Vermont, and the evaluation of the Schools of the Future program sponsored by the Hawaii Community Foundation.

Andrew Swanlund (Instrument and Impact Study Designer) is a senior statistician at AIR. He specializes in psychometrics, statistics, and research design. He serves as analytic support across a variety of evaluation and impact studies including an experimental study of Indiana’s diagnostic assessment system. He has a wealth of experience working with states, districts, and foundations to design and implement rigorous evaluations of their programs.

Lawrence Friedman, Ph.D. (Evaluation Study Monitor) has a Ph.D. in Philosophy and is a managing director at AIR with more than 25 years of experience studying education reform and innovation including leading the evaluation of the nationwide Self-Developing Schools Project for the Soros Foundation–Hungary and the study of technology initiatives in the Los Angeles Unified School District.
Angeles Annenberg Metropolitan Project. He also co-leads an evaluation of the Virtual High School.

**New York City Department of Education:** The New York City Department of Education is the largest system of public schools in the United States. They are a partner in this work in having committed to identifying and brokering the placement of our project in middle schools in addition to supporting our needs for student standardized test and demographic data. (Please see letter of support.)

**City of New York Department of Parks and Recreation:** The New York City Department of Parks & Recreation is a partner in this work in that they jointly operate NYC school playgrounds and have committed to advising the project on how to build the SciGames kits to work well on school playgrounds. They have expressed an interest in using the SciGames on their community playgrounds. (Please see letter of support.)

**Advisory Board:** The following individuals have agreed to advise the project over its five years, attending one meeting a year: Bronwyn Bevan, Ph.D., Director of the Center for Informal Learning and Schools at the Exploratorium; Jan Plass, Ph.D., Professor of Educational Communication and Technology at New York University and co-Director of the NYU Games for Learning Institute; and David Hammer, Ph.D., Professor in Education and in Physics at Tufts University.

**Management Plan**

The following management plan details the how the objectives of the project will be achieved as per the project timeline and its milestones detailed above. Kanter (PI) will be responsible for oversight of the project as a whole, overseeing the work of all subteams and project partners, and the advisory board. Honey will serve as mentor to the PI. The subteam working on the playground games will consist of Kanter, Danahy, Carroll, and the Research Fellow. Danahy will help conceptualize the hardware and software interfaces for the playground games. On this subteam, Carroll will manage the two Parsons Design and Technology graduate students who will be building out the hardware and software for the playground games. The
subteam working on the digital apps will consist of Kanter, Klopfer, Osterweil and the Research Fellow. Osterweil will manage the Learning Games Network designer/developer and graphic artist who will be building out the digital apps paired to each of the playground games. Margolin will lead the work of the subteam working on evaluation, managing Swanlund and Friedman, and conducting the implementation study. Swanlund will develop all instruments (surveys, assessments, and observation tools) and design and conduct the impact study. Friedman will ensure that the study has full access to resources and that the study meets high quality standards. The Research Fellow will be hired to have a background in evaluation to be able to serve as a liaison between this evaluation subteam and the design and implementation subteams. The implementation subteam will be led by the Professional Development Specialist who will design and deliver the teacher professional development program, supported by the Research Fellow. The PI, co-PI, and subteam leaders will meet weekly by phone or videoconference, and the research group as a whole will meet twice monthly by videoconference.