Communication, Society and Media
ISSN 2576-5388 (Print) ISSN 2576-5396 (Online)
Vol. 5, No. 2, 2022
www.scholink.org/ojs/index.php/csm

Original Paper

Putting the Stars within Reach Using NASA 3D Data-Based Models

Kimberly K. Arcand[1*]

[1] NASA's Chandra X-ray Observatory, Center for Astrophysics | Harvard & Smithsonian 60 Garden Street | MS 70 | Cambridge, MA 02138, United States

[*] Kimberly K. Arcand, NASA's Chandra X-ray Observatory, Center for Astrophysics | Harvard & Smithsonian 60 Garden Street | MS 70 | Cambridge, MA 02138, United States

Received: May 20, 2022　　Accepted: May 31, 2022　　Online Published: August 22, 2022

doi:10.22158/csm.v5n2p37　　URL: http://dx.doi.org/10.22158/csm.v5n2p37

Abstract

This study investigated the quantifiable effects of data-based 3D models and prints on spatial reasoning skills and interest in science, technology, engineering, and mathematics (STEM) fields, for n = 100 youths aged 9-12 (99 female and 1 non-binary), primarily from traditionally underrepresented groups in STEM. In a pre-post design, participants engaged in workshops using data-based astrophysical 3D models delivered via computer-based interactions, virtual reality, and 3D prints. Multivariate ANOVAs yielded significantly increased STEM interest but were not significant for increasing spatial ability. The results are discussed in terms of the need to extend exposure and science communications to STEM activities to female youths that are younger than middle school aged.

Keywords

spatial reasoning, 3D printing, women in STEM, astronomy, VR

Putting the Stars Within Reach Using NASA 3D Data-Based Models

1. Introduction

This study examined the effects of data-based astrophysical 3D models delivered via computer based interactions, virtual reality, and 3D prints, on spatial reasoning skills and interest in science, technology, engineering, and mathematics (STEM) for females aged 9-12, in particular from traditionally underrepresented groups in STEM. Underrepresented, or underserved, audiences refer to the demographic status of, and the services that are offered or presented to, segments of a community, typically not currently being served within a larger population that might benefit from such services (Williams et al., 2009). Research to date has not focused extensively on the development of STEM interests and spatial reasoning skills of young females, particularly at the time when such learners are forming potential identities in or with STEM and beginning to think about educational and career-related options. The literature review provides a brief background on the development of data visualizations in astrophysics used in the study, issues of STEM interest and identity, and biases and role models, before focusing on spatial skills investigated in the study for the target demographic group.

Literature Review

3D Data in Science Communication

Astrophysical data, and the visualizations from such data, are often two-dimensional (2D). Our perspective from Earth, including from our farthest-reaching telescopes, primarily offers a flat projection on the sky, but spectral information can restore the 3-dimensionality of the Universe (Ferrand et al., 2016). Researchers from computer scientists to astrophysicists, engineers, and technicians are developing multiple techniques to expand astrophysical data visualization within the 2D space and beyond into 3D often for scientific research purposes, but also to help improve communications of such research (see e.g., Arcand et al., 2017b, 2018, 2019; Diemer & Facio, 2017; Goodman et al., 2009; Madura et al., 2015, 2017; Steffen et al., 2014). This work, however, is still relatively nascent, particularly regarding its impact on specific audiences.

The potential to study celestial objects from multiple viewpoints not only can provide researchers with a better understanding of such but also has the potential help open access to and enable science communication with non-experts (Arcand et al., 2017b, 2020; Ferrand & Warren, 2018). There are data-driven 3D objects in 3D printed or extended reality (XR) forms in space science under development that have been or can be used in science communications including, for example, 3D printed supernova remnants (Arcand et al., 2017; Arcand & Watzke, 2019), binary stars (two stars orbiting each other) (Watzke & Edmonds, 2017), lunar geological maps (Ellison, 2014), Martian craters and meteorites (Capraro & NASA JPL, 2014; Gwinner et al., 2014), other local planetary bodies such as asteroids (e.g., Kim, 2018), and topics as the Cosmic Web (Diemer & Facio, 2017) and the Cosmic Microwave Background (Clements et al., 2016). Astrophysical VR applications for use in science communication have been based in simulated worlds (see e.g., Farr et al., 2009) and standalone applications (see e.g., Arcand et al., 2018; Ferrand & Warren, 2018; Russell, 2020). Non-experts can, for example, travel across parts of the Martian surface (Good, 2017), walk on the Sun (Hinode Science Center at NAOJ, 2018), or work with radio data cubes of a spiral galaxy (Ferrand et al., 2016). As noted previously, potential impacts of such 3D or XR assets on specific audiences like female-identifying youths has not been studied in detail, though there is some indication through informal observations and audience demographics of the potential of these mediums using astronomical data via XR (Ferrand & Warren, 2018) and 3D printing (Arcand et al., 2017).

STEM Interest, Identity & Self-efficacy

Separately from the growth of technical achievements in processing astronomical data, there are numerous reasons why gender parity in astronomy in STEM, is important to achieve. Practical issues such as equal opportunity and pay, along with job security for women in STEM pathways, are key, especially as career opportunities in STEM fields increase (Langdon et al., 2011). Financial security can be important for the economic viability of a family (see Anderson, 2016), particularly as single-mother families increase (U.S. Census Bureau, 2016). Importantly, developing strong critical thinking skills so vital in today's world (Duran & Sendag, 2012) are part of STEM subjects. Women tend to make many of the decisions on health and wellbeing for themselves or their families (Marincola, 2006). These personal consumer choices, voting behaviors, investments, and other decisions affect wider economics, politics, finance, and other areas. This argues for women to be well versed in skills, especially those associated with critical thinking and reasoning, associated with STEM fields. The loss—or absence—of women in STEM has quantifiable effects on economies, as well as effects on the "systems we create", from scientific policy to consumer products (Reilly et al., 2019; Wade & Zaringhalam, 2019, para.4).

STEM interest is a critical component of developing a STEM identity and can be intertwined with issues of confidence and self-efficacy for young females. Self-efficacy is often lower in females than males generally (Gnilka & Novakovic, 2017), and related self-efficacy issues of belonging and confidence (Settles, 2014) can therefore be difficult to navigate in formulating the STEM identity. Research has shown that by as young as kindergarten age, girls believe boys are more brilliant and suited to highly intellectual activities over girls (Bian et al., 2017). Negative stereotypes and implicit biases, particularly as unconsciously reinforced by parents, caretakers, and educators, can affect girls' self-perception in STEM topics (Levine, et al., 2015). Lack of confidence in themselves as participants in STEM, or more generally feeling inadequate, can lead to feelings associated with what has been characterized as imposter syndrome (Heaverlo et al., 2013), which affects adult women (and men) in STEM, and particularly affects minoritized populations (Byars-Winston & Dahlberg, 2019). This relates to Bandura's theory of self-efficacy, which concerns one's belief in one's capabilities to perform tasks in specific domains to bring goals to fruition (Bandura, 1997). Self-efficacy can influence what choices and effort an individual will make to reach goals, and also, importantly, if the individual will persist through difficulties and challenges to obtain those goals (Bandura, 1997; Rittmayer & Beier, 2008). Success in STEM areas requires a strong sense of one's ability to succeed, and it is noted that females can have lower senses of self-efficacy towards STEM than males (Williams & George-Jackson, 2014).

Younger women and underrepresented groups may find that if they are not fitting a more typical STEM profile by middle school, then STEM careers might not look viable or feasible (Hill, 2019).

Furthermore, there is an often referenced “STEM pipeline” which, perhaps inadvertently, asserts a linear path from childhood interest in STEM, straight through to taking STEM courses at university, to finding work in a STEM career (The STEM Pipeline, 2015). In fact, this pipeline applies to only about 50% of STEM workers, and further can serve to discourage diversification of the STEM workforce population (Cannady et al., 2014) by unintentionally propounding a one size fits all.

A personal STEM identity is the belief that one can do well and succeed at STEM subjects (Ayoub, 2017). The social aspect of the STEM identity is where participants can visualize themselves “accepted as a member of a STEM discipline or field” (Kim et al., 2018, p. 3). There may be several factors related to STEM identities that make it challenging, particularly for underserved groups like young women of color, to adopt such an identity for themselves (Steinke, 2017). Negotiating an identity as a girl or woman while also adopting an identity in STEM can be challenging, though also eventually beneficial (Settles, 2014).

The challenges to reaching such an identity, however, may include not being able to picture oneself working in STEM, particularly for underrepresented groups (Chee, 2018). Additionally, the socio-cultural cues from others, such as those who seem to indicate that STEM is not cool and, therefore, is less acceptable among peers (Steinke, 2017), or that women do not belong, particularly women of color (Kim et al., 2018), can be challenging. Importantly, recent research has demonstrated that through interventions, STEM identity challenges can be positively influenced (Kim et al., 2018).

Biases, Role Models and Stereotypes

Multiple studies have shown that STEM skills are learned (Hill et al., 2010) and gender bias still has a negative impact on girls’ noted interest in STEM fields from a very early age. Work from Spencer et al. (1999), for example, demonstrated the strength of stereotype threat for girls in mathematics testing among top math performers, where even the mention of gender bias in a test led to reduced performance for female participants. Group stereotypes (whether based on gender or race) “can threaten how students evaluate themselves, which then alters academic identity and intellectual performance” and can affect “members of any group about whom negative stereotypes exist” (American Psychological Association, 2006, p. 1). Gender bias not only supports the perception that science and mathematics are for boys, but more importantly, that stereotype threat can lead to girls’ performance anxiety and low expectations on academic tests (Doyle, 2016).

A growing body of research has shown that there are significant sociocultural roadblocks for women in both the computer science and engineering fields. These barriers can greatly hinder determination of selection and study (Ceci et al., 2009; Cheryan et al., 2015). A meta-analysis conducted by Cheryan et al. (2017) of 1,200 papers on gender gaps in computer science, engineering, and physics, pointed to three primary issues. These included a masculine culture, which included three sub-issues of incompatibility stereotypes, negative stereotypes, and lack of role models, that can lead to women feeling less welcome in those fields; less early exposure or work within those fields; and reduced self-efficacy. The researchers noted that STEM programs could benefit by addressing problematic cultural issues as well as by building awareness in girls and boys that both are equally capable of succeeding in STEM careers.

Spatial Skills

Focusing down into specific issues in STEM skills, mental manipulation and understanding of 2D or 3D objects has been posited as an important STEM skill, helping to indicate future mathematical success (Hegarty & Waller, 2005; Verdine et al., 2014), science performance (Uttal et al., 2013), and potential pursuit of STEM careers (Uttal & Cohen, 2012). Middle school students with good mental rotation skills particularly tend to perform well in their science classes (Ganley et al., 2014). And Rafi et al. (2005) noted that early childhood activities involving paper folding, unfolding, and other manipulations make important contributions towards enhancing spatial skills.

Research back over three decades has indicated that differences in spatial reasoning between males and females often show reduced or delayed development in females (Linn & Petersen, 1985; Maccoby & Jacklin, 1974; Voyer et al., 1995; Yilmaz, 2009). A meta-analysis of gender differences across 14 studies published between 1975 and 1992 supported that claim, with females shown to be at a statistically significant disadvantage in mental rotation skills in each of the studies (Masters & Sanders, 1993). Recent research has emerged indicating that such spatial skills can be improved through interventions (Gold et al., 2018; Hwang et al., 2009; Uttal et al., 2013; Yeh, 2007), and likely through interventions in virtual environments (Hwang & Hu, 2013). Additionally, 3D or XR experiences may support such learning by providing a hands-on approach and making less abstract such large, abstract and highly spatial concepts as astrophysics through embodied cognition (Cole, Cohen, Wilhelm & Lindell, 2018; Pomerantz, 2019). Those findings underpinned the motivation for the research conducted for this study.

Some literature has suggested an optimal age for spatial skills development would be in early middle school years, that is, ages 11-12 (Ben-Chaim, 1989). Yet, no definitive research was found on the lower limits of total time needed for such interventions, nor specific research particularly targeting underserved young females. The research in this study set out to explore such variables.

Conceptual Framework and Research Question

As this study investigated the quantifiable effects of data-based 3D models and prints on spatial reasoning skills and interest in STEM fields for young females, particularly for those from traditionally underrepresented groups in STEM, it is important to consider young females' participation in STEM activities, the development of spatial reasoning skills, and the presence of encouraging role models, to open potential future pathways in STEM studies.

This study involved working directly with young females to investigate the quantifiable effects of data-based 3D models and prints on improving spatial reasoning skills and interest in STEM fields in under-represented groups, particularly of young female learners. The approach used in this research was pragmatism (Johnson & Onwuegbuzie, 2004), as it allows that there is a subjective reality or "real world," although within that subjective reality there are different world views or perspectives (Johnson & Onwuegbuzie, 2004). Pragmatism, therefore, frames the research question with an eye towards "workability," while understanding that "our values and our politics are always a part of who we are and how we act" (Morgan, 2007, p. 70).

2. Methods and Procedures

2.1 Participants

In the study, a convenience sample of 117 youths aged 9-12 years from female-supporting organizations participated, recruited from local networks known to the researcher: a local girls day camp group (Camp Group = 29), a local private all-girls school (Private School = 59), and local girls scouting organization (Scout Group = 29) in the northeastern region of the United States.

2.2 Materials

There were three instruments used in the study, a demographics questionnaire, a spatial skills evaluation, and a STEM interest inventory. The science materials used in the workshops are described in the procedure, with links provided to the content.

2.3 Demographics Inventory

The demographics inventory developed for study included items for grade, gender, ethnicity, zip code, activity location, self-rated knowledge of STEM, and hours of video games played per day. Age, grade, gender, ethnicity information, zip code, and location were collected as background, for analyses, and to ensure the sample groups represented the targeted population. Self-rating information and estimates of how many hours of video games played per day were collected to later assist with the examination of the data in the spatial skills and inventory responses.

2.4 Spatial Skills Inventory

The spatial skills instrument used was the freely available "Spatial Reasoning Test | 123test.com," (123test, 2019). In pilot testing (n = 4) of comparably aged females, it was determined that using all 10 items from the original instrument was too lengthy for the selected age group, and items 6 through 10 became more difficult for the participants. The evaluation was reduced, therefore, to the first five items.

2.5 STEM Interest Inventory

The STEM interest inventory was the "Instruments for Assessing Interest in STEM Content and Careers" (Tyler-Wood et al., 2010), validated with middle school student participants, and used with permission of the authors. The inventory comprises five opposite word pairs repeated across each topic of science, math, engineering, technology, and STEM careers, with selections across a 7-point scale.

Procedure

The science content used in the study applied information gained from researching the data processing and 3D modeling pipeline. The order of the activities was chosen to build participants towards a knowledge base of a topic not typical in content for children aged 9-12 in the U.S. Care was taken to provide the underpinnings of how the data were collected, what scientists do to the data, and why.

This study was approved by a university ethics committee. Participation was anonymous and the data were aggregated, with no personally identifiable information stored.

The workshops were conducted between February-October, 2019 in concrete sessions of half (2.5 hours) or full (6 hours) days, as determined by the typical meeting time for the group involved. The times listed for the workshops included all activities, from start to finish. This study used a controlled within groups design. Participants were randomly assigned to group A or group B to counterbalance the order for the pre- and post-conditions. Identification numbers were used to ensure that the pre- and post-evaluations were matched.

Group A completed the spatial skills instrument as a pre-evaluation, engaged in the workshop, and then completed the STEM interest inventory as a post-evaluation. Group B completed the STEM interest inventory as a pre-evaluation, engaged in the workshop, and then completed the spatial skills instrument as the post-evaluation. Both Groups A and B completed the demographics inventory as part of the pre-evaluations.

Materials were created and produced by the researcher and her team for NASA's Chandra X-ray Observatory. Materials provided multiple modes of accessibility. VR utilized audio tracks and closed captioning, tactile 3D prints of data sets were provided, videos included audio and text transcripts, images were tagged for screen readers, and digital files were coded to U.S. accessibility standards. Activities were developed to describe and interactively show the kind of work the researcher and others do using astronomical data. Materials were developed with funding from NASA and were required to pass NASA product review, a rigorous internal quality control system from NASA to ensure accuracy of materials.

The half day event was conducted with the researcher and one additional female facilitator who was trained by the researcher. Full-day events were conducted with the researcher and three to four additional trained female facilitators or teacher helpers. Facilitators were current workers in STEM fields. The Private School group had a surplus of volunteer teachers to help throughout the days, and who were trained with a 60-minute webinar prior to the event. The schedule for the workshops follows:

1) Introduction

The pre-evaluation was administered to all participants, followed by an introduction for the science content. The warm-up activity outlined how controllers talk to spacecraft orbiting beyond Earth.

Participants created binary code name tags, translating their names into 1’s and 0’s (http://chandra.si.edu/binary/)

2) Imaging

In the half days, this covered the basics on how to create an image, followed by questions and answers. In the full days, participants completed a computer-based activity using Pencil Code and videos that step through the mechanics of compositing/colorizing NASA data to create an image of stars (http://chandra.si.edu/code/).

3) 3D modeling/printing

In full days, participants used browser-based software Tinkercad to learn 3D modeling before loading 3D NASA data. Users discussed 3D printing and handled 3D prints. Half days had a discussion of the modeling/printing processes in which they handled 3D prints for questions and answers. (http://chandra.si.edu/tinkercad/ and http://chandra.si.edu/3dprint).

4) VR

This segment began with a presentation on creating NASA data-driven VR and what the use of such applications can mean before moving into spatially-aware demonstrations of astronomical VR including supernova remnants and the area around the Milky Way’s supermassive black hole (http://chandra.si.edu/vr).

5) Paper activities

Paper-based activities included origami (http://chandra.si.edu/origami) and paper circuits (http://chandra.si.edu/make/) to help communicate science data concepts (e.g., expansion and unpacking, or neutron stars).

6) Wrap-up

Post-evaluation assessments were presented with a question-and-answer session on science and careers. NASA handouts showcasing women in STEM (http://chandra.si.edu/women/) and handouts on each activity were given out for reflection and sharing.

Informal Observations

Beyond the formal instruments, informal observations were recorded by the researcher in the form of field notes. Though not formal analyses, these observations of participant behaviors and comments were captured throughout the day on the researcher’s iPad and are threaded into the results as supplemental interpretations where applicable.

3. Result

Workshops

There were four workshops, with workshop times and lengths considered in the analyses by examining results for the different groups. The first workshop was a half-day, special event offered at a local Scout

Group headquarters for 4th, 5th, and 6th grade girls (n = 29, aged 9-11). Participants, with their caregivers, self-selected to register individually or as a group of Scouts. Fourteen girls had to leave early due to poor weather and snow. This reduced the total number of completed pre-post sets of data to 15.

The second workshop was conducted as a single full-day workshop hosted at a local college, organized with the Camp Group as part of their summer day-camp program for socio-economically disadvantaged girls. There were 29 participants aged 10-12. Two participants had to leave the workshop early, bringing the total of completed pre-post sets to 27.

The third and fourth workshops took place at a local private girls' school with a "STEAM" (science, technology, engineering, art, and math) center for 59 girls, comprising their entire 4th, 5th, and 6th grade classes (aged 9-12). One 6th grader had to leave early, bringing the total of completed pre-post sets to 58.

Descriptive Data

The data collected were digitized and entered into SPSS version 26 for analysis (IBM Corp., 2019). Results for the full sample (n = 100) and broken down by affiliation are provided in Tables 1 and 2.

As shown, participants in the full group were evenly spread across age and grade, in a narrow band from ages 9 to 12. Participants identified as female with one exception, who self-identified as non-binary. In terms of affiliation, the Camp Group was the only group comprised completely of girls of color (with one participant who chose not to respond to this item).

As a result of the poor weather during one workshop (described above), there were slightly more participants in the spatial inventory pre-STEM inventory post sample (n = 55) than the STEM inventory pre-spatial inventory post sample (n = 45).

All participants were asked to rate themselves on their overall knowledge in STEM topics. On a scale of 1 (low) to 10 (high), the mean ratings indicated a wide range of knowledge for the full sample, with the average rating approaching 7 (see Table 2). When examined by affiliation, the Camp Group participants were somewhat below the other two groups. A one-way analysis of variance determined that these differences were not statistically significant, $F(2, 97) = 2.03$, ns. Although most participants (n = 86; 86%) reported spending 0-2 hours playing video games per day ($M = 1.16$, $SD = 1.68$), there was some confusion vocalized by the girls in this group and noted by the researcher as to what "counted" as playing a video game (e.g., Candy Crush vs. Grand Theft Auto). Thus, this variable was only used in a correlational analysis and not in the primary analyses for investigating the research question.

The scores for the spatial and STEM inventories are presented in Table 3. The total score for the spatial inventory, based on 0 (lowest possible score) to 5 (highest possible score) indicated that the Private School participants had the highest values for the spatial score totals, followed by the Camp Group, and then the Scout Group. For the STEM inventories, the minimum score possible was 0 and the maximum score possible was 135. For the STEM inventory score totals, the Private School placed the highest again, followed by the Scout Group, and then the Camp Group.

Principal Component Analysis

Next, a Principal Component Analysis with a direct oblimin rotation was run on the items of the STEM inventory to examine their factor structure (data are available). A direct oblimin rotation was chosen as any anticipated underlying factors could be correlated. The structure matched what was expected for the inventory, with the items for each subscale loading into an individual factor, aligning with the published results (Tyler-Wood et al., 2010). Based on this analysis, the items for each category were combined into their expected subscales.

Each subscale was then examined in terms of reliabilities, using Cronbach's Alpha. Results are provided in Table 4, along with the published reliabilities for the STEM Inventory (Tyler-Wood et al., 2010). As shown, in each case the data from the participants on the STEM inventory subscales were the same or similar to the published coefficient alphas. Reliabilities were calculated for the Spatial inventory, also shown below. The value reported is generally considered to be at the minimum level of acceptability (Cortina, 1993). As there were no published values for this inventory, no comparison could be made to previous data.

Having established that the reliabilities of the inventories meant that the scales were viable, the researcher proceeded with analyses using the total scores for the Spatial and STEM inventories. Correlations were computed to examine the relationships of hours playing video games with self-rating of STEM knowledge, the Spatial total scores, and the STEM inventory total score, to investigate the extent of any relationship among these variables that might help inform STEM identity. The correlations were done first for the full sample and then by affiliation (full data are available).

For the full sample, video hours were negatively correlated with self-rating of STEM knowledge and the STEM inventory total score, thus suggesting that more hours playing video games were related to lower values for self-rating of STEM knowledge and the STEM inventory total score. Of note is the positive correlation between perception of self-rating of STEM knowledge and the STEM inventory total score.

When the correlation analysis was rerun by affiliation (data are available), only the results for Private School yielded significant relationships. For the full sample, video hours were negatively correlated with self-rating of STEM knowledge, and the STEM inventory total score, suggesting that more hours playing video games were related to lower values for self-rating of STEM knowledge and the STEM inventory total score. Again, for the full sample there was also a positive correlation between self-rating of STEM knowledge and the STEM inventory total score for the Private School sample.

Multivariate Analysis of Variance (MANOVA) Results

These analyses laid the groundwork to allow the examination of research question concerning the quantifiable effects of databased 3D models of on under-represented populations in STEM, using a multivariate analysis of variance (MANOVA). The independent variables in the MANOVA were the three affiliations (Scout Group, Camp Group, and Private School) and two levels of the order of the presentation of the inventories (Spatial inventory pre and STEM inventory post or STEM inventory pre and Spatial inventory post). The two dependent variables were the total scores for the Spatial and STEM inventories. Self-rating of STEM knowledge was used as a covariate to eliminate the possibility of any potential relationship between perceptions of preexisting knowledge and the dependent variables, especially the STEM inventory. Therefore, using this covariate essentially equalized the baseline going into the study in terms of where the participants thought they were in terms of ability, so as not to obscure any potential effects.

The results of the MANOVA (see Table 5) yielded two significant findings. Affiliation was significant, as the pre-post condition. There was no interaction effect of affiliation by pre-post condition.

Univariate Results

The univariate results are shown on Table 6. The descriptive statistics for the multivariate test are shown previously on Table 5 and are presented graphically at the end of the paper.

Figure 1 shows the mean pre-post scores for the Spatial inventory, with separate lines for each affiliation. All three groups did less well on average in the post-evaluation condition, and the participants from the Camp Group were well below the other two groups.

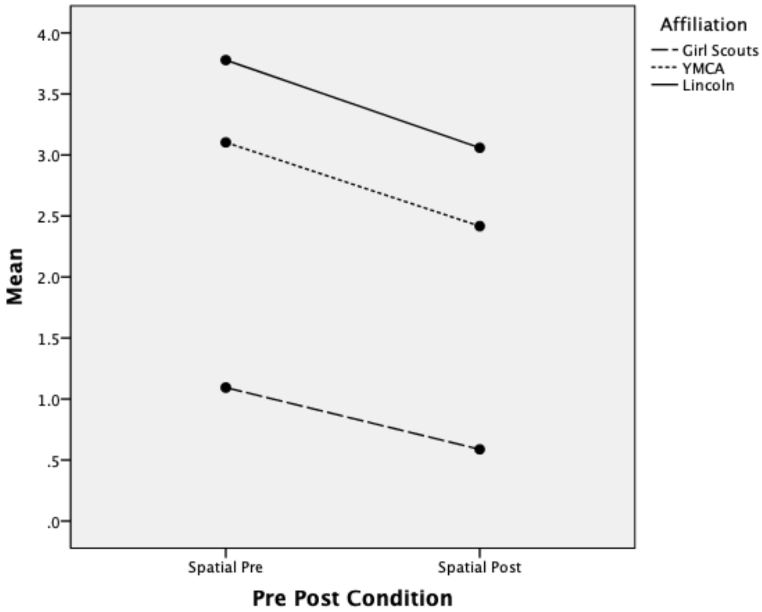

Figure 1. Mean Pre-Post Scores for Spatial Inventory

Figure 2 shows the mean pre-post scores for the STEM inventory, with separate lines for each affiliation. The Private School participants did not exhibit much change from pre- to post, but the other two groups increased nearly 20 points for the Scout Group and 25 points for the Camp Group participants.

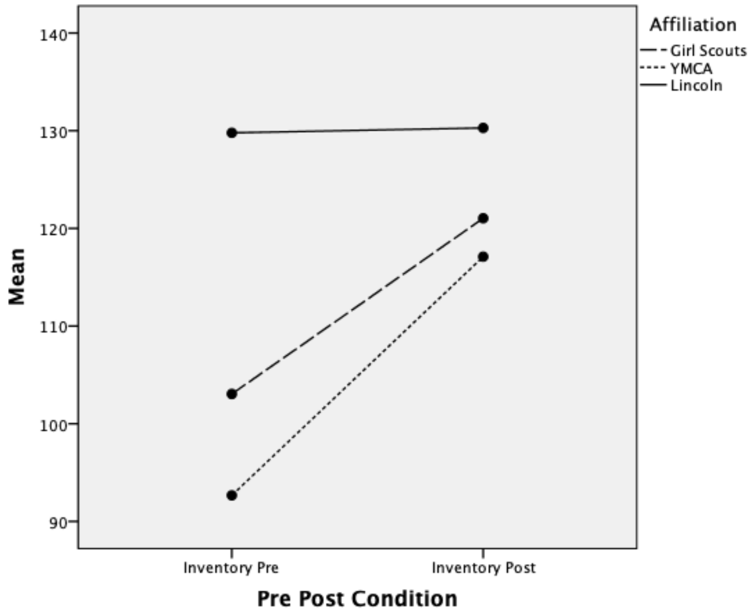

Figure 2. Mean Pre-Post Scores for STEM Inventory

Informal Observations

Informal observations included recording general field notes on participation, vocalizations and energy levels, comments made by the participants to the researcher throughout the day, questions or concerns raised during the activities, as well as summations from the participants volunteered during the wrap up sessions. Comments and questions recorded by the researcher in the field notes were tallied and coded per group and calculated per session as a percentage over time.

Comments throughout the workshops included: "I want to do this when I grow up," "I didn't know I could do this," "This was the best day ever," and "I didn't know I could walk around a black hole." The highest percentage of comments spoken directly to the researcher were from the Camp group; and the Scouts group directed the least number of verbal comments to the researcher in total and as a percentage of time (note the Scouts group had the least time with the researcher overall).

Informal end-of-day dialogues resulted in discussions around activities that the participants enjoyed, concepts learned, and areas to explore. One common theme expressed to the researcher across all groups was enjoyment of the VR and 3D printing (backed up by the recorded observations of the individual activities), though the girls spoke positively about each activity. Questions relating to the researcher's career were numerous, but particularly from the Private School. Science-based questions from black holes to aliens were another common theme across all groups, with the highest frequency from the Private School.

Table 1. Demographic Data for the Full Sample and by Affiliation

Item	Group	n	Percent
Affiliation	Scout Group	15	15
	Camp Group	27	27
	Private School	58	58
Pre-Post Group			
Full Group	Spatial/Inventory	55	55
	Inventory/Spatial	45	45
Scout Group	Spatial/Inventory	7	46.7
	Inventory/Spatial	8	53.3
Camp Group	Spatial/Inventory	16	59.3
	Inventory/Spatial	11	40.7
Private School	Spatial/Inventory	32	55.2
	Inventory/Spatial	6	44.8
Age			
Full Group	9	21	21
	10	25	25
	11	38	38
	12	16	16
Scout Group	9	6	40
	10	8	53.3
	11	1	6.7
	12	--	--
Camp Group	9	--	--
	10	4	14.8
	11	9	33.3
	12	14	51.9
Private School	9	15	25.9
	10	13	22.4
	11	28	38
	12	2	3.4

Item	Group	*n*	Percent
Grade			
Full Group	4	27	27
	5	25	25
	6	32	32
	7	16	16
Gender	Female	99	99
	Nonbinary	1	1
Ethnicity			
Full Group	Caucasian	45	45
	Black/African	11	11
	Hispanic/Latina	9	9
	Other	30	30
	Prefer NtA	5	5
Scout Group	Caucasian	11	73.3
	Black/African	--	--
	Hispanic/Latina	--	--
	Other	2	13.3
	Prefer NtA	2	13.3
Camp Group	Caucasian	--	--
	Black/African	8	29.6
	Hispanic/Latina	8	29.6
	Other	10	37
	Prefer NtA	1	3.7
Private School	Caucasian	34	58.6
	Black/African	3	5.2
	Hispanic/Latina	1	1.7
	Other	18	31
	Prefer NtA	2	3.4
Self-rating			
Full Group	1	2	2
	2	--	--
	3	1	1

Item	Group	*n*	*Percent*
	4	6	6
	5	21	21
	6	19	19
	7	20	20
	8	18	18
	9	7	7
	10	6	6
Scout Group	1	--	--
	2	--	--
	3	--	--
	4	2	13.3
	5	3	20
	6	1	6.7
	7	2	13.3
	8	2	13.3
	9	3	20
	10	2	13.3
Camp Group	1	--	--
	2	--	--
	3	1	3.7
	4	1	3.7
	5	8	29.6
	6	7	25.9
	7	7	25.9
	8	3	11.1
	9	--	--
	10	--	--
Private School	1	2	3.4
	2	--	--
	3	--	--
	4	3	5.2
	5	10	17.2
	6	11	19

	7	11	19
	8	13	22.4
	9	4	6.9
	10	4	6.9

Item	Group	*n*	Percent
Video Hours			
Full Group	0	48	48
	1	24	24
	2	14	14
	3	6	6
	4	2	2
	5	2	2
	6	1	1
	7	2	2
	8	1	1
Scout Group	0	5	33.3
	1	5	33.3
	2	3	20
	3	1	6.7
	4	1	6.7
	5	--	--
	6	--	--
	7	--	--
	8	--	--
Camp Group	0	6	22.2
	1	10	37
	2	5	18.5
	3	1	3.7
	4	1	3.7
	5	1	3.7
	6	1	3.7
	7	1	3.7
	8	1	3.7

Private School	0	37	63.8
	1	9	15.5
	2	6	10.3
	3	4	6.9
	4	--	--
	5	1	1.7
Item	Group	*n*	Percent
	6	--	--
	7	1	1.7
	8	--	--

Table 2. Means and Standard Deviations for Age, Self-rating, and Video Hours for Full Group, and by Affiliation

Item	Mean	Standard Deviation
Age		
Full Group	10.49	1
Scout Group	9.67	0.62
Camp Group	11.37	0.74
Private School	10.29	0.9
Self-rating		
Full Group	6.55	1.81
Scout Group	7.07	2.12
Camp Group	6	1.24
Private School	6.67	1.91
Video Hours		
Full Group	1.16	1.68
Scout Group	1.2	1.21
Camp Group	1.96	2.19
Private School	0.78	1.38

Table 3. Scores for the Spatial Skills and STEM Inventories for the Full Sample and by Pre Post Conditions within Affiliation

Item	Group	*n*	Mean	SD
Spatial Total Score	Full Group	100	2.89	1.6
	Scout Group	15	0.87	1.19
	Camp Group	27	2.78	1.6
Item	Group	*n*	Mean	SD
	Private School	58	3.47	1.23
Inventory Total Score	Full Group	100	121.08	35.5
	Scout Group	15	114.4	32.92
	Camp Group	27	104	33.88
	Private School	58	130.76	33.97
Spatial Pre	Full Group	55	3.24	1.48
	Scout Group	7	1.14	1.35
	Camp Group	16	3.06	1.77
	Private School	32	3.78	0.79
Spatial Post	Full Group	45	2.47	1.66
	Scout Group	8	0.63	1.06
	Camp Group	11	2.36	1.29
	Private School	26	3.08	1.55
Inventory Pre	Full Group	45	116.24	43.28
	Scout Group	8	105.63	40.88
	Camp Group	11	89	42.04
	Private School	26	131.04	39.14
Inventory Post	Full Group	55	125.04	27.36
	Scout Group	7	124.43	18.94
	Camp Group	16	114.31	23.08
	Private School	32	130.53	29.77

Table 4. Reliability Statistics using Cronbach's Alpha for STEM Inventory Subscales and Spatial Skills

Subscale	Cronbach's Alpha	Published Cronbach's Alpha
Science	.91	.84
Math	.91	.88
Subscale	Cronbach's Alpha	Published Cronbach's Alpha
Engineering	.90	.92
Technology	.91	.91
Career	.89	.93
Spatial	.70	--

Table 5. Results of Multivariate Analysis of Variance

Effect	Value	F	Hypostudy *df*	Error *df*	Sig.	Partial Eta Squared
Intercept						
Pillai's Trace	.363	26.26	2	92	<.001	.363
Wilks' Lambda	.637	26.26	2	92	<.001	.363
Hotelling's Trace	.571	26.26	2	92	<.001	.363
Roy's Largest Root	.571	26.26	2	92	<.001	.363
Self-rating						
Pillai's Trace	.110	5.69	2	92	.005	.110
Wilks' Lambda	.890	5.69	2	92	.005	.110
Hotelling's Trace	.124	5.69	2	92	.005	.110

Roy's Largest Root	.124	5.69	2	92	.005	.110
Affiliation						
Pillai's Trace	.428	12.66	4	186	<.001	.363
Wilks' Lambda	.596	13.58	4	184	<.001	.363
Hotelling's Trace	.637	14.50	4	182	<.001	.363
Roy's Largest Root	.566	26.32	2	93	<.001	.363
Effect	Value	F	Hypostudy df	Error df	Sig.	Partial Eta Squared
PRe-Post Condition						
Pillai's Trace	.082	4.11	2	92	.019	.082
Wilks' Lambda	. 918	4.11	2	92	.019	.082
Hotelling's Trace	.089	4.11	2	92	.019	.082
Roy's Largest Root	.089	4.11	2	92	.019	.082
Affiliation*Pre Post						
Pillai's Trace	.030	.71	4	186	.586	.015
Wilks' Lambda	.970	.71	4	184	.586	.015
Hotelling's Trace	.031	.71	4	182	590	.015
Roy's Largest Root	.030	1.41	2	93		.015

Table 6. Results of Univariate Analysis of Variance Between-subjects Effects

Source	Dependent Variable	Type lll Sum of Squares	df	Mean Square	*F*	Sig.	Partial *Eta* Squared
Affiliation	Spatial Total	78.14	2	39.36	22.96	<.001	0.331
	Inventory Total	12464.33	2	6232.17	6.1	0.003	0.116

Pre Post	Spatial Total	7.44	1	7.438	4.34	0.04	0.045
Condition	Inventory Total	3751.09	1	3751.09	3.67	0.058	0.038
Affiliation *	Spatial Total	0.14	2	0.067	0.04	0.961	0.001
Pre Post	Inventory Total	2879.38	2	1439.69	1.41	0.25	0.029
Condition							
Error	Spatial Total	159.39 93	1.71				
	Inventory Total	95069.43	93	1022.25			
Source	Dependent Variables	Type lll Sum of Squares	df	Mean Square	f	Sig.	Partial Eta Squared
Total	Spatial Total	1089	100				
	Inventory Total	1590788	100				
Corrected	Spatial Total	253.79	99				
Total	Inventory Total	124751.36	99				

4. Discussion

Significant Growth in STEM Inventory for Disadvantaged Participants

The mean pre post scores for the STEM inventory showed that the Private School participants did not exhibit much change from the pre- to post evaluations, but the Scout and Camp groups did, with a statistically significant increase of almost 20 points for the Scout Group and 25 points for the Camp Group participants. It might seem counterintuitive to find that the more privileged audience remained relatively stable in their STEM inventory total scores, given that the participants' behavior and observed and recorded comments were highly positive. However, it is useful to note that the private school is a leader in STEAM (Fogarty & Arcand, 2020) learning, with a well-equipped facility with 3D printers, robotics and coding programs, and an excellent staff-to-learner ratio. These participants attended a resource-rich school, and STEAM/learning exposure levels were high, while many, though not all, participants provided zip codes of often wealthy communities, in general. This group of participants came into the workshops strong in STEM interest and left as strong. For them, the workshops may have been more about reinforcing and providing practical examples of STEM topics, careers, role models, etc., than being novel or revelatory regarding STEM pathways.

The significant positive growth for the Camp group firstly, and Scout group secondly, in the STEM inventory, however, gets to the core of the research. The Camp group was the most socio-economically disadvantaged group of the three samples, as determined by Camp selection criteria and the requirements to attend the Camp free of charge. The participants in the program are considered potentially at-risk youth by the Camp. Besides signing up for the overall summer camp program and agreeing to attend the 3D workshop, Camp participants had less individual agency in selecting a day-long STEM activity, as the Camp directors selected the itinerary. This group, however, exhibited the most growth in the STEM inventory overall. As such, it is an example of the potential benefits of even short term (single day) STEM engagement on STEM interest for less socio-economically advantaged groups, and perhaps provides additional evidence regarding reaching groups who are not necessarily self-selecting for STEM activities already.

Next in STEM inventory improvement, the Scout group had more variable individual agency than the Camp group, in that some participants selected to attend the workshop by themselves while others came as part of their chapter’s decision to attend. Overall, the Scout group participants came from mostly and generally middle-class income zip code districts, and was less diverse than either of the other workshop groups. Being members of an out-of-school group likely provided these participants with access to external enrichment programs, but still this half-day program saw a quantifiable effect on the growth of STEM interest.

Lack of Impact of Spatial Reasoning Skills

Although it was hoped that 3D data manipulation activities would yield improvement in spatial reasoning skills from the pre to post groups, ultimately perhaps it was not surprising to see no statistical improvement in the spatial reasoning evaluation. Current literature was not clear on the minimal baseline of interventions required to effect positive change in spatial reasoning skills, and specifically not for this age, gender, and schooling level, although research with older participants tended toward multiple steps of interventions spread over time (see e.g., Contero et al., 2006; Dominguez et al., 2012; Feng et al., 2007; Martín-Gutierrez et al., 2013; Rafi et al., 2005; Wright et al., 2008). The literature, therefore, lacks a baseline for understanding spatial reasoning intervention impacts, particularly for the age and gender groups that participated in this study.

This study, then, adds to the literature on spatial skills improvement for young females, that single, relatively short interventions (6 hours or less) with spatial skills related activities are likely not enough to effect demonstrably positive change. However, effective spatial skills testing instruments for such demographic groups must be reconsidered in approach and methodology before reaching any conclusions in this area, as discussed in the following sections.

It was, however, surprising to see a decrease in the spatial reasoning skills between pre- and post-evaluations, particularly as STEM interest went up as a whole in the study. It is possible that the participants were suffering from frustration with the spatial instrument used (see the next subsections critiquing the evaluation instruments chosen), test fatigue or more broad fatigue from hours of STEM work in a space beyond their more typical daily schedule, or, in the case of the Scouts group, from a late activity after a school day and with the disruption of some classmates leaving early due to poor weather. STEM program evaluation fatigue has been remarked upon generally in U.S. projects (Malyn-Smith et al., 2013). Another more specific study on cognitive fatigue during administration of standardized tests showed that, in Danish public schools as the school day progresses, test scores tended to lower each hour by 0.9 of a standard deviation (Sievertsen et al., 2016). Such results might help explain why there was a decrease in spatial skills demonstrated even during the short Scout event, which was conducted in the evening, ending at 8:00 p.m., and with the younger girls (ages 9 and 10).

It is important, therefore, to reflect on methodology at this point. There was a disparity, and difficulty in completion, of the survey instruments across the groups, with the workshops for the Camp and Scout groups having the most difficulty, as discussed in the next section.

The Importance of Appropriate Instruments for Specific Demographics

Spatial and Inventory Instruments

The spatial skills inventory was the most challenging aspect of the day for all groups, with frustration and fatigue being evident during the exercise. The researcher found it necessary to talk through an example of the mental rotation exercise – not one included on the handout – to the participants, often with much confusion around it even with the demonstration. The researcher and assistant facilitators repeated directions for the evaluation for all groups, and provided noticeably more assistance for the Camp group, followed by slightly less assistance for the Scout group, followed by the least amount for the Private School. It would be useful to compare this data with similarly aged groups across different demographics, including of other genders (male and non-binary).

Previous research has demonstrated that improvements could be made on mental rotation testing for female and male participants (aged 15-22) with attitude-based interventions (Moè, 2009; Moè & Pazzaglia, 2010). In one study, for females, providing positive information on gender before testing (e.g., telling subjects women tend to outperform men) led to improved test results, whereas for males, discussion of task difficulty before testing led to improved results (Moè, 2009). A second study (mean age 17.02) suggested that discussions of effort involved ("anyone can succeed in this task by putting in effort") can improve performance in mental rotation tests (Moè & Pazzaglia, 2010, p. 464). The research suggested, then, that female improvement in spatial skills testing could be directed through targeted encouragement leading to increased self-efficacy. This technique could potentially be applied to future spatial instrument assessments to see if it expands results beyond mental rotation testing to help better capture the crucial data necessary for investigating spatial skills in young female-identifying populations, and particularly those that identify with other groups.

The STEM inventory instrument was less problematic. In each of the workshops, the visual boxing of the STEM inventory questions by topic (science, technology etc.) led some of the younger participants to believe they only needed to select one condition per topic. Others were confused by the switching of the sides for the positive and negative terms. While this is done purposefully to discourage automatic fill-ins without consideration, it was slightly more challenging for the age group in this study. Facilitators clarified directions, asking participants to complete the questions in each topical box.

The Camp and Scout groups needed more assistance and reminders for the instruments overall. The Private School participants tended to socialize their instruments—unprompted by the researcher—seemingly trying as a collective not to leave fellow learners behind. The overall finding supports the literature that argues for “culturally and contextually responsive” instruments in STEM interventions, and particularly those that target underrepresented audiences in STEM (Malyn-Smith et al., 2013, p. 2). Understanding the needs of the specific demographic groups being evaluated, and being sensitive to the perspectives and experiences they may have, should be considered during instrument development, but also during instrument deployment where a real-time, flexible response could be practiced to guide the evaluation process (Malyn-Smith et al., 2013; National Academies of Sciences, Engineering, and Medicine, 2018; University of Birmingham, 2020). As noted earlier, cognitive fatigue could have been a factor in post-evaluation performance. It is also possible that the instruments resembled “high-stakes testing” during formal education, which can be stressful to learners (Malyn-Smith et al., 2013, p. 2).

Instruments that are better situated for collecting such data angled for younger participants who might be underserved in informal STEM engagements (Allen & Peterman, 2019). This result extends the literature of current assessment instruments helping to show that additional research is required to refine existing instruments or develop more learner-specific methodologies that could better capture spatial reasoning skill data or STEM interest data for younger females (ages 9-12) and particularly for socio-economically disadvantaged group subsets within that demographic, who might otherwise experience frustration with existing instruments. In sum, the findings from this study highlight the importance of using or creating age appropriate, and level of learning appropriate ways to be able to evaluate progress.

Additional Findings

There was a significant positive correlation between perception of STEM knowledge and total career score in STEM inventory. This correlation suggests support for the literature regarding self-efficacy and interests in middle school populations (see, e.g., Blotnicky et al., 2018; Fouad & Smith, 1996; Simpkins et al., 2006). However, an alternative explanation for this result is that perhaps young females are drawn towards areas of perceived strength or self-efficacy.

Positive attitudes by learners towards VR technologies have been established for some time (see Mikropoulos et al., 1998). Mantovani (2001) noted the potential for learners’ visualization, making concrete what might seem abstract through other methods of display or interactions, and important for learning that which cannot be easily experienced in the real world. Additionally, VR has the potential to offer avenues of adaptation, collaboration, and evaluation (Mantovani, 2001). VR can also provide critical interactivity in multimedia that may help improve learning function, guiding the user to explore and understand scientific phenomenon (Mayer & Betrancourt, 2005). 3D visualization, therefore, is understood to be an immersive environment that can help solve the "problem of looking at two-dimensional images and requiring participants to understand relationships three-dimensionally" (Bengfort, 2020, para. 12). Though there is also caution to apply that such models should support learners with direction and guidance, particularly for topics such as astrophysics or other similar scientific visualizations that may provide overly complex models (Tversky, Morrison, & Betrancourt, 2002) and require large amounts of information to be processed by the user (Lowe, 2003).

In the case of the workshops presented in this study, traditional, or physical, and computer-based, or 3D, materials for data manipulation were combined so they could work "in concert" (Paranandi, 2002, p. 333) to enrich the learning process and build up to the direct 3D modeling and VR application work, which capped off the final third of the workshops. The observed highly positive response to the 3D/VR segment, therefore, aligns with the literature that supports the use of VR as a communication and education tool (see e.g., Chen, 2006; Pantelidis, 2009), and simultaneously recognizes that much deeper research is required to identify and understand the models and techniques used. Findings from this research suggest that these techniques need to be examined and adapted to facilitate assessment of evaluative methods specific to younger age groups.

Engagement and excitement levels notably increased during the 3D and VR parts of the workshops, and particularly the VR, as recorded in the field notes. If confirmed with more rigorous analysis, this could suggest that applications that provide immersive experiences for meaning making could lead to increased motivation and engagement levels (see Pantelidis, 1995, Pantelidis, 2009; Winn, 1993). The most positively remarked upon part of the workshops involved the 3D and VR activities. The complexity of questions, energy levels and noise increased greatly from those with regard the other activities. Though this result was observational, future programs that involve interactions with young females might want to include more 3D and XR related materials, aligning to current safety standards of a maximum of 15 minutes of VR with 10 minutes of rest afterwards (Hicks, 2018). The results, therefore, provide some observational evidence that the 3D and VR applications were particularly motivating for the participants, supporting the existing literature on this topic, and extending it to the age groups in the study.

Limitations

The researcher acknowledges limitations related to choices made about the study design, including access to only a small portion of the total U.S. geographic area, and a moderate sized database of only 100 participants in a limited age range. Additional limitations of the research include total exposure time, that is, the workshops being single-day events instead of multiple sessions across longer intervals of time and having one workshop of a shorter duration than the others. This research also should not be generalized to different underserved populations, such as those with specific physical disabilities, other underserved groups not included in the research population of this study, or males.

Conclusion

This study showed that interventions that utilize real world data manipulations and 3D applications as part of hands-on activities can be particularly effective for female-identifying youths from underserved groups aged 9-12. Further, interventions shorter than 6 hours in duration for female youths, or that are done in only one session, may not be enough to demonstrably effect spatial skills. Additionally, instrument development that is specialized for these age groups and learning levels is encouraged.

The female-identifying youths in this study participated in STEM activities that provided authentic experiences manipulating actual observational data such as the NASA data used in the study, and they were engaged in such STEM activities with women researchers. There is, however, an issue that needs additional investigation, between groups having privilege and those with less privilege. In comparison with socio-economically disadvantaged females, more privileged female youths might start off higher in interest levels, but then leave just as high, which is encouraging for potential STEM identities and pathways for the ages of the participants in this research.

The availability of the cutting-edge data sets in 3D used in this study that are able to be converted into 3D prints and VR/AR is expanding at a rapid rate. There have been significant improvements in quantity and platform reach even since the data for this study were collected. Such scientific and technological advancements are leading to a larger, expanded pipeline of spatial data modeled into 3D that can potentially further enhance authentic science communication experiences for nonexperts in STEM topics, particularly for underrepresented learners. These also could inform the development of the spatial reasoning skills that are important for STEM. Using 3D data is not a novelty, but rather a tool for communicating and learning, with great potential for expanding scientific knowledge with nonexperts and perhaps to help diversify the potential influx of future researchers. However, it is crucial to recognize that the data and the technology are only one small part of putting the stars within reach for underrepresented young females. People around them, community support, will likely be needed to help drive the most far-reaching, out-of-this-world impact.

Acknowledgement

My deepest thanks to Prof. Lisa Smith, without whom this research topic would never have happened, as well as to Dr. Keryn Pratt for lending me her helpful perspective. I also thank Sara Price for her wonderful help on data digitization and formatting.

I have no known conflict of interest to disclose. This paper was supported by the University of Otago, New Zealand. Materials were developed with funding from NASA under contract NAS8-03060 for the Chandra X-ray Center, with additional support from funding through NASA's Universe of Learning, based upon work supported by NASA under award number NNX16AC65A to the Space Telescope Science Institute, in partnership with Caltech/IPAC, Center for Astrophysics | Harvard & Smithsonian, and the Jet Propulsion Laboratory.

References

123test. (2019). *About 123test*. Retrieved from https://www.123test.com/about_123test/

Allen, S., & Peterman, K. (2019). Evaluating informal STEM education: Issues and challenges in context. In A. C. Fu, A. Kannan, & R. J. Shavelson (Eds.), *Evaluation in informal science, technology, engineering, and mathematics education* (Iss.161, pp. 17-33). New Directions for Evaluation. https://doi.org/10.1002/ev.20354

American Psychological Association. (2006). *Stereotype threat widens achievement gap*. Retrieved from http://www.apa.org/research/action/stereotype.aspx

Anderson, J. (2016). *Quick Figures: Breadwinner mothers by race/ethnicity and state*. (Research Report NO. #Q054. Washington, DC: Institute for Women's Policy Research. Retrieved from https://iwpr.org/publications/breadwinner-mothers-by-raceethnicity-and-state/

Arcand, K., Edmonds, P., & Watzke, M. (2017a). *The tale of Supernova 1987a*. Retrieved from https://chandra.si.edu/deadstar/sn1987a.html

Arcand, K., & Watzke, M. (2016). *3D printing the X-ray universe*. Retrieved from http://chandra.si.edu/3dprint

Arcand, K. K. (2017a). *How to talk to a spacecraft*. Retrieved from https://chandra.si.edu/binary/

Arcand, K. K. (2017b). *[Preliminary data analysis of a recent 3D printed workshop]*. Unpublished raw data.

Arcand, K. K. (2017c). *Universe in 3D: modeling and printing cosmic objects*. Retrieved from http://chandra.si.edu/tinkercad/

Arcand, K. K. (2018a). *Origami Universe*. Retrieved from https://chandra.si.edu/origami/

Arcand, K. K. (2018b). *A Universe of making and doing*. Retrieved from https://chandra.si.edu/make

Arcand, K. K. (2019). *Walking among the stars*. Retrieved from https://chandra.si.edu/vr

Arcand, K. K., Jiang, E., Price, S., Watzke, M., Sgouros, T., & Edmonds, P. (2018). Walking through an exploded star: Rendering supernova remnant Cassiopeia A into virtual reality. *Communicating Astronomy with the Public Journal*, *1*(24), 17-24. Retrieved from https://www.capjournal.org/issues/24/24_17.php

Arcand, K. K., Jubett, A., Watzke, M., Price, S., Williamson, K., & Edmonds, P. (2019). Touching the stars: Improving NASA 3D printed data sets with blind and visually impaired audiences. *JCOM Journal of Science Communication*, *18*(04). https://doi.org/10.22323/2.18040201

Arcand, K.K., Price, S. R., & Watzke, M. (2020). *Holding the cosmos in your hand: Developing 3D modeling and printing pipelines for communication and research* [Manuscript submitted for publication]. Chandra X-ray Observatory, Center for Astrophysics|Harvard & Smithsonian.

Arcand, K., & Watzke, M. (2019). *Tycho's supernova remnant*. Retrieved from https://chandra.si.edu/deadstar/tycho.html

Arcand, K. K., Watzke, M., DePasquale, J., Edmonds, P., & DiVona, K. (2017b). Bringing cosmic objects down to Earth: An overview of 3D modeling and printing in astronomy. *Communicating Astronomy with the Public Journal*, *1*(22), 14-20. Retrieved from https://www.capjournal.org/issues/22/22_14.pdfs

Arcand, K. K., Watzke, M., Rector, T., Levay, Z. G., DePasquale, J., & Smarr, O. (2013). Processing color in astronomical imagery. *Studies in Media and Communication*, *1*(2), 25-34. https://doi.org/10.11114/smc.v1i2.198

Ayoub, C. (2017, May 1). *After-school STEM programs need our help* [Blog post]. Retrieved from https://amysmartgirls.com/afterschool-stem-programs-need-our-help-86a7d7d8b35b

Bandura, A. (1997). *Self-efficacy: The exercise of control* (10th ed.). W. H. Freeman and Company.

Ben-Chaim, D. (1989). Adolescents' ability to communicate spatial information: Analyzing and effecting students' performance. *Educational Studies in Mathematics*, *20*, 121-146.

Bengfort, J. (2020, February 17). Virtual reality advances bring new possibilities to higher education. *EdTech*. Retrieved from https://edtechmagazine.com/higher/article/2020/02/virtual-reality-advances-bring-new-possibilities-higher-education

Bian, L., Leslie, S.-J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389-391. https://doi.org/10.1126/science.aah6524

Blotnicky, K. A., Franz-Odendaal, T., French, F., & Joy, P. (2018). A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career activities on the likelihood of pursuing a STEM career among middle school students. *International Journal of STEM Education*, *5*, Article 22. https://doi.org/10.1186/s40594-018-0118-3

Borkin, M. (2010). *IIC/AstroMed*. Retrieved from http://wayback.archive-it.org/org-935/20170105143005/http://am.iic.harvard.edu/

Borkin, M., Goodman, A. Halle, M., & Alan, D. (2007). Application of medical imaging software to 3D visualization of astronomical data. In R. A. Shaw, F. Hill, & D. J. Bell (Eds.), *Astronomical Data Analysis Software and Systems XVI ASP Conference Series*, Vol. 376 (pp. 621-624). San Francisco, CA: Astronomical Society of the Pacific.

Byars-Winston, A. (2014). Toward a framework for multicultural STEM-focused career interventions. *The Career Development Quarterly*, *62*, 340-357.

Byars-Winston, A., & Dahlberg, M. L. (Eds). (2019). *The science of effective mentorship in STEM*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25568

Cannady, M. A., Greenwald, E., & Harris, K. N. (2014). Problematizing the STEM pipeline metaphor: Is the STEM pipeline metaphor serving our students and the STEM workforce? *Science Education*, *98*(3), 443-460. https://doi.org/10.1002/sce.21108

Capraro, K., & NASA JPL. (2014). *Block Island*. Retrieved from https://nasa3d.arc.nasa.gov/detail/block-island

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, *135*(2), 218-261. https://doi.org/10.1037/a0014412

Chee, C. (2018). *Building STEM identity*. Presentation at 2015 Winter NSF EPSCoR PD/PA/EOD Meeting, Honolulu, HI. Retrieved from https://www.nsf.gov/od/oia/programs/epscor/presentations/PDPA_Winter_2015/STEM_ID.pdf

Chen, C. J. (2006). The design, development and evaluation of a virtual reality based learning environment. *Australasian Journal of Educational Technology*, *22*(1), 39-63. https://doi.org/10.14742/ajet.1306

Cheryan, S., Master, A., & Meltzoff, A. N. (2015). Cultural stereotypes as gatekeepers: Increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in Psychology*, *6*(49), 1-8. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4323745/

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, *143*(1), 1-35. https://doi.org/10.1037/bul0000052

Christian, C. A, Nota, A., Greenfield, P., Grice, N., & Shaheen, N. (2015). You can touch these! Creating 3D tactile representations of Hubble Space Telescope images. *Journal and Review of Astronomy Education and Outreach*, *3*, B33-48.

Clements, D., Sato, S., & Fonseca, A. (2016). Cosmic sculpture: A new way to visualise the cosmic microwave background. *European Journal of Physics*, 38(1). https://doi.org/10.1088/0143-0807/38/1/015601

Cole, M., Cohen, C., Wilhelm, J., Lindell, R. (2018). Spatial thinking in astronomy education research. *Physical Review Physics Education research*, *14*, 010139.

Contero, M., Naya, F., Company, P., & Saorin, J. L. (2006). Learning support tools for developing spatial abilities in engineering design. *International Journal of Engineering Education*, *22*(3), 470-477.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98-104.

DePasquale, J., Arcand, K. K., & Edmonds, P. (2015). High energy vision: Processing X-rays. *Studies in Media and Communication*, *3*(2), 62-71. https://doi.org/10.11114/smc.v3i2.913

Díaz-Merced, W. (2014, September 22). *Making astronomy accessible for the visually impaired* [Blog post]. Retrieved from https://blogs.scientificamerican.com/voices/making-astronomy-accessible-for-the-visually-impaired/

Diemer, B., & Facio, I. (2017). The Fabric of the Universe: Exploring the cosmic web in 3D prints and woven textiles. *Publications of the Astronomical Society of the Pacific*, *129*(975), 1-10. doi:10.1088/1538-3873/aa6a46

Dominguez, M. G., Martín-Gutierrez, J., Gonzalez, C. R., & Corredeaguas, C. M. M. (2012). Methodologies and tools to improve spatial ability. *Procedia-Social and Behavioral Sciences*, *51*, 736-744. https://doi.org/10.1016/j.sbspro.2012.08.233

Doyle, M. (2016, March 10). The truth about why boys and girls need STEM Toys. *The Good Men Project*. Retrieved from https://goodmenproject.com/gender-sexuality/the-truth-about-why-boys-and-girls-need-stem-toys-dg/

Duran, M., & Sendag, S. (2012). A preliminary investigation into critical thinking skills of urban high school students: Role of an IT/STEM program. *Creative Education*, *3*(2), 241-250. https://doi.org/10.4236/ce.2012.32038

Ellison, D. (2014). *Moon nearside farside*. Retrieved from https://nasa3d.arc.nasa.gov/detail/moon-nearside-farside

Farr, W. M., Hut, P., Ames, J., & Johnson, A. (2009). *An experiment in using virtual worlds for scientific visualization on self-gravitating systems*. Retrieved from https://arxiv.org/ftp/arxiv/papers/0905/0905.1066.pdf

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, *18*(10), 850-855.

https://doi.org/10.1111/j.1467-9280.2007.01990.x

Ferrand, G., English, J., & Irani, P. (2016, May-June). *3D visualization of astronomy data cubes using immersive displays*. Paper presented at the CASCA Conference, Winnipeg, Manitoba. Retrieved from http://hci.cs.umanitoba.ca/assets/publication_files/Gilles.pdf

Ferrand, G., & Warren, D. (2018). Engaging the public with supernova and supernova remnant research using virtual reality. *Communicating Astronomy with the Public Journal*, *1*(24), 25-31. Retrieved from https://www.capjournal.org/issues/24/24_25.ph

Fogarty, S., & Arcand, K. (2020, January 10). Making the case for STEAM, not STEM. *Providence Business News*. Retrieved from https://pbn.com/making-the-case-for-steam-not-stem/?bypass=5e174943d326a

Fouad, N. A., & Smith, P. L. (1996). A test of a social cognitive model for middle school students: Math and science. *Journal of Counseling Psychology*, *43*(3), 338-346. https://doi.org/10.1037/0022-0167.43.3.338

Ganley, C., Vasilyeva, M., & Dulaney, A. (2014). Spatial ability mediates the gender difference in middle school students' science performance. *Child Development*, *85*(4), 1419-1432. https://doi.org/10.1111/cdev.12230

Gnilka, P. B., & Novakovic, A. (2017). Gender differences in STEM students' perfectionism, career search self-efficacy, and perception of career barriers. *Journal of Counseling & Development*, *95*(1), 56-66. https://doi.org/10.1002/jcad.12117

Gold, A., Pendergast, P., Ormand, C., Budd, D., Stempien, J., Mueller, K., & Kravitz, K. (2018). Spatial skills in undergraduate students—Influence of gender, motivation, academic training, and childhood play. *Geosphere*, *14*(2), 668-683. https://doi.org/10.1130/ges01494.1

Good, A. (2017). *Take a walk on Mars—In your own living room*. Retrieved from https://www.jpl.nasa.gov/news/news.php?feature=6978

Goodman, A., Rosolowsky, E. W., Borkin, M. A. Foster, J. B., Halle, M. Kauffmann, J., & Pineda, J. E. (2009). A role for self-gravity at multiple length scales in the process of star formation. *Nature*, *457*, 63-66. https://doi.org/10.1038/nature07609

Gwinner, K., Oberst, J., Jaumann, R., & Neukum, G. (2014). *Gale Crater*. Retrieved from https://nasa3d.arc.nasa.gov/detail/gale-crater

Heaverlo, C., Cooper, R., & Laanan, F. S. (2013). STEM development: Predictors for 6th-12th grade girls' interest and confidence in science and math. *Journal of Women and Minorities in Science and Engineering*, *19*(2), 121-142. https://doi.org/10.1615/JWomenMinorScienEng.2013006464

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 121-169). New York, NY:

Cambridge University Press.

Hicks, M. (2018, December 10). When is it safe to buy kids a VR headset? *TechRadar*. Retrieved from https://www.techradar.com/how-to/when-is-it-safe-to-buy-kids-a-vr-headset

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. Retrieved from https://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf

Hill, W. L. (2019, October 2). The myth of the STEM pipeline. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/views/2019/10/02/negative-consequences-pipeline-metaphor-stem-fields-opinion

Hinode Science Center at NAOJ. (2018, July 9). *VR app "Excursion to the Sun" has been released*! Retrieved from http://hinode.nao.ac.jp/en/news/notice/vr-app-excursion-to-the-sun-has-been-released/

Hwang, W. Y., Su, J. H., Huang, Y. M., & Dong, J. J. (2009). A study of multi-representation of geometry problem-solving with virtual manipulatives and whiteboard system. *Educational Technology & Society*, *12*(3), 229-247. Retrieved from https://www.researchgate.net/publication/220374659_A_Study_of_Multi-Representation_of_Geometry_Problem_Solving_with_Virtual_Manipulatives_and_Whiteboard_System

Hwang, W. Y., & Hu, S.-S. (2013). Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving. *Computers & Education*, *62*, 308-319. https://doi.org/10.1016/j.compedu.2012.10.005

IBM Corp. Released. (2019). IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26. https://doi.org/10.3102/0013189X033007014

Kim, A., Sinatra, G., & Seyranian, V. (2018). Developing a STEM identity among young women: A social identity perspective. *Review of Educational Research*, *88*(4), 589-625. https://doi.org/10.3102/0034654318779957

Kim, R. (2018). *Asteroid Vesta*. Retrieved from https://nasa3d.arc.nasa.gov/detail/asteroid-vesta

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). *STEM: Good jobs now and for the future*. Retrieved from http://www.esa.doc.gov/sites/default/files/stemfinalyjuly14_1.pdf

Levine, M., Serio, N., Radaram, B., Chauduri, S., & Talbert, W. (2015). Addressing the STEM gender gap by designing and implementing an educational outreach chemistry camp for middle school girls. *Journal of Chemical Education*, *92*(10), 1639-1644. https://doi.org/10.1021/ed500945g

Linn, M., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*(6), 1479-1498. https://doi.org/10.2307/1130467

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction*, *13*(2), 157-176.

Maccoby, E. E., & Jacklin, C. N. (1974). *Psychology of sex differences*. Stanford University Press.

Madura, T. I. (2017). A case study in astronomical 3-D printing: The mysterious Eta Carinae. *Publications of the Astronomical Society of the Pacific*, *129*(975). https://doi.org./10.1088/1538-3873/129/975/058011

Madura, T., Clementel, N., Gull, T., Kruip, C., & Paardekooper, J. (2015). 3D printing meets computational astrophysics: Deciphering the structure of η Carinae's inner colliding winds. *Monthly Notices of the Royal Astronomical Society*, *449*(4), 3780-3794. https://doi.org/10.1093/mnras/stv422

Malyn-Smith, J., Cedrone, D., Na'im, A., & Supel, J. (2013). *A program director's guide to evaluating STEM education programs: Lessons learned from local, state, and national initiatives*. ITEST Learning Resource Center, Education Development Center, Massachusetts Department of Higher Education, UMass Donahue Institute. Retrieved from http://stelar.edc.org/sites/stelar.edc.org/files/A_Program_Directors_Guide_to_Evaluating_STEM_Eduation_Programs_links_updated.pdf

Mantovani, F. (2001). VR learning: Potential and challenges for the use of 3D environments in education and training. In G. Riva & C. Galimberti (Eds.), *Towards cyberpsychology: Mind, cognitions and society in the Internet age* (pp. 207-226). IOS Press. Retrieved from https://www.semanticscholar.org/paper/VR-Learning-%3A-Potential-and-Challenges-for-the-Use-Mantovani/c7f8d19850f4a88655fe81caea89912dcad66241

Marincola, E. (2006). Why is public science education important? *Journal of Translational Medicine*, *4*(7). https://doi.org/10.1186/1479-5876-4-7

Martín-Gutierrez, J., Navarro Trujillo, R. E., & Acosta-Gonzalez, M. M. (2013). Augmented reality application assistant for spatial ability training. HMD vs computer screen use study. *Procedia-Social and Behavioral Sciences*, *93*, 49-53. https://doi.org/10.1016/j.sbspro.2013.09.150

Masters, M. S., & Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behavior Genetics*, *23*, 337-341.

Mayer, R. E., & Betrancourt, M. (Eds.). (2005). The Animation and Interactivity Principles in Multimedia Learning. In *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.

Mikropoulos, T. A., Chalkidis, A., Katskikis, A., & Emvalotis, A. (1998). Students' attitudes towards educational virtual environments. *Education and Information Technologies*, *3*(2), 137-148. https://doi.org/10.1023/A:1009687025419

Moè, A., & Pazzaglia, F. (2010). Beyond genetics in mental rotation test performance: The power of effort attribution. *Learning and Individual Differences*, *20*(5), 464-468. https://doi.org/10.1016/j.lindif.2010.03.004

Moè, A. (2009). Are males always better than females in mental rotation? Exploring a gender belief explanation. *Learning and Individual Differences*, *19*(1), 21-27. https://doi.org/10.1016/j.lindif.2008.02.002

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, *1*(1), 48-76. https://doi.org/10.1177/2345678906292462

National Academies of Science, Engineering, and Medicine. (2018). *English learners in STEM subjects: Transforming schools, classrooms, and lives*. https://doi.org/10.17226/25182

Pantelidis, V. S. (1995). Reasons to use virtual reality in education. *VR in the Schools*, *1*(1), 9.

Pantelidis, V. S. (2009). Reasons to use virtual reality in education and training courses and a model to determine when to use virtual reality. *Themes in Science and Technology Education*, *2*, 1-2, 59-70. Retrieved from https://www.timtechconsults.com/images/ttcvreducation%20.pdf

Paranandi, M. (2002). An inquiry into computers in design: When cardboard met computer. *In Proceedings of the 7th International Conference on Computer-Aided Architectural Design Research in Asia* (pp. 329-337). Multimedia University Malaysia. Retrieved from https://cumincad.architexturez.net/system/files/pdf/bb35.content.pdf

Pomerantz, J. (2019). XR for Teaching and Learning: Year 2 of the EDUCAUSE/HP Campus of the Future Project. *Educause*. Retrieved from https://www.educause.edu/ecar/research-publications/xr-for-teaching-and-learning-year-2-of-the-educause-hp-campus-of-the-future-project/xr-technologies-for-achieving-learning-goals#7d673cf87e934b6bb509c3ece8d4ea15

Rafi, A., Anuar, K., Samad, A., Hayati, M., & Mahadzir, M. (2005). Improving spatial ability using a Web-based Virtual Environment (WbVE). *Automation in Construction*, *14*(6), 707-715. https://doi.org/10.1016/j.autcon.2004.12.003

Reilly, E. D., Awad, G. H., Kelly, M. M., & Rochlen, A. B. (2019). The relationship among stigma consciousness, perfectionism, and mental health in engaging and retaining STEM women. *Journal of Career Development*, *46*(4), 440-454. https://doi.org/10.1177/0894845318784745

Rittmayer, A. D., & Beier, M. E. (2008). Overview: Self-efficacy in STEM. S*WE-AWE CASEE Overviews*, *1*(12).

Russell, C. (2020, April 13). *Traveling to Our Galactic Center Through Virtual Reality* [Blog post].

Retrieved from https://chandra.si.edu/blog/node/755

Settles, I. H. (2014, October). *Women in STEM: Challenges and determinants of success and well-being*. American Psychological Association. Retrieved from https://www.apa.org/science/about/psa/2014/10/women-stem

Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(10), 2621-2624. doi:10.1073/pnas.1516947113

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: a longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*(1), 70. https://doi.org/10.1037/0012-1649.42.1.70

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4-28. https://doi.org/10.1006/jesp.1998.1373

Steffen, W., Teodoro, M., Madura, T. I., Groh, J. H., Gull, T. R., Mehner, A., Corcoran, M. F., Damineli, A., & Hamaguchi, K. (2014). The three-dimensional structure of the Eta Carinae Homunculus. *MNRAS*, *442*(4), 3316-3328. https://doi.org/10.1093/mnras/stu1088

Steinke, J. (2017). Adolescent girls' STEM identity formation and media images of STEM professionals: Considering the influence of contextual cues. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00716

The STEM Pipeline. (2015). *Inside Higher Ed*. Retrieved from https://www.bu.edu/stem/files/2015/02/The-STEM-Pipeline-booklet1.3-36.pdf

Tyler-Wood, T., Knezek, G., & Christensen, R. (2010). Instruments for Assessing Interest in STEM Content and Careers. *Journal of Technology and Teacher Education*, *18*.

University of Birmingham. (2020). *Evaluating your STEM intervention or activity*. Retrieved from https://www.birmingham.ac.uk/university/colleges/eps/STEM/activities/pedagogic/evaluating.aspx

Uttal, D. H., & Cohen, C.A. (2012). Spatial thinking and STEM education: When, why, and how? In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 52, pp. 148-178). San Diego, CA: Academic Press.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352-402. https://doi.org/10.1037/a0028446

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development*, *85*(3), 1062-1076. https://doi.org/10.1111/cdev.12165

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*(2), 250-270.

Wade, J., & Zaringhalam, M. (2019, August 6). Why we need to keep talking about equality in physics. *Physics World*. Retrieved from https://physicsworld.com/a/why-we-need-to-keep-talking-about-equality-in-physics/

Watzke, M., & Edmonds, P. (2017, September 18). V745 Sco: Two stars, three dimensions, and oodles of energy. *Chandra X-ray Observatory*. Retrieved from http://chandra.si.edu/photo/2017/v745/

Williams, B., Bonnett, E., Jackson, C., & Smith, D. (2009). *Working with underserved and at-risk audiences*. Virginia Cooperative Extension. Retrieved from https://pubs.ext.vt.edu/382/382-512/382-512.html

Williams, M. M., & George-Jackson, C. (2014). Using and Doing Science: Gender, Self Efficacy, and Science Identity of Undergraduate Students in STEM. *Journal of Women and Minorities in Science and Engineering*, *20*. https://doi.org/10.1615/JWomenMinorScienEng.2014004477

Winn, W. (1993). *A conceptual basis for educational applications of virtual reality* (Technical Report TR-93-9). Human Interface Technology Laboratory, University of Washington. Retrieved from http://www.hitl.washington.edu/publications/r-93-9/

Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychonomic Bulletin & Review*, *15*(4), 763-771. https://doi.org/10.3758/pbr.15.4.763

Yeh, A. J. C. (2007). *Knowledge construction of 3D geometry in virtual reality microworlds*. (Doctoral thesis). Queensland University of Technology, Brisbane, Australia.

Yilmaz, H. (2009). On the development and measurement of spatial ability. *International Electronic Journal of Elementary Education*, *1*(2), 83-96. Retrieved from https://eric.ed.gov/?id=EJ1052049