For one or for all?: Survey of educator perceptions of Web Speech-based auditory description in science interactives

Brett L. Fiedler
brett.fiedler@colorado.edu
University of Colorado at Boulder
Boulder, Colorado, USA

Taliesin L. Smith
taliesin.smith@colorado.edu
University of Colorado at Boulder
Boulder, Colorado, USA

Jesse Greenberg
jesse.greenberg@colorado.edu
University of Colorado at Boulder
Boulder, Colorado, USA

Emily B. Moore
emily.moore@colorado.edu
University of Colorado at Boulder
Boulder, Colorado, USA

ABSTRACT

The evolution of Web Speech has increased the ease of development and public availability of auditory description without the use of screen reader software, broadening its exposure to users who may benefit from spoken descriptions. Building off an existing design framework for auditory description of interactive web media, we have designed an optional Voicing feature instantiated in two PhET Interactive Simulations regularly used by students and educators globally. We surveyed over 2000 educators to investigate their perceptions and preferences of the Web Speech-based Voicing feature and its broad appeal and effectiveness for teaching and learning. We find a general approval by educators of the Voicing feature and more moderate statement ratings than expected to the different preset speech levels we presented to them. We find that educators perceive the feature as beneficial both broadly and for specific populations while some acknowledge particular populations for whom it remains ineffective. Lastly, we identify some variance in the perceptions of the feature based on different aspects of the simulation experience.

CCS CONCEPTS

• **Human-centered computing → Accessibility**; *Accessibility technologies*; • **General and reference → Surveys and overviews**.

KEYWORDS

surveys, web speech, educators, auditory description, interactives

ACM Reference Format:
Brett L. Fiedler, Taliesin L. Smith, Jesse Greenberg, and Emily B. Moore. 2022. For one or for all?: Survey of educator perceptions of Web Speech-based auditory description in science interactives. In *19th Web for All Conference (W4A'22), April 25–26, 2022, Lyon, France.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3493612.3520456

W4A'22, April 25–26, 2022, Lyon, France

ACM ISBN 978-1-4503-9170-2/22/04.
https://doi.org/10.1145/3493612.3520456

1 INTRODUCTION

The evolution of modern web APIs such as Web Audio [17] and Web Speech [18] enables rapid and scalable development of auditory display features embedded within web media in ways not previously achievable. In particular, Web Speech, a JavaScript API enabling web page development utilizing a device's speech recognition and synthesis capabilities, can ease the development of built-in accessibility features. Features include simple text-to-speech of on-screen text and auditory description display for multimedia content. Traditionally, auditory or spoken descriptions are provided via screen reader software that can access true text on the screen or true text made available in accessible digital media. While screen reader software has more capabilities than the Web Speech API can afford, it does share some features that can be included as part of a web page in any modern browser. Thus, the Web Speech API introduces new possibilities for who can readily access auditory descriptions of interactive media. The ready access and scalability promote new investigations into how and for whom auditory descriptions can benefit users, along the vein of the Electronic Curb-Cut Effect [10] – accessible feature development that benefits more than the initially designed-for audience.

We designed and developed a Web Speech-based auditory description delivery system (i.e., Voicing) for two popular web-based interactive science simulations for learning. We interviewed diverse learners to verify the potential effectiveness of the feature, and analysis of this design research is in progress. In this study we present a survey investigation into the perceptions of teachers and educators who use the simulations in their practice, for its effectiveness and appeal for themselves and their students."

2 AUDITORY DISPLAY AND PHET INTERACTIVE SIMULATIONS

The PhET Interactive Simulations project (http://phet.colorado.edu) is a research and development project at the University of Colorado Boulder that creates, investigates, and publishes web-based mathematics and science interactives used worldwide. Since 2014, researchers within the PhET project have been designing, developing, and implementing new multimodal features to increase access and inclusion for interactives [11, 15]: including auditory displays (speech and non-speech sounds), haptic and tangible displays [30], customizable visual displays (pan and zoom), and alternative input [16].

Project efforts in auditory display have included research projects focused on the design and development of sound effects and sonification implemented using the Web Audio API [6, 8, 31–34], and text description accessed and read aloud using screen reader software (a feature we call Interactive Description) [24–27] implemented using a novel Parallel DOM architecture [23]. In addition to research findings and technical infrastructure, outcomes have included the publication of 13 simulations with sound effects and sonifications and 10 simulations with Interactive Description (5 of which are accessible using iOS VoiceOver on iPhones and iPads), all available as free and openly-licensed learning resources for educators, students, and parents around the world [20].

Recently, the PhET project has expanded its work in auditory display to include a customizable Voicing feature that uses Web Speech (Web Speech API) to voice simulation information as a user navigates and interacts. The voiced, or spoken, information can complement the visual display – which has minimal to zero text on-screen – during interaction. After initial iterations of design and development of the new Voicing feature in two interactive simulations that was evaluated in a small number of interviews with diverse learners, we wanted to understand more about how educators would perceive the Voicing feature in general, and gain insight into the following questions

- Do educators perceive the Voicing feature as beneficial, and if so, what populations of learners do educators identify to likely benefit from a Voicing feature?
- How does educator preference for Voicing change with the amount of voicing presented?
- How does educator preference for Voicing change based on the interaction design of the simulation?

To do this, we conducted a series of surveys of educators who incorporate and facilitate PhET's interactives in their curriculum. Each survey focused on a different physics interactive. Within each survey, three participant subgroups were each provided with a different variant of the customizable Voicing feature–each variant providing a distinctly different set of available spoken responses (i.e., simulation information spoken aloud when interacting). We analyzed the survey responses for themes related to Voicing aesthetic, usability, and intended audience.

3 AUDITORY DESCRIPTION DISPLAY

Auditory description display for digital resources has progressed across a continuum of visual and interactive contexts. Guidelines and practices, using structures such as alternative text that began with static images [1–3, 5, 29], progressed through pre-recorded moving images (movies and television) [19], and scripted but improvisational contexts such as theater [9]. Modern web technologies have increased the general availability of tools such as Text-to-Speech for any on-screen text for any user on their browser or device on popular application storefronts. Additionally, Machine-learning approaches to descriptions of static content are beginning to see progress for general web accessibility [13, 21, 28, 35]. Descriptions of dynamic digital content, interactive and modifiable by the user, remain a challenge for web technology designers. Aside from the need to anticipate every possible state of the interactive resource that the user could create, there are also technological hurdles that arise between a particular screen reader software, an operating system, the web technology the resource is built upon (e.g., HTML5) and the web browser that renders the resource to a learner.

The PhET Interactive Simulations project published a description design framework for auditory description display for interactive simulations, with possible applications to broader interactive media. Work on this "interactive" description design framework started with Smith's efforts in designing descriptions for non-visual access for highly interactive science simulations as part of their master's thesis [22]. Through their design research, Smith built upon Keane's work on simple interactives [12]. Subsequent description design work iteratively refined a systematic description design framework for designing modular descriptions for complex interactives [26, 27]. This iterative design also led to the novel Parallel DOM architecture [23] and general capabilities for alternative input [16].

We briefly describe the general structures of the interactive description framework to help contextualize the design of the Voicing feature in Section 4. For a complete explanation, the published framework is available [27]. The resulting interactive description and the structure of the Parallel DOM focus on enabling screen reader users to efficiently read, navigate and interact with the simulation, the same as they would use and operate any web page appropriately coded with semantically rich HTML. The design framework juxtaposes two primary description structures to present a user with holistic information about the evolving simulation as they interact: state descriptions and responsive descriptions.

State descriptions are available at all times to be read through and include descriptions of both static and dynamic content. State descriptions, accessible via the "browse" or "read" modes on most screen reader software, include a summary of the visual display and information about the current state of all objects and controls. The information is organized with headings, brief paragraphs, and bullet lists. The information updates silently as the user interacts and the simulation changes, allowing them to return at any time to hear updated descriptions.

In contrast, Responsive Descriptions are read out in direct response to a user's interactions and include descriptions of both the focused object and changes to other objects not focused by the user that are part of the surrounding context. They are delivered automatically when the user navigates to or interacts with interactive objects in the simulation. Interactive objects include standard web UI components such as buttons, checkboxes, sliders, and customized interactions such as draggable objects or bi-manual input. Responsive descriptions are designed to alert the user to relevant change–to objects and any other parts of the simulation as they happen. These responses include changes to objects and content outside the user's keyboard focus.

4 VOICING DESCRIPTION: DESIGN AND IMPLEMENTATION

We will briefly explain some details of the Voicing feature to provide context for this study, limited to the design and implementation as it relates to the simulations as presented to educators in the survey. However, at the time of writing, more exhaustive, mature details of the feature and its design evolution based on interviews with diverse learners are in preparation and will be published at a later time.

Briefly, Voicing in the PhET simulations is optional and intended to serve as a user-customized auditory description display capable of serving a broad range of needs. It is designed to deliver voiced information only in response to or as a direct result of user interaction. Once learners turn the feature on, they may choose the amount of information they hear. They can choose more information by 1.) enabling automatic delivery of responses describing initial object details, changes to those objects, or changes to other parts of the simulation during interaction; 2.) by enabling automatic delivery of available help text for interactive objects; 3.) by directly requesting a description of an overview, the current details of the current state of the simulation, or a suggestive hint on what to do next, all via a series of play buttons presented in a toolbar to the side of the simulation; or 4.) can quickly silence or reactivate all automatic Voicing of responses that happen when directly interacting with the simulations objects and the visual text (if any) presented as part of the simulations visual design.

An example scenario, with minimal spoken description, could be a sighted learner who enjoys or benefits from hearing changes to the simulation as they interact that they may otherwise miss while focused on the object they are moving. Alternatively, the same learner may prefer to hear about the latest current details after the fact as a reminder of what happened while making changes. A more extensive description may benefit a low-vision learner, who is pairing auditory description with visual zoom features, and would like to hear details about the object they are changing and any related contextual changes arising from their interaction outside of their visual range.

In this way, the structure of Voicing is analogous, but not identical, to the bifurcated presentation of interactive description into state and responsive descriptions but can be adjusted piecemeal by each individual based on their needs and preferences. However, the descriptions provided by Voicing are somewhat less extensive, are not always identical, nor are they delivered in the same way as those provided by Interactive Description. Interactive Description is designed from the beginning to take advantage of web technologies that support access using screen reader software for a complete and fully interactive non-visual experience.

While Voicing shares many features with traditional screen readers, it differs in crucial ways that may or may not obviate a user's need or desire for a full-screen reader experience with the simulation through Interactive Description. Our Voicing feature does not explicitly describe document structures such as headings and lists or functional behavior such as buttons, checkbox, sliders, etc. Additionally, Voicing is only delivered in response to user interaction, whether that happens from, for instance, changing the state of an object in the simulation or pressing a button to receive a verbal account of the current state of the simulation. Compared to screen reader software, the Voicing feature does not support cursor access to a semantically structured web page-like document that can be efficiently navigated through to read all essential and supplementary details of the current state of the interactive.

Voicing is enabled through the audio tab of the simulation's preferences menu. As shown in Figure 1A, once Voicing is enabled, users can customize the automatic delivery of voicing provided

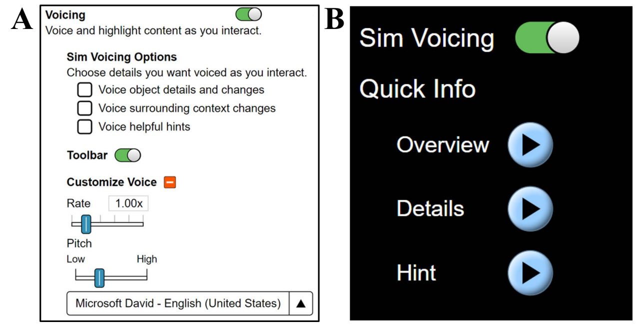

Figure 1: A.) Voicing submenu found within the simulation preference menu, including Voicing toggle, speech output level checkboxes (Sim Voicing Options), toolbar toggle, and voice customization options. Voice customization options are collapsed by default. B.) Interactive components of the collapsible Toolbar that appears on the left side of the simulation after enabling Voicing.

during interaction, displayed as "Voice and highlight content as you interact". Additionally, they may toggle the presence of the 'Toolbar' (Figure 1B), which contains quick access to a Sim Voicing switch that toggles voicing off and on, as well as buttons for on-demand access to the overview, current details, and a hint for interaction. Lastly, there are options to customize the voice, including pitch, rate, and a choice of voices from a list available on the learner's device.

The choice of voicing details (Figure 1A) differ in the types (not necessarily amount) of content a user hears relative to what they are interacting with. With no boxes checked, the feature limits verbal description to any on-screen text that receives intentional focus by the user (i.e., mouse click, touch tap, or keyboard focus), including the Preferences Menu, and the names of any interactive object or control. Checking "Voice object details and changes" verbalizes more details about the interactive object or control that the user is focused on and includes details such as the current value, starting position, or any other information that may help a user understand their direct actions to that element. Checking "Voice surrounding context changes" verbalizes details about changes to the simulation that arise as a result of a user's actions on an object, such as the values and positions of other objects - not under user control - that have are changing, related variables that are changing, or anything else that may help a user understand the indirect consequences of their actions. Checking "Voice helpful hints" adds a brief hint, akin to help text, to each interactive object deemed to need a hint. The delivery and timing of a hint may also depend on the learner's input method and the interaction pattern for the interactive object.

The collapsible Toolbar, Figure 1B, is a collapsible sidebar that appears on the left side of the simulation when Voicing is enabled and contains several buttons related to the Voicing feature. It includes a Sim Voicing toggle that allows a learner to quickly disable or re-enable the speech at any time during their exploration. It also contains three "read-me" or "play" buttons that voice information about the current state of the simulation. An "Overview" button sets the scene and describes what is visually displayed on the screen in playful language. The "Details" button briefly describes the current state of the simulation and updates as the user makes changes to the simulation. The "Hint" button suggests a productive path forward based on the current state of the simulation.

5 METHODS

5.1 Participants and Survey Distributions

Educator feedback was acquired through two surveys delivered through the Qualtrics platform. Each survey contained a single PhET simulation with the Voicing feature. Participants in both surveys were educators who use the PhET Interactive Simulation project website. Visitors to the PhET website can create a user account and opt-in to receive email announcements. During account creation, they can provide information such as role (Teacher, Preservice Teacher, Researcher, Student, etc.), STEM subject specialty, and grade level. The user base was divided into three subsets based on the user's last name entered in the account profile. We emailed an invitation to complete a research survey to the first two subsets (A-G and H-O) of users who selected one or more of the following options: Teacher, Pre-service Teacher, Teacher Educator, Other.

In a second step to confirm participants were educators, an initial survey question asked participants to select their role, and a selection of a non-educator role ended the survey allowing only those selecting educator roles to proceed. The survey was estimated to take about 7 minutes or less to complete for the required sections. No compensation was provided.

The total number of invited participants was 88,858 for Survey 1 (June 2021) and 88,949 for Survey 2 (July 2021). For the required sections of the survey, 1,096 participants completed Survey 1, and 1,212 completed Survey 2. We include partial responses in our statistical analysis.

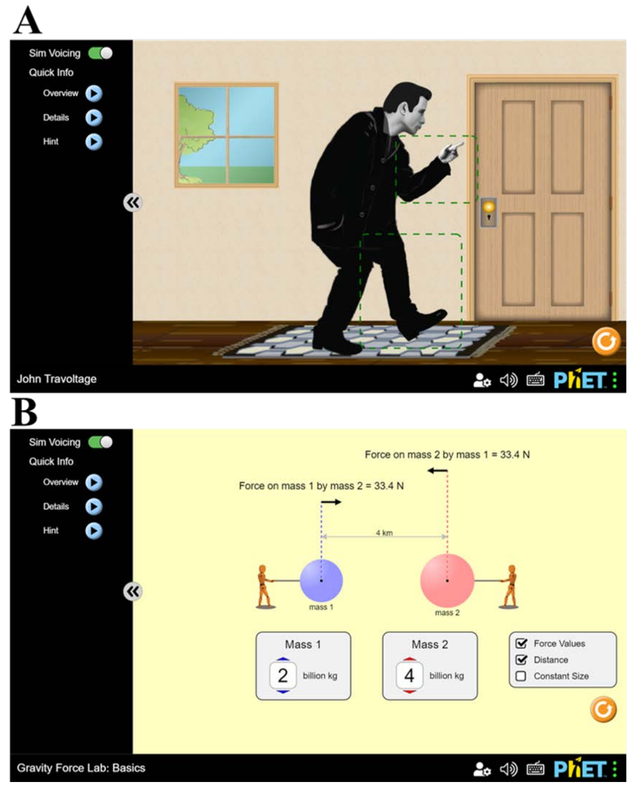

Figure 2: Visual display of A.) John Travoltage and B.) Gravity Force Lab: Basics. Preferences menu is the leftmost icon in the bottom bar (iconic person with cog).

5.2 Simulations

The two surveys distributed differed only by the simulation the educators interacted with. Survey 1 presented the simulation John Travoltage (Figure 2A) and Survey 2 presented the simulation Gravity Force Lab: Basics (Figure 2B), representing the first PhET simulations with the Voicing feature implemented. While Gravity Force Lab: Basics was surveyed second, the timing between surveys was sufficiently short (1 month) and no changes were made to its Voicing content or presentation as a result of the findings of the first survey.

In John Travoltage [33], Figure 2A, a character, John, stands on a rug by a door. Rubbing his foot results in negative charges transferring onto his body, and moving his arm towards the doorknob results in a shock. Learners can explore the relationship between the amount of charge on John's body and the distance between his hand and the doorknob that results in a shock.

Sonifications and sound effects include the sound of the foot rubbing on the rug, a pop sound as negative charges transfer onto John's body, a low continuous hum representing the charges on John's body, a ratchet-like sound when John's arm is moved, and an electrical zap sound as charges are discharged from John's body.

The Voicing feature includes in part: foot and hand positions relative to the rug, John and doorknob, changing amount of electrons on John's body, and amount of electrons discharged during the shock. For example, with all speech enabled, when moving the foot back and forth several times across the rug from its initial position and the arm unchanged, the sim could voice "Leg Swing" as the name of the interactive object, "Foot off rug" and "Foot on rug" as Object changes, and a resulting Context change "Several electrons on body."

In Gravity Force Lab: Basics, Figure 2B, learners can explore the relationships of mass and distance with gravitational force by manipulating the separation distance and individual masses of two mass spheres, held into place by two robot figures. Changes to the separation distance or individual change to the mass of either sphere change the length of the gravitational force arrows centered over each mass sphere, while the arrow direction always points to the other sphere.

Sonifications and sound effects include a tone that varies in pitch with the magnitude of the gravitational force, a percussive sound that changes pitch as mass changes, and collision noises when the spheres are brought together or reach the outer boundaries of the play area.

The Voicing feature includes in part: a read-out of the on-screen text displaying the force values in newtons, separation distance in kilometers, and sphere names and their current mass in billions of kilograms. Changes to separation distance, masses, or force arrows are voiced aloud. For example, with all speech enabled, when moving the blue (leftmost) sphere to the right from the initial position, the sim could voice "4 kilometers from mass 2, move mass 1" as the starting Object value and object name, then "3.3 kilometers from mass 2, Closer, force arrows get bigger. Forces is now 49 newtons" as combined, follow-up Object and Context changes.

Links to the simulation prototypes used by participants in this survey (and other survey data) are available in a public repository (Open Science Framework) [7].

5.3 Study and Survey Design

Investigation of our research questions centered on having educators interact with the Voicing feature by directly embedding the voiced simulation in the survey, and then having them respond to statements directly related to the quantity and quality of spoken content and rate it based on their experience. In this study, we structured the survey and study design on user experience related to the amount of automatic description delivery via interactive responses. However, on-demand description accessed via the Toolbar was also present in both simulations surveyed.

We constrained the default level of speech output (i.e., Sim Voicing) for interactive responses to three presets. Given the high level of customizability and exponential possibilities of Voicing schemes, these presets support conclusions drawn about the amount and types of interactive responses related to participant preference and perception. Participants were randomly assigned to only one of three presets. The first preset, preset A (PA), had both Object Changes and Context Changes checked by default when enabling Voicing. The second preset, preset B (PB), had only Context Changes checked. The third preset, preset C (PC), had no boxes checked and voiced names of interactive objects and on-screen text, which readable when intentionally focused (i.e., mouse click, touch tap, or keyboard focus).

Figure 3 shows a schematic of the survey logic presented to the participants for both survey 1 and survey 2. Participants in either survey were randomly assigned 1 of 3 of the voicing detail presets (PA, PB or PC), but saw the same simulation (John Travoltage or Gravity Force Lab:Basics). The survey included two optional sections. Both optional sections were prefaced with a gating question asking if the participants were willing to give more time to answer questions related specifically to the Toolbar or Preferences Menu. If they selected "No", they continued to the next part of the survey. Analysis of the statements found in these optional sections is not included in this study.

5.3.1 Voicing and Simulation Instructions. Technological limitations prevent Web Speech from being enabled by default on many devices, which required instruction to the participant to enable the Voicing feature prior to interacting with the simulation. Participants were instructed that "You will be asked to change settings in a menu prior to your interaction with the simulation. Please read the directions carefully." On the page containing the embedded simulation, they were directed to "Please enable 'Voicing' and then close the Preferences menu (X)." along with an image of the enabled position of the Voicing toggle. After enabling Voicing, they were instructed to interact with the simulation for 30 seconds (when the next button appeared).

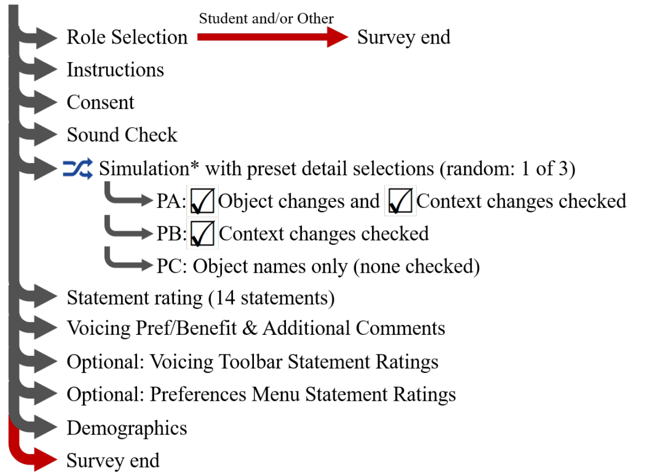

Figure 3: Survey flow chart. *Simulation for survey 1 was John Travoltage, survey 2 was Gravity Force Lab: Basics. Surveys are identical before the Demographics section, except for the presented simulation. Voicing details were preset and randomly presented. Statement order was randomly presented.

5.3.2 Voicing Statements. Participants were presented with 14 Likert-style statements to rate on a 5-point scale from Strongly Disagree to Strongly Agree. Statements were randomly presented and included perceptions and preferences related to the affective experience of Voicing responses in the simulation and the content, length, and performance of the Voicing responses while interacting. All statements are listed as part of Table 1 in Section 6.1.

5.3.3 Voicing Preference and Benefit. In a new section, participants were presented with three additional statements to rate if they thought the Voicing feature would benefit their students, benefit all students, and if they felt the Voicing feature should be available in all PhET simulations. The statements are presented in Table 3 in Section 6.1. A subsequent multiple selection question asked educators to select categories of learners they believed could benefit from using simulations with a Voicing feature like the one they experienced. Selections are presented in Table 4 in Section 6.1.

5.3.4 Additional Comments. In the same section as the prior statements, educators were presented with an optional text entry field to "Please share any additional comments you have." Responses to this question were categorized for the thematic groupings presented in Section 6.2.

5.3.5 Demographics. Educators were asked to provide their age, gender identity, primary level of their students (Elementary to University) and subject, perception of distribution of their students that identify as having a disability, and their primary language.

Demographic statistics and complete surveys as displayed to the participants are available publicly [7].

5.4 Analysis

We generated descriptive statistics for each scale-style response and used Mann-Whitney Wilcoxon (MWW) statistical tests when comparing responses within the same simulation and across different voicing detail presets (PA, PB, or PC) to determine differences in rating distributions. While we only highlight MWW test results, all test results were similarly conclusive at the 95% confidence level for two-sample t-tests. Statistics were generated in the Qualtrics platform, and statistical tests were performed using the R programming language [4].

Themes were generated and assigned to the 509 additional comments educators provided through a text-entry field. Author 1 used the themes to assess the major qualitative ideas and feedback expressed through these responses regarding the educator's experience with the Voicing feature. Educator comments were separated based on simulation (survey) and assigned preset content level (ordered PA, PB, PC) and otherwise presented chronologically by submission. Themes were generated by author 1 according to the content of an educator's response, starting with Survey 1 preset A and working through to Survey 2 preset C. New themes were added when the inferred content of an educator's response did not match an existing theme. Some revisions occurred early to refine or differentiate themes. No further revisions or additional themes were added after reviewing all responses from survey 1. Multiple themes could be assigned to each response, but did not exceed 3 themes in this set. The complete theme list with expanded definitions and example responses is available publicly [7].

6 EDUCATOR'S PERCEPTIONS OF THE VOICING FEATURE

We present a subset of our findings related to our research questions in the order displayed in the survey. When interpreting these results, we consider that educator responses reflect a combination of their preferences in the use of auditory description displays and their perceptions of their students' preferences in the use of auditory description displays. The complete survey data and analysis are available publicly [7].

6.1 Voicing Statements

Table 1 presents the average rating for each statement, displayed as the difference from a neutral rating (3), calculated as Difference = Mean-3, in order to highlight the general trend in agreement for each statement. In this way, negative numbers indicate average disagreement with the statement. First, we find it interesting that averages do not reflect extreme values and remain well within 1 point of the neutral rating (3) for all statements. While there is variance in individual ratings [7], all rating distributions are unimodal about the moderate ratings (2-4, Somewhat Disagree - Somewhat Agree). This finding contrasts with earlier surveys we conducted on non-speech auditory display with a similar population pool [8]. When designing this study, we hypothesized that the Voicing feature would elicit more frequent strong agreement and disagreement ratings (5 or 1 on the scale) than we see in the data. Also, trends in the highest and lowest levels of agreement do not noticeably differ between simulations. The highest agreement occurs with statements 1, 3, and 6 related to how interesting, helpful, and natural it was to hear (some) text read while using the simulation. The lowest agreement occurs with statements 3, 10, and 11 related to, respectively, how overwhelming it was to hear text read, a desire for more text read-aloud, and the sentiment that some information was missing from the voiced content.

The hypothesis tests revealed differences in average statement rating depending on which simulation (John Travoltage or Gravity Force Lab: Basics) and what voicing detail preset (PA, PB, or PC) an educator experienced in their survey. Table 2 presents the outcomes of the non-parametric statistical tests comparing the means of each statement.

Table 2A displays the significant results when comparing preset levels in each simulation (per survey). For example, for John Travoltage, there is a significant difference in how educators rated statement 11 ("I felt like some information from the text read-aloud was missing") if they heard both Object changes and Context changes (PA) versus only Context changes (PB). In this instance, educators were more likely to agree if they only heard Context changes. Likewise, educators were more likely to agree with statement 11 if they only heard Object names and on-screen text (PC) versus Object changes and Context changes (PA). When looking at both simulations, we find it interesting that there is little overlap in statistically significant statements across presets. Rating distributions for John Travoltage appear to have more significant differences between PA and PC and more significant differences overall. The only exception appears for statement 14, in which educators tend to find the amount of text more distracting for both John Travoltage and Gravity Force Lab: Basics with both Object changes & Context changes (PA) enabled, compared to only Object names and on-screen text (PC).

Table 2B displays the significant results when comparing preset levels between the simulations (between Survey 1 and Survey 2). Only 4 of the statements across all three presets were statistically different: three related directly to the amount of text read out, and one related to user choice of read-aloud text. Educators are more likely to agree while using John Travoltage with only Object names (PC) that they would like more text after interacting and are also more likely to agree while using John Travoltage when only Context changes (PB) or only Object names (PC) are read out that they felt some text was missing. However, we note that the average ratings of each of these distributions are neutral or in slight disagreement, so we are hesitant to infer based on these results alone that participants would like to see more (or different) content added for these cases. Lastly, educators are more likely to agree, while using Gravity Force Lab: Basics when reading out both Object changes and Context changes (PA), that they felt the desire to choose word-for-word what was read aloud.

Table 3 presents average ratings and standard deviations for three statements divided by voicing detail preset and simulation presentation. Overall results are favorable. First focusing on differences in voicing detail presets: there is a significant difference in rating distributions between all voicing detail presets for "The Voicing Feature should be added to as many PhET sims as possible" in John Travoltage, toward a higher agreement with Object names and screen text (PC). For this design, the presets decrease in word amount from PA to PC, which may suggest a negative correlation between the amount of text read and desire for the Voicing feature, although more specific probes are needed.

Considering trends between the simulations: overall ratings for Gravity Force Lab: Basics are higher and have less variance. This

Table 1: Difference in average rating from a neutral rating (3, Neither Agree nor Disagree) to highlight the general trend in agreement for each statement, calculated as Difference = Mean-3. Negative numbers indicate average disagreement with the statement. Sample sizes: John Travoltage PA (371), PB (377), PC (357). Gravity Force Lab: Basics PA (406), PB (415), PC (401).

	John Travoltage			Gravity Force Lab: Basics		
Statement	**PA**	**PB**	**PC**	**PA**	**PB**	**PC**
1. It was interesting to hear text read-aloud while interacting	0.56	0.57	0.69	0.63	0.59	0.66
2. It was overwhelming to hear text read-aloud while interacting	-0.16	-0.16	-0.29	-0.12	-0.19	-0.32
3. It was helpful for me to hear text read-aloud while interacting	0.5	0.49	0.64	0.48	0.55	0.53
4. I don't like hearing long phrases read-aloud to me	0.09	0.15	0.11	0.19	0.11	0.11
5. I would prefer to choose word-for-word what is read-aloud and when	0.14	0.22	0.24	0.35	0.31	0.24
6. It seems natural to have some text read-aloud as I interact	0.41	0.54	0.62	0.45	0.52	0.53
7. It's strange to hear this amount of text read-aloud while interacting	-0.01	0.09	-0.04	0.09	0.1	-0.09
8. I felt like removing some of the text read-aloud while interacting	0.15	0.07	0	0.23	0.12	0.06
9. It felt natural to hear this amount of text read-aloud while interacting	0.09	0.18	0.33	0.11	0.17	0.23
10. I wanted more text to be read-aloud after interacting	-0.22	-0.13	-0.02	-0.21	-0.2	-0.2
11. I felt like some information from the text read-aloud was missing	-0.25	-0.1	-0.03	-0.33	-0.3	-0.22
12. I would prefer to hear more... about the effects my interactions have in the sim	0.05	0.15	0.29	0.07	0.06	0.17
13. I would prefer to hear more... about the interactions I make	-0.11	0.02	0.1	-0.1	-0.07	-0.04
14. It was distracting to hear this amount of text read-aloud while interacting	-0.03	-0.03	-0.25	0.03	-0.05	-0.14

difference potentially indicates that the simulation educators experienced impacts their desire to see the Voicing feature added to additional interactive simulations. There is also a consistent, significant decrease in ratings when educators reflect on the benefit of Voicing for their students and students broadly for both simulations. We conjecture this decrease may stem from the ease of imagining a specific example or situation where they want to enable the feature but not anticipating that their context is generalizable. However, more information is needed to infer the differences educators may perceive between their students and the population of students more generally. This decrease does not seem to significantly alter their responses to their desire for the feature in future PhET sims, indicating a desire for the feature for at least their specific circumstances.

Table 4 displays the response percentages for "After using this simulation, which of the following learners do you feel could potentially benefit from using a simulation with a Voicing feature like this one? Please check all learners that you feel could benefit." for Survey 1, John Travoltage. Percentages for Survey 2, Gravity Force Lab: Basics differ less than 1% for each selection. Educator's comprehensive selections of learners that may benefit from the Voicing feature were very similar for the specified populations. Notably, less than 1% of selected options by educators included None of the above. Highlighted frequently in the "Other learners:" optional text box for both surveys are educators calling out specific intellectual/developmental/learning disabilities (e.g., attention deficit hyperactivity disorder, autism spectrum disorders, etc.) and learners at various levels of English fluency (e.g., ESL and English language learners). One population not included in the checkbox selections frequently offered in both surveys' text entry is that "all learners" would benefit from the Voicing feature (lower bound of 20 responses out of 251 all/every learner/student excluding more complex responses).

6.2 Thematic summary of educator's comments about Voicing

We present a broad look at all of the themes observed in 509 educator responses to "Please share any additional comments you have". We avoid quantitative analysis of the frequency of the themes due to the lack of inter-rater analysis and present the list of identified themes for supplementary intrigue. Responses ranged from suggestions, comments, anecdotes, criticism, and praise about or for the Voicing feature, their students, all learners, and the simulations more generally. The list of themes observed in the responses, worded to be self-explanatory and ordered by approximate frequency from more frequent to less frequent, is:

- Voicing as an optional feature
- Voicing is beneficial for specific population
- Information provided by Voicing
- Voicing is ineffective for specific population
- Speech engine quality
- General praise for the feature
- Voice interrupts itself
- Voicing supports simulation scaffolding
- Voicing is distracting
- Assistive technology comparisons or recommendations
- Voicing lags behind interaction
- Voicing conflicts with simulation scaffolding
- Voicing helps to highlight information
- Speech engine rate
- Voicing phrase length
- Personal dislike of the feature
- Voicing is repetitive
- Voicing complements other sensory feedback
- Voicing and sound overlap

We note that this text-entry field existed in the same section as the questions presented in Tables 3 & 4, so the frequent comments

Table 2: Outcomes of significant statistical test results for the mean ratings of each statement. A) Comparing sample means between different presets for the same simulation. An entry indicates which simulation was significant and the higher-rated preset indicated in parenthesis. B) Comparing sample means for the same preset between different simulations. An entry indicates a significant difference between the mean of the preset A, B, or C (column) and an inequality indicating the higher mean rating.

A) Different presets, same simulation

#	PA/PB	PA/PC	PB/PC
1	-	-	-
2	-	GFLB (PA)	-
3	-	-	-
4	-	-	-
5	-	-	-
6	-	JT (PC)	-
7	-	GFLB (PA)	GFLB (PB)
8	-	GFLB (PA)	-
9	-	JT (PC)	JT (PC)
10	-	JT (PC)	-
11	JT (PB)	JT (PC)	-
12	-	JT, PC	-
13	-	JT (PC)	-
14	-	BOTH (PA)	JT (PB)

B) Same preset, different simulations

#	PA	PB	PC
1	-	-	-
2	-	-	-
3	-	-	-
4	-	-	-
5	JT >GFLB	-	-
6	-	-	-
7	-	-	-
8	-	-	-
9	-	-	-
10	-	-	GFLB >JT
11	-	GFLB >JT	GFLB >JT
12	-	-	-
13	-	-	-
14	-	-	-

acknowledging who may or may not benefit from the feature is not surprising, though of interest are the specific populations called out (discussed in Section 7.1).

Of the themes related directly to the speech engine, the quality of the voice (e.g., "robotic" nature) was frequent, and there was some overlap in identifying the Voicing as distracting and the quality of the voice. Participants were not explicitly pointed toward the Customize Voice submenu in the Preferences menu that would allow them to choose a different voice, nor were they made aware that they were limited to the voices available on their device (a limitation of Web Speech).

Interestingly, although infrequent, most of the participants who communicated a personal dislike for the Voicing feature also–at the same time– identified the Voicing as beneficial for a group of people. This overlap gave us a sense that they acknowledge the feature as potentially helpful–just not for them.

Lastly, the most prevalent sentiment, "Voicing as an optional feature", expressed a desire to toggle some or all of the Voicing details. The optional toggling of the feature was already a part of the feature at the time of the survey. While participants were not explicitly encouraged to disable Voicing during the survey, it was required that they toggle it on in order to engage with the Voicing feature during the survey, as it initialized in the off state.

Table 3: Average ratings of statements for educators perceptions of student benefit and educator desire for the Voicing feature in PhET sims for Gravity Force Lab: Basics (GFLB) and John Travoltage (JT). *Mean ratings were statistically significant between all presets for John Travoltage for "The Voicing Feature should be added to as many PhET sims as possible". Sample sizes: John Travoltage PA (371), PB (377), PC (357). Gravity Force Lab: Basics PA (406), PB (415), PC (401).

Statement	JT Mean (SD)	GFLB Mean (SD)
My students would benefit from using the PhET sims with the Voicing feature.	PA: 3.65 (2.37) PB: 3.77 (2.23) PC: 3.78 (2.35)	PA: 3.96 (0.99) PB: 4.06 (0.96) PC: 4.05 (0.97)
All students would benefit from using the PhET sims with the Voicing feature.	PA: 3.24 (2.47) PB: 3.33 (2.34) PC: 3.34 (2.35)	PA: 3.53 (1.17) PB: 3.62 (1.12) PC: 3.68 (1.09)
The Voicing feature should be added to as many PhET sims as possible.	PA: 3.50 (2.39)* PB: 3.63 (2.24)* PC: 3.72 (2.41)*	PA: 3.88 (1.06) PB: 3.93 (1.02) PC: 4.02 (0.99)

Table 4: Response percentages for "After using this simulation, which of the following learners do you feel could potentially benefit from using a simulation with a Voicing feature like this one? Please check all learners that you feel could benefit." for Survey 1 (John Travoltage).

Checkbox Selection	Response %
Learners who have difficulty interpreting visual language or mathematical representations (e.g., dyslexia, dyscalculia, language learners)	25.97%
Learners who have difficulty seeing the visual display (e.g., have low vision)	23.24%
Learners who may have difficulty getting started with using a simulation, or need explicit guidance to enact specific interactions (e.g., younger learners, learners with intellectual or developmental disabilities)	24.59%
Learners with diverse auditory and visual processing needs	21.83%
Other learners:	3.79%
None of the above	0.57%

For the Discussion (Section 7), we will further expand on a subset of these themes that were of particular interest to our research questions. The complete list of themes, including expanded definitions and example quotes, is available publicly [7].

7 DISCUSSION

We present the discussion of our findings in the order of our research questions.

7.1 Do educators perceive the Voicing feature as beneficial, and if so, what populations of learners do educators identify to likely benefit from the Voicing feature?

Through our analysis, we find that educators find the Voicing feature desirable and beneficial both broadly and for specific populations. Educators rated their beliefs that the feature would be beneficial to their students and all students favorably, as well as their desire to see the feature included in more interactive simulations (Table 3). Additionally, we see remarkably low selection frequency for the "None of the above" category when asking educators to identify learner groups that may benefit from the Voicing feature and an even distribution across the predefined options.

Themes we found that are likely unique to the educators involved in this study and whether they would use it in their classroom are how the feature did or did not support pedagogical scaffolding or self-directed science learning. These comments held both positive and negative sentiments, identified as Voicing providing supporting interaction and exploration (much like a teacher or peer may) or removing agency from the student in discovering and reasoning through the scientific content themselves. For example, one educator wrote: "I suggest it be limited or controlled by the teacher if possible so that the conclusions are not readily available for those who are discovering by inquiry." This educator notes their belief that the Voicing explicitly provides conclusions that shortcut them toward the "point" of the simulation.

This concern is a careful balance in designing both visual and auditory elements of the simulation. The Voicing design intends to describe the aspects of the simulation that the learner is changing and a direct description (not interpretation) of what happens on screen due to learner changes. However, we acknowledge it may introduce vocabulary not explicitly in the visual display. Encapsulating the Voicing design and the design of other sim elements is the idea of implicit scaffolding [14] intended to support learner exploration and discovery. Implicit scaffolding encourages interactions that lead to further discovery but do not point directly toward or reward answers.

In determining who may or may not benefit from the Voicing feature, there is additional insight within the "Other learners:" text-entry choice from Table 4 and themes from the Additional Comments responses. We find educators frequently reference the feature as a helpful focus tool or to highlight important information prudent to sim exploration. This sentiment by educators covers three diverse use scenarios: 1) As a conceptual reinforcement or present introductory language for learners uncomfortable with the topic, 2) a way to slow down or diversify the content for individuals who vary in their visual/auditory processing or executive function (e.g., ADHD) and 3) a way to encourage interaction and attention for the general population by activating another (non-visual) sense.

The participant pool is from a global community of teachers and educators. Participants frequently identify the Voicing feature as a learning support and helpful guide for English Language Learners and Non-English readers, assuming the vocal parameters are set in a way to be understood or able to be adjusted. However, significant numbers of respondents, the majority of responses aligned with "Voicing is ineffective for specific population", indicated the reverse. Since the feature was only available in the English language, some participants indicated that it was not helpful for learners in non-English speaking countries or for those with low English fluency. While the on-screen text in all PhET Interactive Simulations is translatable to any language through a translation tool supported by a community of translators, such a tool does not currently exist for the Voicing or Interactive Description features.

Educators noted that the Voicing feature is only beneficial to learners who can hear the spoken description. This excludes learners who are deaf or hard of hearing and learners in contextual situations where it is impossible or difficult to play audio (e.g., in large classroom settings). A frequently mentioned complementary feature for the spoken descriptions was a form of closed captioning displayed during simulation interaction. This feature does not currently exist, but we have actively considered this and are interested in exploring such a feature using scalable web technologies.

Ultimately, we do not find evidence that the participants perceive the tool as strictly designed for a single group of learners (e.g., only those with visual impairments). Instead, we find an emphasis on learner populations not often associated with using auditory description tools. This gives us confidence in the Voicing feature and leads us to believe that it would be helpful to expand the optional feature to more learners by adding it to more simulations.

7.2 How does educator preference for Voicing change with the amount of voicing presented to them?

Surprisingly, we do not find significant evidence in the ratings or in the comments to make firm conclusions about correlations between the amount of spoken description, but the data does provide some leads. Most of the significant differences in rating distributions across the Voicing statements were found in the comparisons between Object changes & Context changes (PA) and Object names and on-screen text (PC) (Table 2). Respectively, these presets correlate with the delivery of more spoken description (PA) and the delivery of less spoken description (PC) during interaction in the case of these two simulations. This assumption of more or less spoken description does not account for any on-demand description activated by the user through their possible intermittent use of the "play buttons" in the Toolbar, and may not be true for simulations with more on-screen text or more complex object names.

In this scenario, however, the presence of multiple significant differences between these two presets lends evidence to a difference in interpretation of the Voicing feature based on the amount of spoken description during interaction. For example, participants were more likely to agree that they would like to remove some of the text read-aloud in Gravity Force Lab: Basics when hearing both

Object changes & Context changes (PA), than if they heard only Object names and screen text (PC). Similar results appear for John Travoltage in statements 10-14, all of which indicate a preference for less text when exposed to PA compared to PC.

However, it remains surprising how few significant differences there are and how small the changes are between each preset (i.e., PA, PB, and PC). This may result from participant acknowledgment that the feature is optional or "intended" for an audience other than themselves, resulting in more neutral ratings across the board. Voicing is a feature that is likely uncommon in most educator's day-to-day and never before present in a simulation they have used. The presence of any voicing may play more of an important role than the actual amount of voicing played in the surveyed simulations. Alternatively, although we predicted statement agreement would be primarily determined by phrase length or amount, educator perceptions may be founded in other variables not accounted for in this study. A focused qualitative inquiry into educators' perceptions of the feature may provide more insight.

7.3 How does educator preference for Voicing change based on the interactive design of the simulation?

The interactive design varies for every simulation–expectations for when, how often, and in what ways learners move objects or press buttons depend on the presentation and nature of the science topic of the simulation. Do changes arise while the learner is moving an object? Are there one or more elements that continue to change after a learner has changed an object? Or is there some element that continues to change regardless of learner input? John Travoltage and Gravity Force Lab: Basics share commonalities but differ when asking these questions.

In either simulation, spoken description does occur if a learner continuously moves an object (e.g., dragging a sphere back and forth across the play area), but moving the leg multiple times across the carpet in John Travoltage is a requisite to exploring the full context of the simulation. In the case of Voicing design for John Travoltage, the contextual changes appear in two separate chunks, one is the build-up of charges on John's body while moving the leg, and the second is the discharge event, which can happen when exploring with either the leg or the arm (and even after they have stopped moving). Learners may not initially anticipate that a discharge is possible with both interactions. A discharge that happens while exploring with the leg may be more unexpected and thus possibly may cause misalignment with expectations. In the case of Gravity Force Lab: Basics, the simulation design only needs one chunk of contextual changes. Once a learner stops interacting and hears the latest change to the force, nothing more can happen in the simulation. There are no surprise context changes that might cause misalignment.

Some participant responses tied to a misalignment in expectation are under the theme of "Voice interrupts itself". One educator wrote: "The main problem I noticed was the overlap of sentences. I can predict students would get distracted by that. However, I think that's [sic] more of an issue for Travoltage because of the very dynamic situation." This educator acknowledges the "dynamic" nature of John Travoltage and connects it to the restarting spoken description.

As previously discussed, more statements had significant differences for John Travoltage than for Gravity Force Lab: Basics between PA and PC (Table 2), with most of the differences coming from statements related to the amount of text-read aloud during interaction. In Table 3, the trend in ratings for the benefit to students and desire for the feature in more simulations for Gravity Force Lab: Basics is the same as John Travoltage; however, the magnitude of the mean is consistently higher for participants using Gravity Force Lab: Basics than John Travoltage, indicating an overall preference for the Voicing as they were presented it in the former.

These findings suggest that the interactive design of the simulation may influence educator perceptions of the Voicing feature, although these two simulations do not represent all interaction design patterns that a simulation may have, nor are the designs mutually exclusive. However, we remark again that educators found the Voicing feature, in general, to be both beneficial and desirable as a feature for the simulations.

8 CONCLUSIONS AND FUTURE WORK

This study found that educators approve of an optional Voicing feature designed for interactive science simulations for teaching and learning. While generally favorable toward Voicing, educators' perceptions were less statistically extreme across preset Voicing detail levels than anticipated. We find that educators perceive the feature as beneficial both broadly and for specific populations, while some acknowledge particular populations for whom it remains ineffective. Notably, the populations identified are not limited to populations commonly associated with using an auditory description of interactive media. Lastly, we identify some variance in the perceptions of the feature based on different aspects of the interactive design of the simulation, which provokes exciting questions for those designing similar Voicing features in interactive media.

We are interested in making further improvements to the design of the Voicing feature and a deeper examination of how the feature can meet the needs of diverse learners and usage in diverse learning environments. On the design side, future work will explore options for visually presenting the voiced information on the screen to support more diverse learners. More immediately, we will apply what we have learned thus far to balance the types of information spoken, the amount of information spoken, and the general usability of the feature while interacting with the simulation. We will be continuing interviews directly with learners over the evolution of the feature as we implement it in more simulations. In future design work, we will improve the Voicing design and implementation process by better understanding how the Voicing responses relate to the State and Responsive descriptions of our description design framework. As we learn more about the design, uses, and benefits of this feature, we will share guidelines for designing web-based Voicing features.

ACKNOWLEDGMENTS

We thank the many educator participants who contributed their time and perspectives in filling out the surveys. This material is based on work supported by the National Science Foundation under DRL-#1814220.

REFERENCES

[1] DIAGRAM Center. 2015. *Image Description Guidelines.* http://diagramcenter.org/table-of-contents-2.html

[2] DIAGRAM Center. 2022. *Poet Image Description.* https://poet.diagramcenter.org/index.html

[3] World Wide Web Consortium. 2022. *Web Content Accessibility Guidelines (WCAG) Overview.* https://www.w3.org/WAI/standards-guidelines/wcag/

[4] Team R Core. 2022. *R Foundation for Statistical Computing.* https://www.r-project.org/

[5] Richard Ely, Robert Wall Emerson, Theresa Maggiore, Madeleine Rothberg, Trisha O'Connell, and Laurel Hudson. 2006. Increased Content Knowledge of Students with Visual Impairments as a Result of Extended Descriptions. *Journal of Special Education Technology* 21 (2006). Issue 3. https://doi.org/10.1177/016264340602100304

[6] Brett L. Fiedler, Emily B. Moore, Tiara Sawyer, and Bruce N. Walker. 2021. Multimodality and inclusion: Educator perceptions of physics simulation auditory display. *2021 Physics Education Research Conference Proceedings*, 123–128. https://doi.org/10.1119/perc.2021.pr.Fiedler

[7] Brett L. Fiedler, Taliesin L. Smith, and Emily B. Moore. 2022. *Surveys of Auditory Display in PhET Interactive Simulations (2020-2021).* https://doi.org/10.17605/OSF.IO/H84VZ

[8] Brett L. Fiedler, Bruce N. Walker, and Emily B. Moore. 2021. To sonify or not to sonify?: Educator perceptions of auditory display in interactive simulations. *2021 International Conference on Auditory Display Conference Proceedings.* https://doi.org/10.21785/icad2021.xxx

[9] Andrew Holland. 2008. *Audio description in the theatre and the visual arts: Images into words.* https://doi.org/10.1057/9780230234581

[10] Steve Jacobs. 2002. *The electronic curbcut effect.* http://www.icdri.org/technology/ecceff.htm

[11] Mina C. Johnson-Glenberg, Caroline Savio-Ramos, Katherine K. Perkins, Emily B. Moore, Robb Lindgren, Douglas Clark, Corey Brady, Pratim Sengupta, Mario Martinez-Garza, Deanne Adams, Stephen Killingsworth, Grant Van Eaton, Matthew Gaydos, Amanda Barany, Kurt Squire, and Nathan Holbert. 2014. Science sims and games: Best design practices and fave flops. *Proceedings of International Conference of the Learning Sciences, ICLS* 3. Issue January.

[12] Kyle Keane and Christina Laverentz. 2014. Interactive scientific graphics: Recommended practices for verbal description. (2014).

[13] Alexander Mathews, Lexing Xie, and Xuming He. 2016. SentiCap: Generating image descriptions with sentiments. *30th AAAI Conference on Artificial Intelligence, AAAI 2016.*

[14] Emily B. Moore, Timothy A. Herzog, and Katherine K. Perkins. 2013. Interactive simulations as implicit support for guided-inquiry. *Chem. Educ. Res. Pract.* 14 (2013), 257–268. Issue 3. https://doi.org/10.1039/C3RP20157K

[15] Emily B. Moore and Clayton Lewis. 2015. Opportunity: Inclusive design for interactive simulations. *ASSETS 2015 - Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility.* https://doi.org/10.1145/2700648.2811387

[16] Emily B. Moore, Taliesin L. Smith, and Jesse Greenberg. 2018. Keyboard and screen reader accessibility in complex interactive science simulations: Design challenges and elegant solutions. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10907 LNCS. https://doi.org/10.1007/978-3-319-92049-8_28

[17] Mozilla. 2022. *Web Audio API.* https://developer.mozilla.org/en-US/docs/Web/API/Web_Audio_API

[18] Mozilla. 2022. *Web Speech API.* https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API

[19] Jaclyn Packer, Katie Vizenor, and Joshua A. Miele. 2015. An overview of video description: History, benefits, and guidelines. *Journal of Visual Impairment and Blindness* 109 (2015). Issue 2. https://doi.org/10.1177/0145482x1510900204

[20] PhET Interactive Simulations. 2022. *PhET Simulations with Inclusive Features.* https://phet.colorado.edu/en/simulations/filter?a11yFeatures=accessibility&sort=alpha&view=grid

[21] Shivam Singh, Anurag Bhandari, and Nishith Pathak. 2018. Accessify: An ML powered application to provide accessible images on web sites. *Proceedings of the 15th Web for All Conference : Internet of Accessible Things, W4A 2018.* https://doi.org/10.1145/3192714.3192830

[22] Taliesin Smith. 2016. Access, Action, & Agency: Inclusive Design for the Non-visual Use of a Highly Interactive Simulation. http://openresearch.ocadu.ca/id/eprint/713

[23] Taliesin L. Smith, Jesse Greenberg, Sam Reid, and Emily B. Moore. 2018. Parallel DOM architecture for accessible interactive simulations. *Proceedings of the 15th Web for All Conference : Internet of Accessible Things, W4A 2018.* https://doi.org/10.1145/3192714.3192817

[24] Taliesin L. Smith, Clayton Lewis, and Emily B. Moore. 2016. A balloon, a sweater, and a wall: Developing design strategies for accessible user experiences with a science simulation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9739. https://doi.org/10.1007/978-3-319-40238-3_15

[25] Taliesin L. Smith, Clayton Lewis, and Emily B. Moore. 2016. Demonstration: Screen reader support for a complex interactive science simulation. *ASSETS 2016 - Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility.* https://doi.org/10.1145/2982142.2982154

[26] Taliesin L Smith, Clayton Lewis, and Emily B Moore. 2017. Description strategies to make an interactive science simulation accessible. *Journal on Technology and Persons with Disabilities* (2017), 225–238. http://scholarworks.csun.edu/handle/10211.3/190214

[27] Taliesin L. Smith and Emily B. Moore. 2020. Storytelling to Sensemaking: A Systematic Framework for Designing Auditory Description Display for Interactives. *Conference on Human Factors in Computing Systems - Proceedings.* https://doi.org/10.1145/3313831.3376460

[28] Volker Sorge, Mark Lee, and Sandy Wilkinson. 2015. End-to-end solution for accessible chemical diagrams. *W4A 2015 - 12th Web for All Conference.* https://doi.org/10.1145/2745555.2746667

[29] Abigale Stangl, Nitin Verma, Kenneth R. Fleischmann, Meredith Ringel Morris, and Danna Gurari. 2021. Going beyond One-Size-Fits-All Image Descriptions to Satisfy the Information Wants of People Who are Blind or Have Low Vision. *ASSETS 2021 - 23rd International ACM SIGACCESS Conference on Computers and Accessibility.* https://doi.org/10.1145/3441852.3471233

[30] Jennifer L Tennison, Jesse Greenberg, Emily B Moore, and Jenna L Gorlewicz. 2021. Haptic Paradigms for Multimodal Interactive Simulations. *The Journal on Technology and Persons with Disabilities* (2021), 110.

[31] Brianna Tomlinson, Prakriti Kaini, Bruce Walker, Jared Batterman, and Emily Moore. 2018. Supporting Simulation Use for Students with Intellectual and Developmental Disabilities. *Journal on Technology and Persons with Disabilities* 6 (2018).

[32] Brianna J. Tomlinson, Prakriti Kaini, Siyan Zhou, Taliesin L. Smith, Emily B. Moore, and Bruce N. Walker. 2018. Design and evaluation of a multimodal science simulation. *ASSETS 2018 - Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility.* https://doi.org/10.1145/3234695.3241009

[33] Brianna J. Tomlinson, Brittany E. Noah, and Bruce N. Walker. 2018. BUZZ: An auditory interface user experience scale. *Conference on Human Factors in Computing Systems - Proceedings* 2018-April. https://doi.org/10.1145/3170427.3188659

[34] Brianna J. Tomlinson, Bruce N. Walker, and Emily B. Moore. 2020. Auditory Display in Interactive Science Simulations: Description and Sonification Support Interaction and Enhance Opportunities for Learning. *Conference on Human Factors in Computing Systems - Proceedings* (2020), 1–12. https://doi.org/10.1145/3313831.3376886

[35] Shaomei Wu, Jeffrey Wieland, Omid Farivar, and Julie Schiller. 2017. Automatic alt-text: Computer-generated image descriptions for blind users on a social network service. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW.* https://doi.org/10.1145/2998181.2998364