applied sciences

MDPI

Article

Voice-Interactive 2D Serious Game with Three-Tier Scaffolding for Teaching Acoustics in Primary Schools: A Randomized Comparison of Knowledge, Motivation, and Cognitive Load

Minyu Che [1], Hongrun Li [2], Zhiwei Chen [3], Qiang Li [4,*] and Nayoung Kim [2,*]

[1] Department of Design, Sungkyunkwan University, Seoul 03063, Republic of Korea; cheminyu@skku.edu
[2] Department of Game Graphic Design, Hongik University, Sejong 30016, Republic of Korea; polestarclark@g.hongik.ac.kr
[3] Department of Communication Design, Hongik University, Sejong 30016, Republic of Korea; chenzhiwei950818@g.hongik.ac.kr
[4] School of Design and Art, Shenyang Aerospace University, Shenyang 110136, China
* Correspondence: qiangli@sau.edu.cn (Q.L.); nayoung@hongik.ac.kr (N.K.)

Abstract

Misconceptions about sound are common among primary-school pupils, but research on voice-interactive game-based learning remains limited, especially regarding the role of scaffolding. We investigated whether a voice-interactive 2D platformer with a three-tier scaffolding model improves learning about loudness, pitch, and echo. In a classroom-feasible randomized comparison, 45 third-graders were assigned to a scaffolded serious game (SSG), a non-scaffolded serious game (NSG), or traditional hands-on materials instruction (TRAD) on matched sound content. Outcomes were an immediate eight-item knowledge test and learner-centered ratings of perceived learning, flow, intrinsic motivation, and extraneous cognitive load (ECL). The knowledge test showed low internal consistency, so results involving this measure should be interpreted with caution. SSG yielded higher immediate learning than NSG and matched traditional instruction. Across experience measures, only intrinsic motivation differed, favoring NSG. Hierarchical regression revealed a motivation-by-structure effect: scaffolding strengthened the positive association between intrinsic motivation and test scores, whereas ECL was not predictive. Findings indicate that voice-interactive serious games can match near-term learning achieved with physical materials, and well-calibrated scaffolds help convert motivation into accurate encoding. We also map sound constructs to gameplay mechanics and provide a compact, classroom-feasible, replicable evaluation design for primary classrooms.

Keywords: serious games; scaffolding; voice interaction; primary-school acoustics; intrinsic motivation; extraneous cognitive load

check for updates

Academic Editors: Douglas O'Shaughnessy and Lamberto Tronchin

Received: 11 September 2025
Revised: 31 October 2025
Accepted: 31 October 2025
Published: 4 November 2025

Citation: Che, M.; Li, H.; Chen, Z.; Li, Q.; Kim, N. Voice-Interactive 2D Serious Game with Three-Tier Scaffolding for Teaching Acoustics in Primary Schools: A Randomized Comparison of Knowledge, Motivation, and Cognitive Load. *Appl. Sci.* **2025**, *15*, 11761. https://doi.org/10.3390/app152111761

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Primary science experiences shape later scientific literacy and attitudes [1]. Without explicit correction, misconceptions about sound and waves commonly persist [2]. Learners often conflate loudness with pitch or treat sound as if it were a self-propagating substance [3]. Science aspirations formed between ages 10–14 are highly stable and strongly predict later participation, underscoring the gatekeeping role of primary schooling [4]. Taken together, these findings suggest that, in primary science education, adopting serious games alongside other computer-based strategies can enhance conceptual learning, sustain meaningful engagement, and help identify and correct students' misconceptions [5].

Globally, primary school science education increasingly emphasizes active, inquiry-driven exploration, including work with sound. Constructivist accounts show that children learn science best through hands-on investigation and personal meaning-making [6]. Sound is widely represented in curricula, and textbooks commonly organize it by formation, propagation, structure, and perception [7]. Yet research on sound learning remains limited [8], and persistent misconceptions endure: many learners treat sound as a substance rather than a vibratory process, and younger pupils often struggle with propagation [9,10]. Evidence from pre-service teachers shows similarly mixed, often materialistic conceptions, underscoring the need for stronger pedagogical support [8]. In Chinese science classrooms, exam and evaluation priorities and resource conditions may limit opportunities for open-ended science inquiry, including work on sound, suggesting that additional support in curriculum, teacher development, and assessment could be helpful [11,12]. These conditions risk marginalizing sound-based science learning and point to reforms that better align curricula, assessment, and inquiry-based pedagogy in the primary years.

In primary school science education, researchers increasingly use digital technologies such as virtual reality, game-based platforms, and interactive digital media to investigate and demonstrate sound phenomena [7,13–15]. Digital formats often deliver superior learning experiences compared with traditional instruction [16,17]. Accordingly, games have been widely integrated into practice, from collaborative computer games to augmented reality inquiry, and studies consistently report gains in achievement, motivation, and positive affect [18–20]. A common explanation is the high engagement that games afford, with challenging goals, immediate feedback, and rapid cycles of action and consequence sustaining attention and effort [21,22]. 2D platformer serious games with a level-based (“stage-clear”) structure, explicit objectives, calibrated difficulty, and on-the-spot feedback can sustain engagement while scaffolding incremental knowledge acquisition [23,24]. In practice, each short stage introduces or rehearses one concept; unambiguous goals and immediate feedback make the target behavior and errors transparent; and difficulty is raised as competence grows—an arrangement that fosters flow and learning [23,24]. Importantly, effects can vary by learner profile: experimental evidence indicates strong emotional and achievement benefits for low and middle achievers, underscoring the value of adjustable difficulty and explicit scaffolds [18]. Beyond achievement levels, the effectiveness of gamification hinges on its fit with learners’ player profiles. Empirical work in engineering education indicates that instructors’ profiles often diverge from those of students, risking designs that unintentionally mirror instructor preferences rather than learner needs; accordingly, gamified environments should be aligned with student rather than instructor profiles [25]. Moreover, collaborative arrangements that couple gameplay with knowledge organization further amplify gains in achievement, motivation, and self-efficacy, suggesting that social regulation and externalized reasoning complement game-based practice [20]. In CRYSTAL ISLAND: Uncharted Discovery, an exploration-oriented science adventure, primary pupils showed significant gains in content knowledge and problem-solving after gameplay [26]. Similarly, the fifth-grade, mission-based variant of CRYSTAL ISLAND improved fifth graders’ science knowledge and self-efficacy [27]. Yet sound-focused educational games, especially 2D task-driven titles that model wave propagation and other acoustic properties, remain rare due to development costs and classroom adoption barriers [28]. Consequently, research on teaching sound has emphasized documenting students’ misconceptions, an essential foundation for instruction that targets and corrects them; however, relatively few studies have designed and rigorously evaluated interactive interventions that address these misconceptions [29]. However, empirical evaluations of 2D platformer serious games for primary science, especially sound, remain scarce.

Embedding problem-solving in playful, game-based environments can transform complex challenges into engaging learning experiences [30]. Many game-based learning studies ground design in flow, self-determination, cognitive load, and scaffolding, aligning goals, feedback, and challenge while managing extraneous load and offering just-in-time, contingent supports [24,31–35]. However, relatively few studies explicitly operationalize these theories in science education games, particularly serious games that teach sound-related concepts to primary pupils.

In this study, we designed a 2D platformer serious game, grounded in Brunerian scaffolding as articulated by Wood, Bruner, and Ross [31,34], and operationalized via a three-tier software-scaffolding framework [33,35] to teach primary pupils scientific knowledge about sound and to evaluate its effectiveness for improving learning performance and knowledge acquisition. We compared three instructional conditions with matched content: a scaffolded serious game (SSG), a non-scaffolded serious game (NSG), and traditional hands-on materials instruction (TRAD). We measured perceived learning (PL), flow, intrinsic motivation, extraneous cognitive load (ECL), and knowledge test performance. Following flow theory, we conceptualize flow as a state of deep absorption and effortless control that arises when perceived challenge matches skill under clear goals and immediate feedback [36]. Guided by Cognitive Load Theory (CLT) [37–39], following Sweller and colleagues, we distinguish intrinsic load (IL; content complexity relative to prior knowledge), ECL (avoidable demands imposed by the presentation and interaction format), and germane cognitive load (GCL; effort invested in schema construction). Because our manipulation varied only the presentation and interaction format while holding content complexity constant, we designated ECL as the primary outcome. Intrinsic load was held constant by design and therefore not measured, and germane load was not measured. Accordingly, we refrain from mechanistic claims about load components. Complementing this CLT perspective, we ground our interpretation of intrinsic motivation in Self-Determination Theory (SDT) [40,41]. SDT posits that the basic psychological needs of autonomy, competence, and relatedness underlie intrinsic motivation. Designs perceived as autonomy-supportive (e.g., meaningful choice, non-controlling language) and competence-supportive (e.g., clear goals, diagnostic feedback) tend to enhance enjoyment and persistence, whereas controlling supports can dampen intrinsic motivation. In this study, we therefore treat enjoyment as a proximal indicator of intrinsic motivation within the SDT framework; we did not directly measure need satisfaction.

In sum, we compared SSG, NSG, and TRAD to investigate the added value of three-tier scaffolding in a 2D platformer designed to teach primary pupils about sound, including its effectiveness in improving learning performance and knowledge acquisition. We measured PL, flow, intrinsic motivation, ECL, and knowledge test performance. Positioned at the intersection of human–computer interaction, interactive systems and game design, and the learning sciences, this work couples an applied system contribution with evidence-based analysis of learning.

2. Related Work

This section systematically reviews research on serious games in primary-school science and on scaffolding within serious games, synthesizes evidence on engagement mechanisms and learning outcomes, identifies gaps in voice-interactive approaches to acoustics and in comparative effectiveness versus hands-on materials, and lays the theoretical foundation for the present study.

2.1. Serious Games in Primary Science Education

Serious games are digital or non-digital games designed primarily for learning, training, or behavior change [42]. Gameplay is structured by explicit objectives and aligned content, rules, and feedback, and entertainment supports the learning aims rather than serving as an end in itself. In child-focused research, the term has gained traction, and converging evidence links well-designed digital games to learner engagement and knowledge construction [43,44]. In this literature, outcomes are typically operationalized as knowledge tests, brief self-report scales, and simple in-game behavioral metrics [45]. Building on these conventions, recent frameworks emphasize embedding motivational and collaborative elements alongside disciplinary content to strengthen learning [46]. Experimental comparisons further show advantages for intrinsic motivation and flow under game conditions relative to traditional instruction, although links to achievement can vary across designs [45,47]. Within flow theory, platformer structures with progressive difficulty, clear sub-goals, and immediate feedback instantiate the conditions conducive to flow in primary pupils, which can, in turn, support persistence and strategy use [36]. From an SDT perspective, game features such as meaningful choice, optimal challenge, and informative feedback can satisfy autonomy and competence needs, thereby strengthening intrinsic motivation in primary pupils [40,41]. Situated within a narrative-based learning perspective, the quest storyline is intended to organize causal sequences and support meaning-making. This aligns with Bruner's view of narrative as a fundamental mode of thought and with reviews and meta-analyses showing that narrative-based (and narrative-gamified) environments can enhance conceptual understanding, motivation, and engagement in science education [5,31,48–51]. Importantly, enjoyment and user experience can predict learning in child samples [43], underscoring the value of learner-centered measures. Complementing this evidence, Buckley and Doyle [52] reported that a gamified intervention produced significant pre–post knowledge gains and that intrinsic-motivation subtypes were positively associated with participation, suggesting that motivational profiles can condition the benefits of game-based learning.

Sound is a core topic in primary science curricula. Reported gains span a 2D image schema game for early primary pupils [53], a gamified STEM "Sound" unit in upper primary grade 6 [54], and a light-and-sound game that pupils reported as supporting perceived learning, enjoyment, and motivation [29]. However, many evaluations rely on touch-based interaction, small samples, or weak controls, making it difficult to attribute learning to game mechanics rather than digital novelty. Voice-controlled titles such as Flappy Voice and classroom noise-feedback systems indicate that microphone input can engage pupils, yet these systems target speech therapy or behavior management rather than scientific knowledge about sound [55,56]. The evidence based on serious games that teach acoustics through voice-based interaction remains limited; moreover, we identified no studies contrasting such games with TRAD. Work on children's conceptions also shows persistent misconceptions about sound propagation [57]. Taken together, these observations delineate a clear gap that motivates the present study.

Consequently, we investigated whether a voice-interactive serious game improved primary pupils' scientific knowledge about sound. We compared three instructional conditions, namely SSG, NSG, and TRAD, on matched content to benchmark effects against a widely used classroom baseline where concrete artifacts and embodied manipulation are integral to teaching sound. This design addresses two gaps: the scarcity of voice-based serious games for teaching acoustics and the absence of comparisons with hands-on instruction. We also examine learner-centered correlates, including PL, flow, intrinsic motivation, and ECL, to situate any achievement differences within a broader profile of experience.

2.2. Scaffolding in Serious Games

Scaffolding, originally articulated by Wood, Bruner, and Ross, refers to temporary, contingent support that enables learners to perform tasks just beyond their independent capability, with assistance progressively withdrawn as competence increases [31,34]. In serious games, this approach is associated with improved learning outcomes [58]. Within digital game-based learning (DGBL), supports are typically organized into three tiers; we operationalize them here as active, contingent, and implicit [33,35]. Within the active tier, scaffolds that are overly explicit or delivered preemptively (i.e., before learners engage with the task) can increase perceived difficulty, undermine autonomy, and reduce enjoyment [59]. This interpretation is consistent with SDT: controlling scaffolds may frustrate autonomy and reduce intrinsic motivation, whereas contingent and implicit scaffolds can support competence while preserving autonomy [40,41]. Among ten-year-old learners, introducing an external conceptual scaffold before a business simulation improved post-game problem solving relative to both play-only and play-then-study conditions, although perceived learning declined. Flow and enjoyment did not differ, and in the play-then-study condition, flow predicted problem solving [60]. In an augmented-reality (AR) science board game, collaboration-oriented scaffolding increased antecedents of flow and perceived ease of use, whereas a competition-oriented design yielded greater learning effectiveness [61]. A randomized trial detected no overall advantage of adaptive over nonadaptive scaffolding because nonadaptive participants often received incidentally tailored support; exploratory analyses showed tailored information reduced cognitive load and shortened scenario time [62]. Active and contingent scaffolds interact with cognitive style to lower extraneous load and improve performance [63]. Adaptive scaffolding improves learning and strengthens cognitive and emotional engagement [64], and pupils report that diagnosis, hints, and explanations aid understanding and exploration [65].

Applications of scaffolding theory in serious games for primary acoustics remain underexplored. While findings converge across contexts, tasks, modalities, and age groups, effects remain contingent on scaffold design, including timing, explicitness, and adaptivity, and on implementation details.

3. Purpose

Although sound holds distinctive value in primary science education, its ephemeral and invisible qualities often lead to marginalization in curriculum implementation and instructional design. Recent scholarship recommends a shift from outcome-oriented to process-oriented instruction that explicitly represents sound through activities such as vocalization and visualization. This orientation enables learners to build a robust understanding of abstract concepts from experience by making them sound perceptible through visualization and related activities [3]. Building on this rationale, we developed a computer game titled SonicScape. Set in Echo Valley, the game casts the player as a mage who, guided by a white-fox guardian spirit, confronts a wizard whose spells have thrown the valley into disorder. To evaluate the game's potential to support learning about sound, this study compares three instructional conditions, namely SSG, NSG, and TRAD.

The present study had four aims:

1. To design and implement a voice-interactive 2D-platformer serious game in two versions (SSG and NSG) to teach primary pupils core sound concepts (loudness, pitch, echo);
2. To compare SSG, NSG, and TRAD on immediate post-instruction knowledge test performance under matched content;
3. To compare PL, flow, intrinsic motivation, and ECL across the three conditions;
4. To examine whether learning-experience measures predict knowledge test performance and whether the scaffolding condition (SSG vs. NSG) moderates these associations.

4. Materials and Methods

4.1. Learning Objectives and Constructs

The conceptual framework for this study is grounded in three sound constructs that feature prominently in primary science curricula. In China's Compulsory Education Science Curriculum Standards (2022) [66], the topic "Sound" falls within the matter science domain under the thematic unit "Propagation of Sound and Light" and is scheduled for Stage II (Grades 3–4). (1) Loudness (sound pressure level). Perceived loudness increases nonlinearly with root-mean-square (RMS) sound pressure. Levels are expressed on a logarithmic decibel scale (dB SPL), illustrating the nonlinear relation between physical magnitude and psychophysical intensity [67]. (2) Pitch (fundamental frequency). Pitch is the perceptual correlate of a waveform's fundamental frequency (f0) [68]. (3) Echo (temporal separation of reflections). Sound waves obey the law of reflection. When the round-trip path delay is approximately 0.1 s, the auditory system segregates direct and reflected components, and a discernible echo can be heard. This principle is widely used in architectural acoustics and in animal echolocation [69]. Taken together, these three constructs align with the core disciplinary ideas on sound articulated in the Next Generation Science Standards [70], and frame common alternative conceptions among primary pupils, thereby providing a robust theoretical foundation for the learning objectives in the present serious game study.

4.2. Instructional Conditions

4.2.1. Scaffolded Serious Game (SSG)

We designed SonicScape, a 2D platformer serious game that integrates narrative goals, mapping between mechanics and concepts, immediate feedback, and progressive challenges. The design targets three core sound concepts (loudness, pitch, and echo) and assigns each to specific tasks to form an actionable and transferable learning sequence. The narrative unfolds in Echo Valley, where the player, guided by a white-fox guardian spirit, assumes the role of a mage confronting a disruptive wizard. Each region foregrounds one acoustic concept and grants a corresponding ability for collecting note fragments; assembling all fragments completes the overall quest. As discussed in Section 2.1, this narrative framing supports meaning-making and situates the design within a narrative-based learning perspective.

To operationalize these goals, the three regions anchor both mechanics and learning objectives. For loudness (sound pressure level), vocalizing to meet or exceed a threshold triggers environmental responses and clears obstacles, and amplitude-dependent shock waves visualize the effect of sound energy on the environment. For pitch (fundamental frequency), real-time pitch is mapped to a traversable curvilinear terrain, enabling learners to perceive the correspondence between frequency and geometry through a cycle of vocalize, observe, and adjust. For echo (temporal separation of reflections), an Echo Reveal mechanic exposes hidden paths and obstacles and supports navigation under low visibility conditions. Levels are sequenced progressively so that initial tasks manipulate a single variable, intermediate tasks require two-factor coordination, and advanced tasks integrate multiple abilities; for example, players localize with echo and then shape terrain with pitch to obtain a fragment. This pacing calibrates difficulty to learner skill to sustain flow and to avoid losses associated with tasks that are either overly complex or trivial [71]. Figure 1 illustrates the progression.

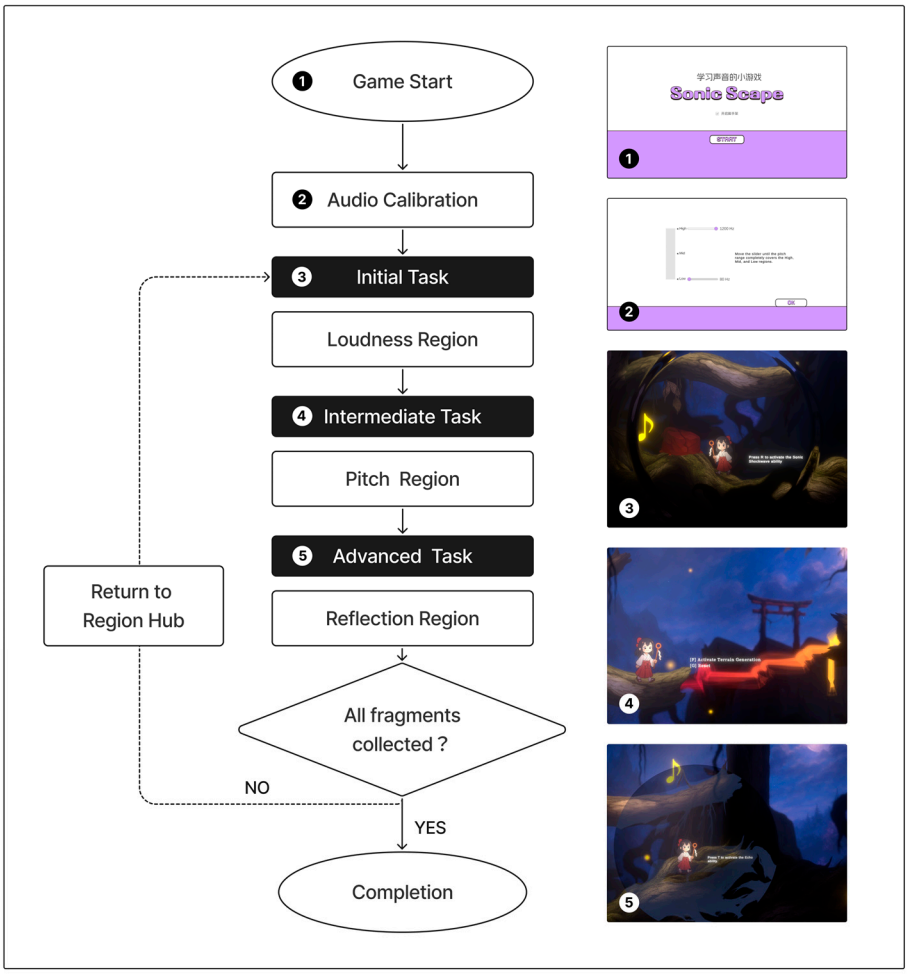


Figure 1. Progression of the game illustrated with interface screenshots. 1. Game start. 2. Audio calibration and microphone test: adjust the slider so that the player's voice level falls within the target range on the meter and confirm microphone input. 3. Initial task—Loudness: press R to activate the microphone and speak loudly to exceed the SPL threshold; amplitude-dependent shock waves visualize the effect of sound energy and clear obstacles. 4. Intermediate task—Pitch: press F to enable pitch control; raise or lower the pitch of your voice to map real-time pitch onto a traversable curvilinear terrain; press G to reset the terrain. 5. Advanced task—Echo: hold T to emit a sound; the Echo Reveal mechanic exposes hidden paths and obstacles to support navigation under low-visibility conditions.

In line with meta-analytic evidence that gamified learning yields small but significant improvements in cognitive, motivational, and behavioral outcomes and that specific design elements can moderate these effects, we implemented three feedback channels to support engagement and performance [50]. Sound visualization shows current pitch and stability as a linear trace. An amplitude-triggered shock wave with post-processing represents sound energy through radius and on-screen intensity for destruction and unlocking. Echo Reveal increases the visibility of a hidden layer as the amplitude rises and gradually exposes routes. These channels are paired with progress-based rewards and ability unlocks, including fragments, new abilities, and access to new areas. A green indicator confirms attainment of target values.

To guide performance without interrupting exploration, we implemented a three-tier scaffolding model that is often used to balance challenge and frustration in serious games [72]. Active scaffolds provide explicit goals, stepwise hints, and post-error explanations. Contingent scaffolds supply just-in-time prompts that respond to player actions. Implicit scaffolds integrate symbolic cues into the environment; for example, in dark ar-

eas, a glowing musical note suggests a navigable path. Figure 2 presents the three-tier scaffolding mechanisms alongside scenes in SSG and NSG conditions.



Figure 2. Three-tier scaffolding designs and game scenes (SSG = scaffolded serious game; NSG = non-scaffolded game).

At the implementation level, the project is built in Unity using the URP 2D rendering pipeline (Universal Render Pipeline). Scenes comprise foreground, midground, background, and character layers, with script-based parallax to enhance spatial depth. MicManager calls the Microphone API to capture audio, uses the fast Fourier transform (FFT) to compute spectra, and outputs CurrentAmplitude and CurrentPitch for use by the game mechanics. Pitch-to-terrain mapping relies on LineRenderer and 2D colliders to convert curves into traversable geometry in real time. Amplitude-driven shock waves use full-screen post-processing shaders to control radius and intensity. The Echo Reveal mechanism uses an auxiliary camera to render a hidden layer to a RenderTexture and blends it in the main camera's post-processing stage so that routes gradually appear as amplitude increases. All experimental sessions were run on PCs.

4.2.2. Non-Scaffolded Serious Game (NSG)

To isolate the contribution of scaffolding, we implemented a non-scaffolded version of the game (NSG) as a minimal guidance baseline condition. Relative to Section 4.2, all core elements were held constant, including regions and objectives, tasks and audio assets, the interface layout, and time on task. We then removed all scaffolding functions. Active scaffolds were disabled, so region entrances no longer displayed text or video instructions. Contingent scaffolds were disabled, so there were no frequency or amplitude hints, no demonstration trajectories after errors or unmet thresholds, and no green light confirmation when target values were reached. Implicit scaffolds were removed, so dark areas did not display a glowing musical note, and no path hints were embedded in environmental

mechanics. Only generic outcome feedback remained to preserve playability, and the proctor did not provide process guidance. This configuration isolates the net effect of scaffolding while minimizing media and dosage confounds. See Figure 2.

4.2.3. Traditional Hands-On Materials Instruction (TRAD)

In the control condition, pupils received a 30-minute traditional lesson using physical teaching aids that covered the same sound concepts (loudness, pitch, and echo). Loudness was demonstrated by contrasting soft and forceful drum strikes and by having pupils listen at multiple distances to illustrate distance-related attenuation. Pitch was taught by anchoring one end of a ruler to the desk and plucking the free end at varying overhang lengths to illustrate changes in frequency. Echo was demonstrated in a corridor or outdoor area by having pupils clap or call toward a distant reflective surface (e.g., a far wall or the end of a long hallway). A distinct repetition (echo) was perceived when the reflection was temporally separable from the direct sound (approximately 0.1 s), whereas shorter delays were described as reverberation. The lesson consisted of direct explanation and demonstration, followed by brief pupil participation, and no scaffolds or gamified elements were provided. Total duration and content coverage matched those of the game-based groups to ensure comparability. For completeness, Figure 3 provides a side-by-side photographic overview of the experimental setups across all three conditions (SSG, NSG, TRAD).



(**a**) (**b**) (**c**)

Figure 3. Experimental setups by condition: (**a**) SSG; (**b**) NSG; (**c**) TRAD (traditional hands-on materials instruction).

4.3. Measures

4.3.1. Knowledge Test Design

To evaluate learning outcomes across conditions, immediately after the SSG, NSG, and TRAD sessions, all participants completed an eight-item sound knowledge test comprising five single-best-answer multiple-choice items and three true or false items. The test used dichotomous scoring, with 1 indicating a correct response and 0 indicating an incorrect response. Items were aligned with three core constructs, namely loudness and amplitude, pitch and frequency, and reflection and echo. Distractors targeted common misconceptions. Two experts in primary science education reviewed and refined the stems and distractors to support content alignment with the three constructs. Item stems are presented below. Response options are omitted. Q1 Which phenomenon directly shows that sound is produced by vibration? Q2 When a drum is struck more forcefully, what happens to the sound? Q3 A whistle produces a sharp, high sound because the sound wave it generates _____. Q4 Which rubber band produces the lowest pitch? Q5 What is the repeated sound heard after shouting in a valley called? Q6 to Q8 were true or false items. Q6 Placing a hand on a working loudspeaker and feeling it vibrate indicates that sound originates from vibration. Q7 Plucking a rubber band more forcefully is louder because the amplitude is greater. Q8 When a sound wave meets a smooth rock wall, part of it is reflected, producing an echo. The mapping between items and constructs was as follows. Loudness

corresponded to Q1, Q2, Q6, and Q7. Pitch corresponded to Q3 and Q4 echo corresponded to Q5 and Q8. Full item stems with all response options and construct tags are provided in Appendix A (Table A1).

4.3.2. Questionnaire Design

Learning-experience measures comprised PL, flow, intrinsic motivation, and ECL. We used a 12-item questionnaire on a 5-point Likert-type scale with emoji anchors (1 = angry face, 5 = smiling face). Each dimension contained three items. PL used three items adapted from Cho et al. [73], modified from Cho et al. [74], and contextualized to the sound-science lesson (e.g., "I can clearly explain the sound concepts learned today"; Cronbach's α = 0.702). Flow was assessed with the reduced Flow Short Scale (r-FKS; 3 items—FKS items 6, 8, 9) indexing absorption and fluency (e.g., "I am totally absorbed in what I am doing") [75]. Consistent with flow theory, this scale treats absorption and fluency as a session-level state [36]. The FKS family is typically administered on a 7-point Likert scale; to harmonize all questionnaires and reduce burden for primary pupils, we administered the r-FKS on a 5-point format (1 = strongly disagree to 5 = strongly agree). Internal consistency in our sample was α = 0.71. Item scores were averaged to yield a flow-intensity index (higher scores = stronger flow). This adaptation is supported by evidence that 4–7 response categories optimize reliability/validity with minimal gains beyond seven [76], and by pediatric measurement reviews endorsing 5-point scales for primary pupils [77]. Intrinsic motivation was operationalized as perceived enjoyment, assessed with the three items introduced by Davis, Bagozzi, and Warshaw [78]. In Davis et al. [78], enjoyment is conceptualized as an instance of intrinsic motivation, whereas the perceived usefulness and perceived ease-of-use scales in that paper were drawn from Davis [79]. To adapt the enjoyment items for a primary-school learning context, we replaced the referent "using the system" with the item "I have fun doing this sound-based learning task." We also harmonized the response format from the original 7-point scale to a 5-point Likert-type agreement scale, consistent with our other instruments; no item meanings were changed. Internal consistency in our sample was Cronbach's α = 0.71. From an SDT perspective, we treat enjoyment as a proximal indicator of intrinsic motivation [40,41]; we did not directly assess basic psychological need satisfaction (autonomy, competence, relatedness). ECL was measured with the child-adapted three-item scale by Altmeyer et al. [80], derived from Klepsch et al. [81], using a 5-point response format appropriate for primary-school pupils (e.g., "In this sound-based learning task, the important information or cues were not obvious, making it difficult for me to know what to pay attention to."; items were negatively worded; Cronbach's α = 0.747). Consistent with CLT, these items capture avoidable processing demands introduced by the presentation and interaction format, rather than content difficulty. To ensure that ECL reflected extraneous rather than intrinsic demands, lesson content and duration were held constant across SSG, NSG, and TRAD, and standardized instructions were used. Intrinsic load was held constant by design and not measured; germane load was not measured. All scales used the same 5-point format (1–5). For comparability across constructs, ECL was reverse-scored (higher values = lower extraneous load), whereas higher values on PL, flow, and intrinsic motivation indicate higher levels of those constructs. The questionnaire was administered immediately after the session; the full instrument is provided in Appendix A, Table A2.

5. Procedures

Participants were recruited from the Jinghu District Youth Science Activity Center in Wuhu, China. We enrolled 51 third-grade pupils (N = 51). Both written parental consent and age-appropriate pupil assent were obtained on paper and signed before data collection.

A brief screening verified that none of the pupils had previously studied sound-related content. After excluding 6 invalid cases, the final sample comprised 45 pupils (27 boys, 18 girls; aged 8–9 years). Pupils were randomly assigned by lot to SSG, NSG, and TRAD (each N = 15).

Sessions ran on laptops equipped with RTX 3070 GPUs to ensure stable performance. All tasks were completed individually in quiet rooms to minimize external interference. Each group first received a brief introduction to basic principles of sound. Pupils in SSG then played the version with built-in hints; pupils in NSG used the same game without scaffolding; the TRAD group received teacher-led instruction using simple props (e.g., a drum and a ruler). Immediately afterward, all pupils completed a paper-based knowledge test on sound properties consisting of static images and questions printed on A4 paper. Procedures were explained in advance, and pupils responded individually.

After the test, pupils completed questionnaires assessing PL, flow, intrinsic motivation, and ECL using a 5-point Likert-type scale with emoji anchors. Following data screening, the valid analytic sample remained N= 45 (SSG N= 15, NSG N = 15, TRAD N = 15). All data were analyzed in IBM SPSS Statistics 27 to compare knowledge-test performance and questionnaire outcomes across conditions.

6. Results and Discussion

6.1. Knowledge Test Analysis

The eight-item knowledge test was blueprint-aligned to three target concepts (loudness, pitch, and echo) and refined through expert review; accordingly, we use the total score as a content-aligned outcome measure rather than a fully validated scale, and we report item-level difficulty and discrimination to support interpretation [82]. Internal consistency for the eight-item test was low in this sample (n = 45): KR-20 = 0.08 (α = 0.17 for standardized items). Given the brevity of the test and its coverage of multiple sub-constructs, this coefficient is a conservative estimate for a single composite; therefore, we provide item-level indices. Item difficulty (proportion correct) ranged from very easy (Q1 = 0.98; Q5 = 0.93; Q6 = 0.93) to typical/moderate (Q2 = 0.78; Q3 = 0.60; Q7 = 0.64; Q8 = 0.60) and relatively difficult (Q4 = 0.36). Discrimination (upper–lower difference in p) was strong for Q7 (0.688), Q8 (0.688), Q2 (0.563), and Q3 (0.525); marginal for Q5 (0.188) and Q4 (0.163); and poor for the very easy Q1 (0.063) and Q6 (0.063), where the lower group was already near ceiling ($p \approx 0.94$), leaving too little variance to differentiate learners. Accordingly, we interpret the total score as a heterogeneous index of mastery and rely on distribution-free tests for between-condition comparisons. Because two items (Q1 and Q6) were near-ceiling and showed poor discrimination, the composite total may be sensitive to these items; group contrasts should therefore be interpreted with caution.

Shapiro–Wilk tests indicated non-normality for knowledge scores in NSG (W = 0.734, p = 0.001) and TRAD (W = 0.880, p = 0.047). Given n = 15 per group and the bounded 0–8 scale, we used the Kruskal–Wallis H test to compare SSG, NSG, and TRAD. The omnibus result was significant, H(2) = 7.12, p = 0.028, ε^2 = 0.12 (medium).

SSG showed the strongest overall distribution (Mean Rank = 29.47; M = 6.47; Mdn = 7), TRAD was intermediate (22.20; 5.73; 6), and NSG was lowest (17.33; 5.27; 6). Full descriptive statistics appear in Table 1. The maximum score reached 8 in SSG and TRAD but only 6 in NSG, indicating a restricted upper range without scaffolds. A plausible mechanism is that SSG externalized goals, cues, and corrective feedback, reducing extraneous processing and focusing attention on test-relevant features. In contrast, NSG demanded greater self-regulation and strategy discovery, diverting effort to exploration and control. TRAD's linear explanations and demonstrations matched the test's static prompts, yielding comparable medians and ceilings but a lower overall distribution than SSG.

Table 1. Descriptive statistics and Kruskal–Wallis H test for the knowledge test across groups.

Group	N	Mean ± SD	Median (IQR)	Min–Max	Mean Rank
SSG	15	6.47 ± 1.30	7 (3)	4–8	29.47
NSG	15	5.27 ± 0.88	6 (2)	4–6	17.33
TRAD	15	5.73 ± 1.03	6 (1)	4–8	22.20

Notes: SSG, scaffolded serious game; NSG, non-scaffolded serious game; TRAD, traditional hands-on materials instruction.

Post hoc Dunn–Bonferroni comparisons confirmed these patterns. SSG outperformed NSG (Mean Rank Diff. = 12.13, Z = 2.651, p_adj = 0.024, r = 0.48, 95% CI [0.15, 0.72]). In contrast, SSG vs. TRAD (7.27, Z = 1.588, p_adj = 0.337, r = 0.29, 95% CI [−0.08, 0.59]) and NSG vs. TRAD (−4.87, Z = −1.063, p_adj = 0.863, r = −0.19, 95% CI [−0.52, 0.18]) were not significant. This pattern is consistent with scaffolding theory. In SSG, just-in-time supports channel attention and time on target toward assessed concepts, whereas in NSG, their absence increases self-regulatory demands and cognitive load. Detailed results are presented in Table 2.

Table 2. Post hoc pairwise comparisons (Dunn–Bonferroni; mean ranks).

Comparison	Mean Rank Diff.	SE	*Z*	*p*	Bonferroni *p*	*r*	*r* [95%CI]
SSG vs. NSG	12.133	4.577	2.651	0.008	0.024 *	0.48	[0.15, 0.72]
SSG vs. TRAD	7.267	4.577	1.588	0.112	0.337	0.29	[−0.08, 0.59]
NSG vs. TRAD	−4.867	4.577	−1.063	0.288	0.863	−0.19	[−0.52, 0.18]

Notes: Two-sided tests. "Bonferroni p" are Bonferroni-adjusted p-values; asterisks mark significance after adjustment (* $p < 0.05$). Mean-rank difference is (first−second).

Overall, the pattern is clear and indicates that scaffolding yields measurable benefits. The scaffolded game outperformed the non-scaffolded version and performed on par with traditional instruction on the immediate knowledge test. A plausible mechanism is that scaffolds externalize goals, cues, and corrective feedback, which reduces extraneous load and directs attention to test-relevant features. In contrast, the non-scaffolded version imposes greater self-regulatory and strategic demands and can therefore cap the upper tail. Traditional instruction aligns closely with the test format through linear explanations and demonstrations, which accounts for comparable medians and ceilings despite a lower overall distribution than SSG.

6.2. Learning Experience

Beyond the knowledge test, we compared learning-experience outcomes across conditions, Shapiro–Wilk tests showed departures from normality for at least one group on each outcome (PL: TRAD W = 0.872, p = 0.036; Flow: NSG W = 0.850, p = 0.018; intrinsic motivation: NSG W = 0.793, p = 0.003; ECL: TRAD W = 0.871, p = 0.035); therefore Kruskal–Wallis H tests were used for between-condition comparisons (Table 3). Overall, only intrinsic motivation differed significantly across groups, with NSG exceeding SSG, whereas PL, flow, and ECL showed no significant differences. TRAD generally fell between the two game conditions. This pattern suggests that greater autonomy in NSG may foster stronger intrinsic motivation.

Table 3. Learning experience outcomes across conditions (Kruskal–Wallis on mean ranks).

Measure	SSG Mean Rank	NSG Mean Rank	TRAD Mean Rank	H(2)	*p*	ε^2 [95% CI]	Post hoc (Bonf.)
PL	17.87	26.33	24.80	3.639	0.162	004 [0.00, 0.20]	—
Flow	21.77	26.27	20.97	1.476	0.478	0.00 [0.00, 0.09]	—
Intrinsic motivation	15.73	28.83	24.43	8.072	0.018 *	0.14 [0.00, 0.35]	NSG > SSG (Bonferroni p = 0.015 *, r = 0.51, 95% CI [0.18, 0.74])
ECL	26.63	19.53	22.83	2.233	0.327	0.01 [0.00, 0.14]	—

Notes: PL, perceived learning; ECL, extraneous cognitive load. Kruskal–Wallis tests on mean ranks; Dunn–Bonferroni–adjusted post hoc where applicable. Effect sizes: ε^2 (omnibus; 95% CIs via BCa bootstrap, within-group resampling, B = 5000) and r (pairwise; 95% CIs via Fisher z). * $p < 0.05$. ECL is reverse-scored (higher ECL = lower extraneous load).

Analysis of PL indicated that the omnibus Kruskal–Wallis H test was not significant, H(2) = 3.64, p = 0.162, ε^2 = 0.04. Mean ranks were NSG (26.33) > TRAD (24.80) > SSG (17.87), as reported in Table 3. Descriptively, NSG exceeded SSG on PL, possibly reflecting the greater autonomy afforded by NSG (open exploration and self-paced strategy selection), which some pupils may have interpreted as greater learning, whereas SSG's step-by-step guidance may have reduced perceived effort despite smoother task progress. Given the nonsignificant omnibus result, this interpretation is tentative. From a Cognitive Load Theory perspective, two mechanisms may account for this pattern. First, an expertise-reversal account: as schemas develop, highly explicit guidance can become partly redundant, dampening the subjective sense of "figuring things out." Second, a variability effect: the less scripted NSG may have afforded more varied problem states and self-directed strategy choices, supporting schema construction and elevating perceived learning. Even so, the observed trend is consistent with prior work [60], suggesting that stronger or preemptive scaffolding can depress perceived learning.

Flow did not differ significantly across groups, H(2) = 1.48, p = 0.478, $\varepsilon^2 \approx 0.00$. Summary statistics appear in Table 3. Consistent with Barzilai and Blau [60], who reported negligible differences across scaffolding conditions (p = 0.990, $\eta^2 \approx 0$), flow appears to reflect individual immersion and concentration rather than being determined directly by scaffolding. In our short, single-session task, the descriptive ordering favored NSG > SSG > TRAD, but the group-level effect was near zero. NSG provided challenge and autonomy, whereas SSG provided clear goals and immediate feedback; these distinct antecedents of flow likely offset one another across conditions. Under flow theory, state flow depends on a balance between challenge and skill, supported by clear goals and immediate feedback; our conditions likely equated these antecedents [38]. A brief acclimation period and n = 15 per group may also have compressed between-group differences.

For intrinsic motivation, a Kruskal–Wallis H test showed a significant group effect, H(2) = 8.07, p = 0.018, ε^2 = 0.14. The corresponding descriptive statistics and post hoc results are reported in Table 3. Dunn–Bonferroni post hoc comparisons indicated NSG > SSG (adjusted p = 0.015, r = 0.51), whereas TRAD did not differ from either group (all adjusted p-values > 0.05). Relative to SSG, NSG afforded choices over path, strategy, and pace, plausibly strengthening perceived autonomy and competence. By contrast, SSG's dense, preemptive scaffolds, while easing task execution, may have been experienced as controlling, thereby reducing perceived autonomy and dampening curiosity-driven engagement, an interpretation consistent with SDT, which posits the basic psychological needs of autonomy, competence, and relatedness as the basis of intrinsic motivation [40,41]. This pattern contrasts with reports that adaptive, within-task scaffolding can increase interest and persistence [64,65]. A plausible explanation is that our scaffolds were pro-

cedural and highly controlling across three tiers (active, contingent, implicit), which at times reduced choice and agency and misaligned with pupils' immediate needs, thereby interrupting exploration and flow. Hence, the divergence likely hinges on the degree of control, adaptivity, and timing, rather than on the number of scaffold layers. Beyond guidance control, differences in intrinsic motivation across SSG, NSG, and TRAD may also reflect profile–design fit. Prior work in engineering education shows that instructors' player profiles often diverge from those typically attributed to students, implying that designs can unintentionally mirror designer preferences rather than learner needs; misalignment of this kind may depress intrinsic motivation for some learners [25]. Although we did not measure player profiles and our sample comprised primary pupils, we treat profile-aligned scaffolding as a plausible boundary condition and a target for future tests.

ECL (reverse-scored; higher values indicate lower load) did not differ significantly across conditions, $H(2) = 2.23$, $p = 0.327$, $\varepsilon^2 = 0.01$. Descriptively, reverse-scored ECL was highest in SSG and lowest in NSG (SSG > TRAD > NSG). Group-level statistics are reported in Table 3. Although prior studies suggest that adaptive or well-integrated scaffolding can reduce ECL [62,63], no group effect emerged here. To unpack this null result, we consider CLT-consistent mechanisms across conditions. SSG: pop-up or side-panel hints that are spatially separated from task elements can induce split-attention; frequent or ill-timed prompts act as micro-interruptions; and auto-fading overlays impose transient-information demands, together adding search, integration, and re-orientation costs. NSG: While prompt-induced split-attention/transience costs are absent, learners face higher navigation and visual-search demands (self-directed exploration, locating functions, trial-and-error), which are extraneous to the learning goal. TRAD: attention must be coordinated across teacher talk, worksheets, and physical apparatus; verbal explanations are often short-lived (another form of transient information), though slower pacing and stable layouts may partially offset switching costs. Taken together, distinct sources of extraneous load likely offset one another across conditions, yielding similar group means. Accordingly, the integration, persistence, spatial contiguity, and timing of scaffolds—rather than their sheer number—are key to their net effect on ECL in this setting.

6.3. Learning Experience Hierarchical Regression Scaffolding × Intrinsic Motivation

Intrinsic motivation and ECL were z-standardized. Scaffold was coded SSG = 1, NSG = 0. PL and flow were excluded to avoid post-treatment bias. VIFs were ≈ 5 (all < 10).

The main-effects model explained 31% of the variance, $R^2 = 0.31$, adj. $R^2 = 0.23$, $F(3, 26) = 3.87$, $p = 0.021$, as shown in Table 4. Scaffold was a significant and practically meaningful predictor ($b = 1.48$, $p = 0.004$), indicating that the SSG group scored on average 1.48 points higher on the 0–8 knowledge test. This supports prior findings that explicit conceptual scaffolds improve posttest performance without undermining enjoyment or flow [60].

Table 4. Knowledge Test regression results (main-effects model).

Predictor	*B*	*SE B*	*β*	95% CI for *B*
Intrinsic motivation (z)	0.38	0.24	0.30	[−0.11, 0.87]
ECL (z)	0.13	0.22	0.11	[−0.32, 0.58]
Scaffold (SSG = 1, NSG = 0)	1.48 **	0.48	0.60	[0.49, 2.47]

Notes: Two-sided *p*. *B* unstandardized (*β* standardized); predictors z-standardized; outcome 0–8; scaffold: SSG = 1, NSG = 0. Model fit: $R^2 = 0.31$; adj. $R^2 = 0.23$; $F(3,26) = 3.87$, $p = 0.021$; $N = 30$. ** $p < 0.01$. ECL is reverse-scored (higher ECL = lower extraneous load).

Adding the intrinsic motivation × Scaffold interaction improved fit ($\Delta R^2 = 0.11$, F_change(1, 25) = 4.70, $p = 0.040$), yielding a full model of $R^2 = 0.42$, adj. $R^2 = 0.33$; F(4,

25) = 4.49, $p = 0.007$ (Table 5). Scaffold remained significant (B = 1.41, $p = 0.004$), and the interaction was significant (B = 1.03, $p = 0.040$). Simple-slope probes showed that in NSG, the intrinsic motivation was nonsignificant, whereas in SSG, a 1-SD increase in intrinsic motivation predicted ≈ 0.73 higher points (≈9% of the 0–8 scale). This motivation-by-structure synergy is consistent with evidence that clear goals and immediate feedback amplify the benefits of intrinsic motivation [52]. From an SDT perspective, overly explicit, preemptive prompts may be experienced as controlling, potentially dampening enjoyment, whereas contingent and implicit supports can preserve autonomy while bolstering competence [40,41].

Table 5. Knowledge Test regression results (interaction model).

Predictor	*B*	*SE B*	*β*	95% CI for *B*
Intrinsic motivation (*z*)	−0.30	0.38	−0.24	[−1.08, 0.49]
ECL (*z*)	0.00	0.21	0.00	[−0.43, 0.43]
Scaffold (SSG = 1, NSG = 0)	1.41 **	0.45	0.57	[0.48, 2.34]
Intrinsic motivation (z) × Scaffold	1.03 *	0.48	0.61	[0.04, 2.02]

Notes: Full model: $R^2 = 0.42$, adj. $R^2 = 0.33$, $F(4, 25) = 4.49$, $p = 0.007$; $\Delta R^2 = 0.11$, F_change (1, 25) = 4.70, $p = 0.040$; $N = 30$. (Parallel model with ECL × Scaffold: $\Delta R^2 = 0.02$, $p = 0.425$; max VIF ≈ 5.) * $p < 0.05$; ** $p < 0.01$. ECL is reverse-scored (higher ECL = lower extraneous load).

ECL was not significant in either model. A plausible account is that scaffolding reduced extraneous processing by externalizing goals, cues, and corrective feedback. Overall, the results fit self-determination and cognitive-load perspectives. Structured scaffolds focus attention on target criteria so that intrinsic motivation is more effectively converted into strategy use and accurate encoding, as summarized in Tables 4 and 5. Consistent with the group-level analysis, reductions in extraneous processing from externalizing goals and cues were likely counterbalanced by split-attention and transient-information costs arising from the presentation, frequency, and timing of pop-up or on-screen prompts in this single-session task [24,32,83], yielding a near-zero ECL effect.

7. Limitations and Future Work

This study has several limitations that qualify the interpretation and generalizability of the findings. First, we did not administer a pretest or other baseline measures. Although pupils were randomly assigned and had no prior formal instruction on sound (screened at enrollment), group equivalence before the intervention cannot be verified; therefore, posttest differences may partly reflect unmeasured pre-existing disparities rather than the intervention alone. Second, the small group sizes (n = 15) limit statistical power and yield wide confidence intervals, increasing the risk of Type II errors and reducing the precision and replicability of statistically significant effects; effect estimates should therefore be viewed as imprecise. Third, potential digital novelty cannot be ruled out. As noted in Section 2.1, voice-interactive gameplay may transiently elevate engagement; our single-session design does not disentangle novelty from instructional mechanisms, so novelty could plausibly contribute to both learning and motivational differences. Fourth, we employed only an immediate posttest and did not include delayed retention or transfer assessments; thus, the durability and generalization of gains remain unknown. Fifth, the single-session, single-site setting (PC laptops in one science activity center) constrains ecological validity; results may not extend to longer classroom implementations, other age groups, or different hardware/software environments.

Additional measurement limitations apply. The 8-item knowledge test prioritized content coverage over internal consistency and produced a low KR-20. In addition, two items (Q1 and Q6) were near-ceiling and poorly discriminating, which may have influ-

enced the composite total. The 3-item, child-adapted ECL scale and other self-reports may underdetect brief, prompt-induced costs at the moment level; and intrinsic and germane cognitive load were not measured, so we refrain from claims about load components. In addition, we did not measure player profiles (e.g., Bartle/HEXAD), which may moderate the motivational impact of gamified elements; consequently, profile-by-condition interactions (SSG/NSG/TRAD) remain untested in our data [25]. Finally, scaffolds were pre-designed and event-triggered rather than learner-modeled; explicit prompts may at times have been overly controlling, which could dampen autonomy for some pupils. Generalizability is limited. Inferences should be constrained to immediate, short-term outcomes observed under conditions similar to those studied (single session, single site, small groups, immediate posttest with a short instrument, and no pretest). Digital novelty associated with voice interaction cannot be ruled out as a competing explanation.

Future work will incorporate: (i) pretesting (content knowledge and relevant covariates) with blocked/stratified randomization and/or covariate adjustment to strengthen causal inference; (ii) larger samples based on a priori power analyses; (iii) repeated-exposure protocols with delayed retention and near- and far-transfer tests (or crossover designs) to separate novelty from instructional effects; (iv) multisite deployments across classrooms and devices to improve external validity; (v) richer process measures (e.g., log-based indicators of search/switching, step-level mental-effort checks) to capture transient extraneous load; and (vi) autonomy-preserving, adaptive, and fading scaffolds (integrated, in situ, learner-controlled) aligned with CLT and SDT. (vii) include a child-appropriate player-profile instrument and test Profile $\times$ Condition (SSG/NSG/TRAD) interactions on intrinsic motivation and performance; complement these analyses with log-based indicators (e.g., optional-prompt use, path choices).

8. Conclusions

The present work makes four original contributions. (1) We introduce a voice-interactive 2D platformer that operationalizes three core acoustics constructs as concrete mechanics—amplitude-triggered shockwaves for loudness, pitch-to-terrain mapping for fundamental frequency, and an echo-reveal for reflections—together with a reusable three-tier software-scaffolding controller (active, contingent, implicit) and an implementable microphone $\rightarrow$ FFT $\rightarrow$ RMS/$f_0$$\rightarrow$mechanics pipeline in Unity. (2) We provide a classroom-feasible randomized comparison that benchmarks a scaffolded game against an otherwise identical non-scaffolded version and a hands-on lesson under matched content and time, enabling like-for-like inference in primary classrooms. (3) We identify evidence consistent with a motivation-by-structure moderation, tested within the game arms only (SSG vs. NSG; N = 30), whereby scaffolding strengthens the positive link between intrinsic motivation and immediate knowledge scores ($\Delta R^2 \approx 0.11$), while ECL is non-predictive in this setting. (4) We contribute a compact, replicable evaluation kit comprising an eight-item, blueprint-aligned knowledge check (not a validated scale) with item-level difficulty and discrimination statistics, alongside learner-centered scales (Appendix A) to support replication. All empirical claims are interpreted cautiously in light of the low internal consistency (KR-20) of this brief test; given the test's brevity and heterogeneity, between-group inferences prioritize non-parametric comparisons and effect sizes rather than scale-score precision.

In a single session implementation at one extracurricular site (N = 45; n = 15 per group) with an immediate eight-item posttest, the scaffolded voice-interactive game (SSG) produced the strongest score distribution, significantly outperforming the non-scaffolded version (NSG; medium effect, $r \approx 0.48$, 95% CI [0.15, 0.72]) and performing on par with traditional hands-on instruction (TRAD). Across learner-experience measures, only intrinsic motivation differed reliably (NSG > SSG), whereas perceived learning, flow, and reverse-

scored ECL (higher values = lower extraneous load) showed no between-group differences. Hierarchical regression likewise indicated a motivation-by-structure interaction (SSG vs. NSG), whereas ECL remained non-predictive.

Interpreted through SDT, competence-supportive and autonomy-preserving scaffolds may, in this short, single-session context, help convert enjoyment—a proximal indicator of intrinsic motivation—into more accurate encoding [40,41]. Consistent with flow theory, the near-zero group-level differences in flow are compatible with a comparable challenge–skill balance across SSG and NSG when goals and feedback are clear [36]. From a CLT perspective, the absence of between-condition differences in ECL and its non-predictive relation to scores do not provide evidence that SSG's advantage over NSG was driven by reduced extraneous demands; distinct sources of extraneous load may have offset one another in this brief task. Because intrinsic load was held constant by design and germane load was not measured, we refrain from claims about load-component mechanisms.

Overall, the present study offers an exportable construct-to-mechanic mapping for teaching sound, a ready-to-use audio-to-mechanics engineering pipeline, and a benchmarked evaluation blueprint that lowers the barrier to running controlled classroom trials in primary science.

Author Contributions: Conceptualization, M.C. and H.L.; Data curation, M.C.; Formal analysis, M.C.; Investigation, Z.C.; Methodology, M.C.; Project administration, M.C.; Resources, Z.C.; Software, H.L.; Validation, M.C.; Visualization, M.C.; Writing—original draft preparation, M.C.; Writing—review and editing, M.C.; Supervision, Q.L. and N.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board of the School of Design and Art, Shenyang Aerospace University (approval date: 19 June 2025).

Informed Consent Statement: Written informed consent was obtained from the parents/legal guardians of all child participants, and age-appropriate assent was obtained prior to participation.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy considerations for child participants and institutional ethical restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Sound Knowledge Test—Items and Response Options.

Q	Construct	Item Stem (Unchanged)	A	B	C	D	Key
1	Loudness	Which phenomenon directly shows that sound is produced by vibration?	A light bulb glowing	A guitar string vibrating	A rainbow appearing	Water heating up	B
2	Loudness	When a drum is struck more forcefully, what happens to the sound?	Quieter	Louder	Higher pitch	Lower pitch	B
3	Pitch	A whistle produces a sharp, high sound because the sound wave it generates _____.	Higher frequency	Greater amplitude	Lower frequency	Slower propagation	A
4	Pitch	Which rubber band produces the lowest pitch?	Thin and short	Thick and long	Thin and long	Thick and short	B
5	Echo	What is the repeated sound heard after shouting in a valley called?	Noise	Fog	Echo	Thunder	C

Table A1. *Cont.*

Q	Construct	Item Stem (Unchanged)	A	B	C	D	Key
6	Loudness	Placing a hand on a working loudspeaker and feeling it vibrate indicates that sound originates from vibration.	True	False	—	—	True
7	Loudness	Plucking a rubber band more forcefully is louder because the amplitude is greater.	True	False	—	—	True
8	Echo	When a sound wave meets a smooth rock wall, part of it is reflected, producing an echo.	True	False	—	—	True

Table A2. Item-level inventory (sources, adaptations, anchors, scoring).

Construct	ID	Item (EN)	Source(s)	Notes	Response Format and Anchors	Scoring Direction
PL	1	I feel that I learned a lot about sound in this lesson.	Cho et al. [73]; adapted via Cho et al. [74]	Retained original five-point Likert format; contextualized to a sound-based lesson; child-friendly wording	Items were rated on a five-point Likert scale with emoji-aligned anchors to aid on-screen comprehension: 1 = strongly disagree 😠, 2 = disagree ☹️, 3 = neither agree nor disagree 😐, 4 = agree 🙂, 5 = strongly agree 😄. analyses used numeric responses.	Positive
	2	I can clearly explain the sound concepts learned today.	[73,74]	As above	As above	Positive
	3	I can apply this knowledge to other science activities.	[73,74]	As above	As above	Positive
Flow	1	I am totally absorbed in what I am doing.	Bartholomeyczik et al. [75]	7→5 points; kept semantics; child-friendly anchors	As above	Positive
	2	I know what to do at each step.	[75]	As above	As above	Positive
	3	I feel that everything is under control.	[75]	As above	As above	Positive
Intrinsic motivation	1	I find this sound-based learning task enjoyable.	Davis, Bagozzi, and Warshaw (1992) [78]	Referent changed to sound-based learning task; 7 → 5 points; child-friendly anchors; meanings unchanged	As above	Positive
	2	I find the process of doing this sound-based learning task pleasant.	[78]	As above	As above	Positive
	3	I have fun doing this sound-based learning task.	[78]	As above	As above	Positive
ECL	1	In this sound-based learning task, the important information or cues were not obvious, making it difficult for me to know what to pay attention to.	Altmeyer et al. (child-adapted) [80], derived from Klepsch et al. (adult) [81]	Child-adapted ECL; facet = hidden; negatively worded	As above	Negative; reverse-score
	2	In this sound-based learning task, the key information or cues were not easy to see or hear, making it difficult for me to complete the task.	Ref. [80] (child), derived from [81] (adult)	Child-adapted ECL; facet = not visible/not audible; negatively worded	As above	Negative; reverse-score
	3	In this sound-based learning task, the prompts and steps were unclear, leaving me unsure about what to do next.	Ref. [80] (child), derived from [81] (adult)	Child-adapted ECL; facet = presentation-made-it-difficult; negatively worded	As above	Negative; reverse-score

References

1. Wang, T.; Berlin, D. Construction and Validation of an Instrument to Measure Taiwanese Elementary Students' Attitudes toward Their Science Class. *Int. J. Sci. Educ.* **2010**, *32*, 2413–2428. [CrossRef]
2. Chang, H.; Chen, J.; Guo, C.; Chen, C.; Chang, C.; Lin, S.; Su, W.; Lain, K.; Hsu, S.; Lin, J.; et al. Investigating Primary and Secondary Students' Learning of Physics Concepts in Taiwan. *Int. J. Sci. Educ.* **2007**, *29*, 465–482. [CrossRef]
3. Veith, S.I. What's the Matter with Sound? - How Primary School Students Perceive the Nature of Sound. *Res. Sci. Educ.* **2023**, *53*, 919–934. [CrossRef]
4. DeWitt, J.; Archer, L.; Osborne, J. Science-Related Aspirations Across the Primary–Secondary Divide: Evidence from Two Surveys in England. *Int. J. Sci. Educ.* **2014**, *36*, 1609–1629. [CrossRef]
5. Clark, D.B.; Tanner-Smith, E.E.; Killingsworth, S.S. Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Rev. Educ. Res.* **2016**, *86*, 79–122. [CrossRef]
6. Martin, D.J.; Jean-Sigur, R.; Schmidt, E. Process-Oriented Inquiry—A Constructivist Approach to Early Childhood Science Education: Teaching Teachers to Do Science. *J. Elem. Sci. Edu.* **2005**, *17*, 13–26. [CrossRef]
7. Aygün, M.; Hacıoğlu, Y. Teaching the Sound Concept: A Review of Science and Physics Education Postgraduate Theses in Turkey. *AJE* **2022**, *9*, 257–276. [CrossRef]
8. Yerdelen, S.; Sungur, S. Pre-Service Science Teachers' Conceptions of Sound: The Role of Task Value Beliefs. *Sci. Educ. Int.* **2020**, *31*, 295–303. [CrossRef]
9. Eshach, H.; Schwartz, J.L. Sound Stuff? Naïve Materialism in Middle-school Students' Conceptions of Sound. *Int. J. Sci. Educ.* **2006**, *28*, 733–764. [CrossRef]
10. Watt, D.; Russell, T. *Primary SPACE Research Reports: Sound*; Liverpool University Press: Liverpool, UK, 1990; ISBN 978-0-85323-456-2.
11. Xu, S.; Reiss, M.J.; Lodge, W. The Development of an Analytical Model for Science Classroom Creativity in China. *Res. Sci. Technol. Educ.* **2024**, *43*, 996–1021. [CrossRef]
12. Zhang, B.; Krajcik, J.S.; Sutherland, L.M.; Wang, L.; Wu, J.; Qian, Y. Opportunities and Challenges of China's Inquiry-Based Education Reform in Middle and High Schools: Perspectives of Science Teachers and Teacher Educators. *Int. J. Sci. Math. Educ.* **2005**, *1*, 477–503. [CrossRef]
13. Eramo, G.; Pastore, S.; De Tullio, M.; Rossini, V.; Monno, A.; Mesto, E. The Sound of Science: A Sonification Learning Experience in an Italian Secondary School. *Front. Educ.* **2025**, *9*, 1502396. [CrossRef]
14. Guiotto Nai Fovino, L.; Zanella, A.; Di Mascolo, L.; Ginolfi, M.; Carpita, N.; Trovato Manuncola, F.; Grassi, M. Evaluating the Effectiveness of Sonification in Science Education Using Edukoi. *Pers. Ubiquit. Comput.* **2024**, *28*, 693–711. [CrossRef]
15. Xiao, Y.; Jiang, C. Conceptual Change in Preschool Science Education: Evaluating a Serious Game Designed with Image Schemas for Teaching Sound Concept. In Proceedings of the International Conference on Human-Computer Interaction, Copenhagen, Denmark, 19–24 July 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 503–520.
16. Koç, A.; Kanadlı, S. Effect of Interactive Learning Environments on Learning Outcomes in Science Education: A Network Meta-Analysis. *J. Sci. Educ. Technol.* **2025**, *34*, 681–703. [CrossRef]
17. Lestari, D.P.; Supahar; Paidi; Suwarjo; Herianto. Effect of Science Virtual Laboratory Combination with Demonstration Methods on Lower-Secondary School Students' Scientific Literacy Ability in a Science Course. *Educ. Inf. Technol.* **2023**, *28*, 16153–16175. [CrossRef]
18. Chen, S.; Jamiatul Husnaini, S.; Chen, J.-J. Effects of Games on Students' Emotions of Learning Science and Achievement in Chemistry. *Int. J. Sci. Educ.* **2020**, *42*, 2224–2245. [CrossRef]
19. Chiang, T.H.; Yang, S.J.; Hwang, G.-J. An Augmented Reality-Based Mobile Learning System to Improve Students' Learning Achievements and Motivations in Natural Science Inquiry Activities. *J. Educ. Technol. Soc.* **2014**, *17*, 352–365.
20. Sung, H.-Y.; Hwang, G.-J. A Collaborative Game-Based Learning Approach to Improving Students' Learning Performance in Science Courses. *Comput. Educ.* **2013**, *63*, 43–51. [CrossRef]
21. Hamari, J.; Shernoff, D.J.; Rowe, E.; Coller, B.; Asbell-Clarke, J.; Edwards, T. Challenging Games Help Students Learn: An Empirical Study on Engagement, Flow and Immersion in Game-Based Learning. *Comput. Hum. Behav.* **2016**, *54*, 170–179. [CrossRef]
22. Volejnikova-Wenger, S.; Andersen, P.; Clarke, K.-A. Student Nurses' Experience Using a Serious Game to Learn Environmental Hazard and Safety Assessment. *Nurse Educ. Today* **2021**, *98*, 104739. [CrossRef]
23. Fanfarelli, J.R.; Vie, S. Medulla: A 2D Sidescrolling Platformer Game That Teaches Basic Brain Structure and Function. *Well Play.* **2015**, *4*, 7–29.
24. Kiili, K. Digital Game-Based Learning: Towards an Experiential Gaming Model. *Internet High. Educ.* **2005**, *8*, 13–24. [CrossRef]
25. Vergara, D.; Antón-Sancho, Á.; Fernández-Arias, P. Player Profiles for Game-based Applications in Engineering Education. *Comp. Applic. Eng.* **2023**, *31*, 154–175. [CrossRef]
26. Lester, J.C.; Spires, H.A.; Nietfeld, J.L.; Minogue, J.; Mott, B.W.; Lobene, E.V. Designing Game-Based Learning Environments for Elementary Science Education: A Narrative-Centered Learning Perspective. *Inf. Sci.* **2014**, *264*, 4–18. [CrossRef]

27. Meluso, A.; Zheng, M.; Spires, H.A.; Lester, J. Enhancing 5th Graders' Science Content Knowledge and Self-Efficacy through Game-Based Learning. *Comput. Educ.* **2012**, *59*, 497–504. [CrossRef]
28. Tene, T.; Vique López, D.F.; Valverde Aguirre, P.E.; Cabezas Oviedo, N.I.; Vacacela Gomez, C.; Bellucci, S. A Systematic Review of Serious Games as Tools for STEM Education. *Front. Educ.* **2025**, *10*, 1432982. [CrossRef]
29. Yazicioglu, S.; Çavus Güngören, S. Game-Based Activities Related to Light and Sound Unit and Students' Views. *J. Inq. Based Act.* **2021**, *11*, 51–68.
30. Prensky, M. *Digital Game-Based Learning*; McGraw-Hill: New York, NY, USA, 2001; ISBN 978-0-07-136344-0.
31. Wood, D.; Bruner, J.S.; Ross, G. The Role of Tutoring in Problem Solving. *J. Child Psychol. Psychiatry* **1976**, *17*, 89–100. [CrossRef]
32. Plass, J.L.; Homer, B.D.; Kinzer, C.K. Foundations of Game-Based Learning. *Educ. Psychol.* **2015**, *50*, 258–283. [CrossRef]
33. Quintana, C.; Reiser, B.J.; Davis, E.A.; Krajcik, J.; Fretz, E.; Duncan, R.G.; Kyza, E.; Edelson, D.; Soloway, E. A Scaffolding Design Framework for Software to Support Science Inquiry. In *Scaffolding*; Psychology Press: New York, NY, USA, 2018; pp. 337–386.
34. Bruner, J. Child's Talk: Learning to Use Language. *Child Lang. Teach. Ther.* **1985**, *1*, 111–114. [CrossRef]
35. Saye, J.W.; Brush, T. Scaffolding Critical Reasoning about History and Social Issues in Multimedia-Supported Learning Environments. *ETR&D* **2002**, *50*, 77–96. [CrossRef]
36. Csikszentmihalyi, M.; Csikzentmihaly, M. *Flow: The Psychology of Optimal Experience*; Harper & Row: New York, NY, USA, 1990; Volume 1990.
37. Sweller, J.; Van Merrienboer, J.J.; Paas, F.G. Cognitive Architecture and Instructional Design. *Educ. Psychol. Rev.* **1998**, *10*, 251–296. [CrossRef]
38. Paas, F.; Renkl, A.; Sweller, J. Cognitive Load Theory and Instructional Design: Recent Developments. *Educ. Psychol.* **2003**, *38*, 1–4. [CrossRef]
39. Sweller, J. Cognitive Load Theory. In *Psychology of Learning and Motivation*; Elsevier: Amsterdam, The Netherlands, 2011; Volume 55, pp. 37–76. ISBN 0079-7421.
40. Ryan, R.M.; Deci, E.L. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemp. Educ. Psychol.* **2000**, *25*, 54–67. [CrossRef]
41. *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*; Ryan, R.M., Deci, E.L., Eds.; Guilford Press: New York, NY, USA, 2017; ISBN 978-1-4625-3896-6.
42. Abt, C.C. *Serious Games*; Viking Press: New York, NY, USA, 1970; ISBN 978-0-670-63490-3.
43. Espinosa-Curiel, I.E.; Pozas-Bogarin, E.E.; Martínez-Miranda, J.; Pérez-Espinosa, H. Relationship Between Children's Enjoyment, User Experience Satisfaction, and Learning in a Serious Video Game for Nutrition Education: Empirical Pilot Study. *JMIR Serious Games* **2020**, *8*, e21813. [CrossRef]
44. Holmes, J.B.; Gee, E.R. A Framework for Understanding Game-Based Teaching and Learning. *On The Horiz.* **2016**, *24*, 1–16. [CrossRef]
45. Admiraal, W.; Huizenga, J.; Akkerman, S.; Dam, G.T. The Concept of Flow in Collaborative Game-Based Learning. *Comput. Hum. Behav.* **2011**, *27*, 1185–1194. [CrossRef]
46. Asadzadeh, A.; Shahrokhi, H.; Shalchi, B.; Khamnian, Z.; Rezaei-Hachesu, P. Serious Educational Games for Children: A Comprehensive Framework. *Heliyon* **2024**, *10*, e28108. [CrossRef]
47. Chang, C.-C.; Liang, C.; Chou, P.-N.; Lin, G.-Y. Is Game-Based Learning Better in Flow Experience and Various Types of Cognitive Load than Non-Game-Based Learning? Perspective from Multimedia and Media Richness. *Comput. Hum. Behav.* **2017**, *71*, 218–227. [CrossRef]
48. Bruner, J. *Actual Minds, Possible Worlds*; Harvard University Press: Cambridge, MA, USA, 1986.
49. Bruner, J. The Narrative Construction of Reality. *Crit. Inq.* **1991**, *18*, 1–21. [CrossRef]
50. Sailer, M.; Homner, L. The Gamification of Learning: A Meta-Analysis. *Educ. Psychol. Rev.* **2020**, *32*, 77–112. [CrossRef]
51. Soares, S.; Gonçalves, M.; Jerónimo, R.; Kolinsky, R. Narrating Science: Can It Benefit Science Learning, and How? A Theoretical Review. *J. Res. Sci. Teach.* **2023**, *60*, 2042–2075. [CrossRef]
52. Buckley, P.; Doyle, E. Gamification and Student Motivation. *Interact. Learn. Environ.* **2016**, *24*, 1162–1175. [CrossRef]
53. Merwade, V.; Eichinger, D.; Harriger, B.; Doherty, E.; Habben, R. The Sound of Science: An Engineering Design Challenge Teaches Students About Sound. *Sci. Child.* **2014**, *51*, 30–36. [CrossRef]
54. Dedetürk, A.; Saylan Kırmızıgül, A.; Kaya, H. The Effects of Stem Activities On 6th Grade Students' Conceptual Development of Sound. *JBSE* **2021**, *20*, 21–37. [CrossRef]
55. Abdoulqadir, C.; Loizides, F. Interaction, Artificial Intelligence, and Motivation in Children's Speech Learning and Rehabilitation Through Digital Games: A Systematic Literature Review. *Information* **2025**, *16*, 599. [CrossRef]
56. Reis, S.; Correia, N. The Perception of Sound and Its Influence in the Classroom. In *Human-Computer Interaction—INTERACT 2011*; Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6946, pp. 609–626. ISBN 978-3-642-23773-7.

57. Ravanis, K.; Kaliampos, G.; Pantidos, P. Preschool Children Science Mental Representations: The Sound in Space. *Educ. Sci.* **2021**, *11*, 242. [CrossRef]
58. Wouters, P.; Van Oostendorp, H. A Meta-Analytic Review of the Role of Instructional Support in Game-Based Learning. *Comput. Educ.* **2013**, *60*, 412–425. [CrossRef]
59. Charsky, D.; Ressler, W. "Games Are Made for Fun": Lessons on the Effects of Concept Maps in the Classroom Use of Computer Games. *Comput. Educ.* **2011**, *56*, 604–615. [CrossRef]
60. Barzilai, S.; Blau, I. Scaffolding Game-Based Learning: Impact on Learning Achievements, Perceived Learning, and Game Experiences. *Comput. Educ.* **2014**, *70*, 65–79. [CrossRef]
61. Lin, Y.-C.; Hou, H.-T. The Evaluation of a Scaffolding-Based Augmented Reality Educational Board Game with Competition-Oriented and Collaboration-Oriented Mechanisms: Differences Analysis of Learning Effectiveness, Motivation, Flow, and Anxiety. *Interact. Learn. Environ.* **2024**, *32*, 502–521. [CrossRef]
62. Faber, T.J.E.; Dankbaar, M.E.W.; Van Den Broek, W.W.; Bruinink, L.J.; Hogeveen, M.; Van Merriënboer, J.J.G. Effects of Adaptive Scaffolding on Performance, Cognitive Load and Engagement in Game-Based Learning: A Randomized Controlled Trial. *BMC Med. Educ.* **2024**, *24*, 943. [CrossRef]
63. Chang, C.-C.; Yang, S.-T. Interactive Effects of Scaffolding Digital Game-Based Learning and Cognitive Style on Adult Learners' Emotion, Cognitive Load and Learning Performance. *Int. J. Educ. Technol. High Educ.* **2023**, *20*, 16. [CrossRef]
64. Chen, C.-H.; Law, V.; Huang, K. Adaptive Scaffolding and Engagement in Digital Game-Based Learning. *Educ. Tech. Res. Dev.* **2023**, *71*, 1785–1798. [CrossRef]
65. Sun, L.; Ruokamo, H.; Siklander, P.; Li, B.; Devlin, K. Primary School Students' Perceptions of Scaffolding in Digital Game-Based Learning in Mathematics. *Learn. Cult. Soc. Interact.* **2021**, *28*, 100457. [CrossRef]
66. Ministry of Education of the People's Republic of China. *Compulsory Education Science Curriculum Standards*, 2022 ed.; Beijing Normal University Press: Beijing, China, 2022.
67. Fastl, H.; Zwicker, E. *Psychoacoustics: Facts and Models*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2006; Volume 22, ISBN 3-540-23159-5.
68. Oxenham, A.J. Pitch Perception. *J. Neurosci.* **2012**, *32*, 13335–13338. [CrossRef]
69. Everest, F.A. *Master Handbook of Acoustics*; The Mcgraw Hill Companies: Columbus, OH, USA, 2022; ISBN 1-260-47360-0.
70. States, N.L. *Next Generation Science Standards: For States, by States*; National Academies Press: Washington, DC, USA, 2013; ISBN 0-309-27230-0.
71. Kiili, K.; De Freitas, S.; Arnab, S.; Lainema, T. The Design Principles for Flow Experience in Educational Games. *Procedia Comput. Sci.* **2012**, *15*, 78–91. [CrossRef]
72. Li, Q.; Wang, P.; Liu, Z.; Zhang, H.; Song, Y.; Zhang, Y. Using Scaffolding Theory in Serious Games to Enhance Traditional Chinese Murals Culture Learning. Comput. Animat. *Virtual* **2023**, *35*, e2213. [CrossRef]
73. Cho, Y.H.; Huh, S.Y.; Jo, G.T. Influence of Individual Differences on Learning with Digital Textbooks. In Proceedings of the International Conference on Computers in Education, Metro Manila, Philippines, 26–30 November 2018.
74. Cho, Y.H.; Yim, S.Y.; Paik, S. Physical and Social Presence in 3D Virtual Role-Play for Pre-Service Teachers. *Internet High. Educ.* **2015**, *25*, 70–77. [CrossRef]
75. Bartholomeyczik, K.; Knierim, M.T.; Weinhardt, C.; Oettingen, G.; Ebner-Priemer, U. Capturing Flow Experiences in Every-day Life: A Comparison of Recall and Momentary Measurement. *J. Happiness Stud.* **2024**, *25*, 66. [CrossRef]
76. Lozano, L.M.; García-Cueto, E.; Muñiz, J. Effect of the Number of Response Categories on the Reliability and Validity of Rating Scales. *Methodology* **2008**, *4*, 73–79. [CrossRef]
77. Naegeli, A.N.; Hanlon, J.; Gries, K.S.; Safikhani, S.; Ryden, A.; Patel, M.; Crescioni, M.; Vernon, M. Literature Review to Characterize the Empirical Basis for Response Scale Selection in Pediatric Populations. *J. Patient Rep. Outcomes* **2018**, *2*, 39. [CrossRef] [PubMed]
78. Davis, F.D.; Bagozzi, R.P.; Warshaw, P.R. Extrinsic and Intrinsic Motivation to Use Computers in the Workplace[1]. *J. Appl. Soc. Pyschol.* **1992**, *22*, 1111–1132. [CrossRef]
79. Davis, F.D. Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. *MIS Q.* **1989**, *13*, 319–340. [CrossRef]
80. Altmeyer, K.; Barz, M.; Lauer, L.; Peschel, M.; Sonntag, D.; Brünken, R.; Malone, S. Digital Ink and Differentiated Subjective Ratings for Cognitive Load Measurement in Middle Childhood. *Brit. J. Edu. Psychol.* **2023**, *93*, 368–385. [CrossRef]
81. Klepsch, M.; Schmitz, F.; Seufert, T. Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Front. Psychol.* **2017**, *8*, 1997. [CrossRef]

82. Dijkstra, J.; Galbraith, R.; Hodges, B.D.; McAvoy, P.A.; McCrorie, P.; Southgate, L.J.; Van Der Vleuten, C.P.; Wass, V.; Schuwirth, L.W. Expert Validation of Fit-for-Purpose Guidelines for Designing Programmes of Assessment. *BMC Med. Educ.* **2012**, *12*, 20. [CrossRef]
83. Sweller, J. Cognitive Load During Problem Solving: Effects on Learning. *Cogn. Sci.* **1988**, *12*, 257–285. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.