

Sparking Early Literacy Learning through Block Construction and Puzzle Solving

Joey Bishop
jbishop47@gatech.edu

Abstract—The decline of fine motor skills (FMS) in young children is a growing concern, impacting crucial developmental areas like early literacy and academic achievement. To address this, we introduce **StoryBlocks**, a tangible user interface (TUI) that promotes early literacy and FMS development. StoryBlocks blends physical interlocking blocks with an iPad game where young learners solve storybook puzzles. Iterative qualitative studies revealed the prototype was highly engaging, facilitating self-expression and enthusiastic re-engagement with puzzles. Furthermore, the Intelligent Tutor System (ITS) proved pivotal in maintaining momentum by providing timely instructional hints. However, initial results are inconclusive on literacy development due to study limitations, and technical hurdles were noted. The facilitator's dual role as a technical proxy and pedagogical guide was indispensable.

1 INTRODUCTION

Research highlights the significant impact of fine motor skills (FMS) on child development, including academic achievement and early reading (Gaul & Issartel, 2016). A clear link exists between FMS and language learning, as both require planning and coordinating coherent sequences (Winter, Stoeger, & Suggate, 2024). Children with high FMS proficiency exhibit stronger expressive and oral narrative skills, whereas those with poor FMS may face academic struggles, increased anxiety, and low self-esteem (Gaul & Issartel, 2016).

Despite the reported significance of FMS in early development, opportunities for children to engage in FMS are declining at home and in school. Children's current environment can be quite passive with "increased opportunity for engagement in sedentary behaviors", including digital devices such as smartphones, tablets, televisions, and video game consoles (Gaul & Issartel, 2016). Winter, Stoeger, & Suggate (2024) argue that more fine motor activities

should be incorporated in kindergarten and school. Fine motor tasks make up between 30% and 60% of a child's school day, with up to 85% of those tasks being pencil and paper activities (Gaul & Issartel, 2016). If children are not provided with sufficient opportunities to engage in FMS, then it can be expected that "the downward trend in FMS might spill over to other developmental areas" like language learning (Winter, Stoeger, & Suggate, 2024). Without an environment that encourages fine motor skills, children might miss key opportunities in language development.

This problem is exacerbated by edutainment video games. Well-designed video games rooted in GBL offer many benefits, including the development of soft skills, critical thinking, self-confidence, self-discovery, and motivate learning (Anastasiadis, Lampropoulos, & Siakas, 2018; Kusuma, Padmadewi, & Saputra, 2024; Sari, Duygu, Şen, & Kirindi, 2020). Despite this, video games are largely a sedentary activity that limit physical play (Adipat et al., 2021), and miss the learning benefits of integrating cognitive activities with physical play (Wu et al., 2025).

Digital experiences that attempt to facilitate language development rely on existing FMS, which misses key developmental connections between FMS and language. In an environment where FMS proficiency is on the decline, these digital experiences do not facilitate any physical play, and might become less effective as learning tools. If these digital experiences are able to incorporate physical mediums that do encourage physical play and fine motor tasks, then children's FMS can be exercised while learning language. When provided at a critical age, this could prepare children for future academic success and wellbeing.

1.2 Terminology

Fine motor skills (FMS) are "small muscle movements" requiring eye-hand coordination (Luo et al., 2007). Examples include drawing, writing, and block building (Suggate, Stoeger, & Pefke, 2016).

Tangible user interfaces (TUIs) augment reality by linking digital data to physical objects (Rodić & Granić, 2021). These systems range from RFID-enabled blocks to music-emitting tables (Kleiman, Pope, & Blikstein, 2013; Markova, Wilson, & Stumpf, 2012). A key component of TUIs is manipulatives—physical objects

learners interact with to communicate digital information via technologies like Bluetooth or radio frequency (Rodić & Granić, 2021).

Intelligent Tutoring Systems (ITS) are artificial agents that serve as tutors in computer-based instruction. Complex ITS, which require designers to model learner mental states, must be "flexible, autonomous, and adaptive" to support more than basic "drill and practice" (Conati, 2009). Being digital, ITS permit personalization through immediate, delayed, or on-demand feedback (Conati, 2009; VanLehn, 2006). ITS are structured with an outer loop for selecting tasks and an inner loop for analyzing steps and providing feedback; the inner loop is only necessary when providing feedback to the user (VanLehn, 2006).

2 RELATED WORK

2.1 RoyoBlocks

Kleiman, Pope, & Blikstein (2013) created *RoyoBlocks* to support creative writing in pre-literate children. This system uses 60 RFID-enabled wooden word blocks alongside an Arduino-powered stuffed animal to help children construct sentences within a literature-rich environment. Educational materials also accompany the project to facilitate collaborative learning between children and their parents or teachers.

2.2 Scribblenauts

The *Scribblenauts* game series is in the action puzzle genre where players write in a magical notebook to summon whatever they can imagine. Although it is a Commercial Off the Shelf (COTS) game, the installment *Scribblenauts Unlimited* has been employed in vocational schools to aid in learning English as a foreign language (Kusuma, Padmadewi, & Saputra, 2024).

3 STORYBLOCKS: A MIXED MEDIA LANGUAGE LEARNING ACTIVITY

The developed artifact is a mixed-media TUI called StoryBlocks where young learners can explore words through physical blocks to solve storybook puzzles. The digital experience is supported by an iPad application developed in Unity. Physical manipulatives are composed of 3D-printed, grammar-enforcing interlocking blocks. A facilitator guide has been written for parents and teachers to guide learner behavior.

3.1 Digital Experience

The iPad application was written in C# through the Unity game engine, relying on Unity's Augmented Reality (AR) toolkit to support image recognition and Apple's native Text-to-Speech (TTS) library to generate audio narration.

3.1.1 Game Design

The application includes a sandbox game, two puzzle games, and a navigation menu to facilitate movement between these three games, seen in Figure 1. The sandbox game allows learners to explore the words and word interactions outside of the context of a puzzle, providing a space to gain familiarity with the TUI. The "picnic" puzzle game asks the player to help an anteater cross a river to reach a picnic basket, and requires only nouns and adjectives to solve. The "treasure" puzzle asks the player to help a zebra reach a jewel that is being guarded by an elephant, and requires nouns and verbs to solve.

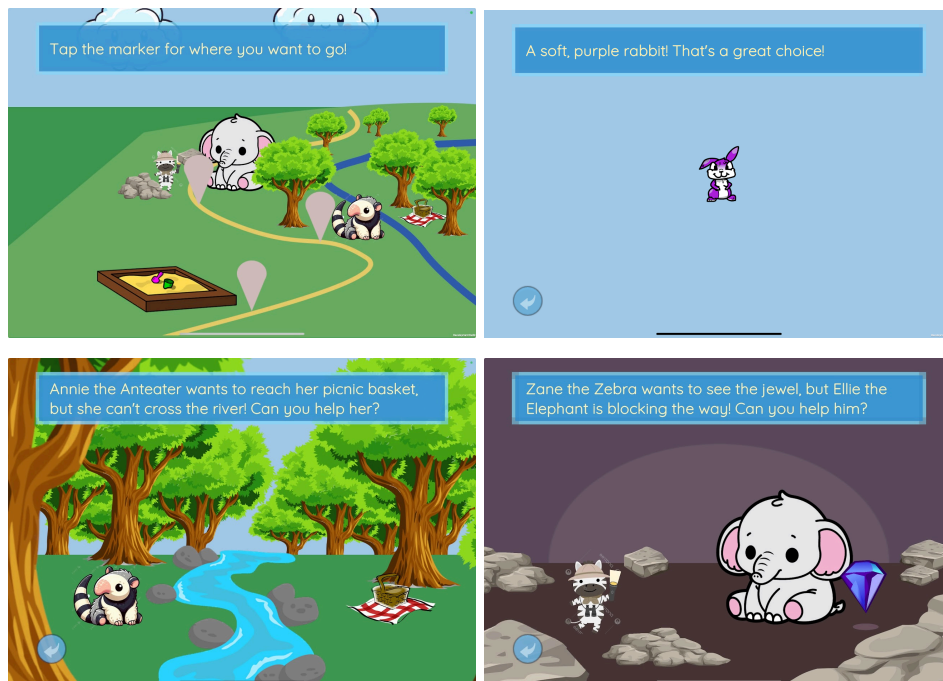


Figure 1—Top left: Navigation menu screen. Top right: Sandbox game screen with a scanned object. Bottom left: Picnic game setup screen. Bottom right: Treasure game setup screen.

The puzzle games capitalize on the affordances of game-based learning and variation to motivate interaction and replayability in the learner. Each puzzle game presents a storybook problem that the learner must solve through a

combination of words. Some word combinations guarantee wins, other word combinations will allow the learner to try again if they “miss” the answer, and other word combinations will fail the puzzle entirely. Failures are categorized separately from misses because they render the puzzle unsolvable; this distinction is similar to how a ball remains irrelevant to an elephant, whereas a singing dragon actively disturbs the elephant because of the noise.

Unique word combinations result in different scripted events. For example, the picnic puzzle game can be won with “boat”, which allows the anteater to cross the river, but it can also be won with “dragon”, which picks up the anteater and carries her across the river. This variation in scripted events encourages the learner to replay and explore different word combinations.

3.1.2 Intelligent Tutoring System Design

The application features a flexible ITS providing immediate TTS and UI feedback, a method suited for the target age group. Each game utilizes an outer loop and at least one inner loop following VanLehn’s semantic structure (2006).

In the sandbox game, the outer loop covers object creation, while the inner loop ensures a noun is present. The ITS scans word selections and provides immediate feedback, restarting the inner loop until a noun is detected, at which point it resumes the outer loop. The ITS then exits the inner loop to complete object creation. Puzzle tasks expand upon basic object creation by assessing the generated object against win, loss, or retry (miss) states. This creation cycle acts as an inner loop within the puzzle’s outer loop, repeating only if a miss occurs. Evaluation feedback is customized to the puzzle and word choice, providing increasingly detailed guidance as the player continues to miss.

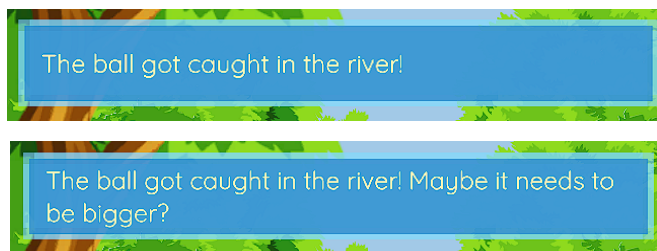


Figure 2—Top: ITS feedback related to the user’s object. Bottom: ITS feedback on retry with a hint.

During the initial attempt, feedback identifies the specific component of the object that caused the miss. By the second attempt, the system proposes a

modification to achieve a win, and for all subsequent tries, it offers an explicit win condition (Figure 2). While win and loss outcomes still receive immediate feedback, the level of detail provided does not increase further once the puzzle is definitively solved or failed.

3.2 Word Blocks



Figure 3—Each interlocking plastic block has a particular shape and color, and includes a unique word and word code.

The iPad app restricts touch input to menu navigation, utilizing nineteen 3D-printed interlocking plastic blocks (Figure 3) for primary interaction. Each block features a unique square code recognized via image recognition. These word-embossed blocks use interlocking mechanisms to encourage stacking and connection, fostering FMS development.

3.2.1 Supported Words

Nineteen words are supported by the TUI. The selection of words, seen in *Appendix 10.1: Supported words*, was based on young learners' familiarity with the terms, the diversity of concepts they represent, and the potential for imaginative expression within the iPad application. While the word codes are a necessity to input the learner's selected blocks, their design provides iconography that connects with the associated word, as seen in Figure 4.

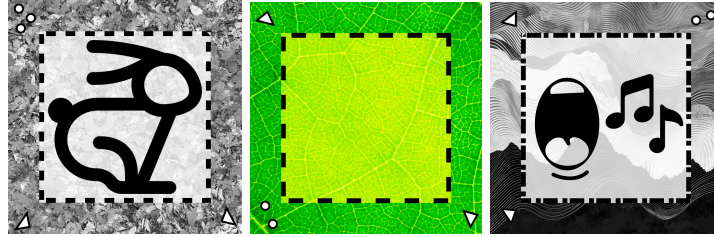


Figure 4—From left to right: The code for “rabbit”, the code for “green”, and the code for “sings”.

3.2.2 Block Design

The physical manipulatives of the TUI include three distinct block types tailored to specific parts of speech, utilizing mechanical connection points to enforce grammatical structure. Adjective blocks, measuring 3”x3”x1”, feature two triangle-shaped female ports. Both verb and noun blocks measure 6”x2”x1”; however, verbs utilize three hourglass-shaped male connectors, whereas noun blocks incorporate four triangle-shaped male connectors and three hourglass-shaped female ports (Figure 5).



Figure 5—From left to right: Adjective female ports, noun male connectors, noun male ports, and verb male connectors.

This framework dictates that adjectives interface with nouns solely via triangle connections, while verbs link to nouns exclusively through hourglass points. Consequently, the system permits a single noun to be modified by multiple adjectives but limited to one associated verb. Although the blocks do not provide

explicit linguistic definitions, the physical design effectively scaffolds fundamental grammar rules for early learners.

3.3 Facilitator Guide

A printable guide accompanies the TUI for parents and teachers to facilitate interactions with the learner. This includes activity suggestions that encourage play and practical reminders to support a healthy and engaging play session. Several activities suggested in the guide do not involve the iPad application, rather focusing on interacting solely with the word blocks. This is intentional to motivate movement and avoid sedentary behavior as a direct response to criticisms of edutainment games.

4 METHODOLOGY

The evaluation occurred in two stages: an initial trial to assess technical function and engagement, followed by a second iteration focused on literacy outcomes. Data was gathered through observational notes and photography.

4.1 Participants

The study involved a four-year-old male learner and a twenty-five-year-old female facilitator. The student was enrolled in daycare, and not pre-school, and the facilitator had no prior background in teaching or parenting.

4.2 Experiment Design

A four-stage structure was applied to both experiments: initial setup, unstructured block play, sandbox-style exploration, and structured puzzle tasks. During the setup phase, the facilitator received a digital version of the facilitator guide and instructions for the subsequent stages, with the requirement to use the facilitator guide throughout.

The remaining three stages involved active participation from both the facilitator and the learner. To assess natural block interaction and facilitator influence, blocks were initially distributed throughout the room for free play. The first iPad-based interaction occurred during sandbox exploration, allowing researchers to observe how learners connected physical blocks to the digital interface. The final stage, puzzle solving, evaluated the learner's ability to relate

specific blocks to digital gameplay events and the role of the facilitator in this process.

5 RESULTS

The following sections synthesize the data collected during the study. Comprehensive observational notes are provided in *Appendix 10.2: Observation notes*.

5.1 Early Literacy Learning

The learner primarily utilized the iPad application and the accompanying iconography to interpret word meanings, as their selection of blocks was based more on the visual word codes than the printed text.

In this context, the blocks functioned similarly to multimodal flashcards, pairing words with icons to convey ideas. This approach allowed for the correct identification of words like "dragon" but also led to significant misunderstandings. For instance, when the facilitator asked for the "purple" block, the learner could not locate it because the iconography featured grapes rather than a simple color representation, prompting a redesign of the word codes. Letter casing also presented challenges, as the learner frequently confused lowercase "l" with uppercase "I" while attempting to spell words.

The integration of the iPad application further blurred the distinction between individual words and the objects they represented. After creating a "blue ball" in the sandbox game, the learner identified the "blue" block itself as the "blue ball," failing to recognize the blocks as separate descriptors contributing to a whole. Similar behavior occurred during spelling tasks; although the learner could successfully spell "rabbit," they claimed it spelled "sticky rabbit sleeps." While this suggests a natural grasp of part-of-speech ordering, it highlights the need for clearer differentiation between words and objects in the interface.

5.2 Player Engagement

In block free play, the learner began to engage with the word blocks to build objects. During the first play session, the learner initially stacked blocks near each other. With some encouragement from the facilitator, the learner recognized that blocks can be connected, and identified the "triangles" on some blocks, while the hourglass was "a different shape". He then began to connect blocks of similar

color, without any encouragement, and stacked them in a row. In the subsequent play session, the learner laid out all blocks with the words face-up. With the assistance of the facilitator to recognize some words, the learner began to connect blocks together to build concepts, such as a “blue dragon” or a “big green ball”.

The learner also remained very engaged during the length of the iPad phases. He and the facilitator tried several combinations of adjectives, nouns, and verbs to explore how objects would appear. During the sandbox game, the learner often made objects that were blue, which was his favorite color. In this way, the activity became a form of self-expression. Furthermore, the learner enjoyed the puzzle games, asking excitedly to play again, and tried several combinations of objects to see what would solve the puzzles.

5.3 Game Design

Observations regarding block interaction indicated that the learner often attempted to place blocks in close proximity without utilizing the designated connection points; however, the facilitator's use of the phrase "connect the blocks" consistently prompted the correct behavior. Furthermore, facilitator frustrations were noted when words failed to scan, particularly as she would often tap the screen in response. Implementing an on-screen indicator for scanning progress or a camera view, as well as a notification that screen tapping is ineffective, would likely alleviate these issues.

Issues were also observed within the specific puzzle tasks. In the picnic puzzle, the learner's initial instinct was to suggest the anteater hop across rocks rather than using the blocks provided. The treasure puzzle proved confusing as the level design provided no clear indicators how to solve the puzzle. With the hints provided by the ITS, however, the facilitator and learner were able to find a solution. Additionally, there is currently no mechanical incentive for learners to utilize texture adjectives, as none of these words contribute to solving the puzzles.

6 LIMITATIONS

The study faced several limitations, primarily regarding the participant selection. Evaluation was restricted to a very small group consisting of a single four-year-old male relative as the learner and a single twenty-five-year-old female relative as the facilitator, who lacked professional experience in teaching

or parenting. Furthermore, the environmental and technical scope of the experiments was narrow. No empirical data was collected during the experiments, and follow-up literacy evaluations were absent from this study. Data collection occurred solely in home settings rather than educational institutions. Hardware was limited to a single 3rd generation iPad Air. Technical challenges further impacted the study, as the AR image tracking proved to be rigid, which prevented the system from scanning large block constructions. The physical blocks were difficult for the young learner to connect manually; while the facilitator helped with block construction, this mitigated the intent to motivate fine motor play.

7 CONCLUSION

The persistent decline of fine motor skills and early literacy, driven largely by increasingly sedentary lifestyles and passive digital engagement, highlights a critical need for developmental interventions that unify physical and cognitive play. StoryBlocks addresses this challenge as a TUI designed to simultaneously foster literacy and FMS development. By integrating grammar-enforcing interlocking physical blocks with an ITS-supported iPad game, StoryBlocks provides an interactive environment where young learners can solve storybook puzzles through tactile construction.

The impact of StoryBlocks on early literacy learning and fine motor skills is inconclusive with the limitations of the study. However, the prototype design was highly engaging for the learner, involving both physical and digital elements. The learner remained very engaged during the iPad phases, trying several combinations of adjectives, nouns, and verbs to explore how objects would appear in the digital world. The sandbox game became a form of self-expression, and the learner excitedly asked to play the puzzle games again, trying multiple object combinations. Physically, in block free play, the learner began to engage with the word blocks to build objects, eventually connecting blocks with the facilitator's assistance to build unique concepts. Furthermore, the ITS was pivotal in informing interaction and suggesting solutions, which helped maintain momentum during the activity. For instance, when the "treasure" puzzle proved confusing due to a lack of clear level design indicators, the hints provided by the ITS helped the facilitator and learner find a solution.

The facilitator occupied a central and indispensable role throughout the activity, functioning effectively as both a physical proxy for driving iPad interactions and a pedagogical guide for scaffolding the learner's conceptual exploration. This dual-layered involvement proved crucial for navigating various technical hurdles, such as rigid AR image tracking, while simultaneously ensuring the learner remained focused on the intended literacy goals rather than becoming distracted by the interface's limitations. By suggesting ideas and managing the digital environment, the facilitator bridged the gap between the physical word blocks and the virtual storybook puzzles, maintaining the session's momentum.

Word block design can be significantly improved to better support learning. Different letter casing created spelling confusion for the learner, and the presence of word codes often meant that blocks functioned like flashcards, resulting in less motivation to actually read the words. Furthermore, the distinction between individual words and the objects they represented often became blurred in the learner's mind. Finally, the blocks themselves were difficult for the learner to connect, which introduced friction points and required facilitator involvement, thereby limiting fine motor play. Despite these challenges, grammar-enforcing mechanics appeared successful as they were parroted by the learner during play.

8 FUTURE WORK

Future work will focus on improving the distinction between words and objects to avoid conflation. This may involve creating more distinct block textures and designs to better separate words, as well as implementing in-game animations that visually demonstrate multiple words combining to form a single object. The evaluation scope will be extended in future studies. This includes increasing the sample size and conducting research within classroom settings. Empirical evaluations will be integrated into future research to observe learner literacy over time. This longitudinal approach will provide data on the effectiveness of the tool for language acquisition and fine motor skill development. Finally, the project will explore the use of Near-Field Communication (NFC) or Radio-Frequency Identification (RFID) communication to replace existing AR image tracking, removing the need for word codes. Future studies will investigate how the lack of word codes—mitigating the "flashcard effect"—impacts learner engagement and overall understanding.

9 REFERENCES

1. Adipat, S., Laksana, K., Busayanon, K., Ausawasowan, A., & Adipat, B. (2021). Engaging students in the learning process with game-based learning: The fundamental concepts. *International Journal of Technology in Education*, 4(3), 542–552. <https://doi.org/10.46328/ijte.169>
2. Anastasiadis, T., Lampropoulos, G., & Siakas, K. (2018). Digital Game-based Learning and Serious Games in Education. *International Journal of Advances in Scientific Research and Engineering (Ijasre)*, 4(12). <https://doi.org/http://doi.org/10.31695/IJASRE.2018.33016>
3. Conati, C. (2009). *Intelligent Tutoring Systems: New Challenges and Directions*. <https://www.ijcai.org/Proceedings/09/Papers/012.pdf>
4. Gaul, D., & Issartel, J. (2016). Fine motor skill proficiency in typically developing children: On or off the maturation track? *Human Movement Science*, 46, 78–85. <https://doi.org/10.1016/j.humov.2015.12.011>
5. Kleiman, J., Pope, M., & Blikstein, P. (2013). Royoblocks. *Proceedings of the 12th International Conference on Interaction Design and Children*, 543–546. <https://doi.org/10.1145/2485760.2485861>
6. Kusuma, D. A., Padmadewi, N. N., & Saputra, I. N. (2024). Problem-based learning using scribblenauts unlimited for teaching English at Vocational High School. *Jurnal Pendidikan Bahasa Inggris Undiksha*, 12(3), 310–317. <https://doi.org/10.23887/jpbi.v12i3.91906>
7. Luo, Z., Jose, P. E., Huntsinger, C. S., & Pigott, T. D. (2007). Fine Motor Skills and mathematics achievement in East Asian American and European American kindergartners and first graders. *British Journal of Developmental Psychology*, 25(4), 595–614. <https://doi.org/10.1348/026151007x185329>
8. Markova, M. S., Wilson, S., & Stumpf, S. (2012). Tangible user interfaces for learning. *International Journal of Technology Enhanced Learning*, 4(3/4), 139. <https://doi.org/10.1504/ijtel.2012.051578>
9. Rodić, L. D., & Granić, A. (2021). Tangible interfaces in early years' education: A systematic review. *Personal and Ubiquitous Computing*, 26(1), 39–77. <https://doi.org/10.1007/s00779-021-01556-x>
10. Sari, U., Duygu, E., Şen, Ö. F., & Kirindi, T. (2020). The effects of STEM education on scientific process skills and stem awareness in simulation based Inquiry Learning Environment. *Journal of Turkish Science Education*. <https://doi.org/10.36681/tused.2020.34>

11. Suggate, S., Stoeger, H., & Pufke, E. (2016). *Relations between playing activities and Fine Motor Development*. *Early Child Development and Care*, 187(8), 1297–1310. <https://doi.org/10.1080/03004430.2016.1167047>
12. VanLehn, K. (2006). *The Behavior of Tutoring Systems*. <https://cs.uky.edu/~sgware/reading/papers/vanlehn2006behavior.pdf>
13. Winter, R. E., Stoeger, H., & Suggate, S. P. (2024). Fine motor skills and their link to receptive vocabulary, expressive vocabulary, and narrative language skills. *First Language*, 44(3), 244–263. <https://doi.org/10.1177/01427237241233084>
14. Wu, Y., Dang, X., Qiao, L., Li, A., Zhang, H., & He, J. (2025). Enhancing English language acquisition through Physical Education and motor imagery: Assessing enjoyment, English retention, and skill accuracy in adolescent learners. *Frontiers in Psychology*, 16. <https://doi.org/10.3389/fpsyg.2025.1672418>

10 APPENDICES

10.1 Supported words

Nineteen words are supported by the TUI (Table 1). The selection of words was based on young learners' familiarity with the terms, the diversity of concepts they represent, and the potential for imaginative expression within the iPad application.

Table 1—Words supported by StoryBlocks, including their part of speech and conceptual family

Part of Speech	Conceptual Family	Word
Noun	living	Dragon
Noun	living	Rabbit
Noun	nonliving	Ball
Noun	nonliving	Boat
Adjective	color	Blue
Adjective	color	Green
Adjective	color	Orange
Adjective	color	Purple
Adjective	color	Red
Adjective	color	Yellow
Adjective	size	Big
Adjective	size	Small
Adjective	texture	Messy
Adjective	texture	Sticky
Adjective	texture	Soft
Verb	boring	Cleans
Verb	boring	Sleeps
Verb	fun	Plays
Verb	fun	Sings

10.2 Observation notes

Evaluation 1 notes: 04/04

- Learner needed encouragement to realize blocks go together
- Tried to put blocks on sideways
- Tried to put blocks on backwards
- Had difficulty putting blocks together, even though connections were correct
- Learner put same color blocks together (all red)
- Noun and verb blocks can be put backwards
- Learner asked for more red
- Recognized connection points
 - Looked through all adj to look at connection points, realized they were all triangle
- Learner theorized that all yellow blocks would make a sun, didn't really take into account codes or words
 - With facilitator input, understood more that the blocks had more meaning
- Tried putting blocks together in space but not connecting them. Stacked them on top of each other or just next to each other
- Learner got confused by pictures. Lemon, grapes, etc. did not equate to adjectives "yellow", "grapes"
- Scanning is still very flimsy. Doesn't see all codes that are in front of it. Ex: "blue boat" either didn't see boat or didn't see blue
- Puzzle mode: learner insisted that you can hop on the rocks to cross the river
- Puzzle mode did not begin scanning
 - It actually did work, but the iPad camera was pointed away from the blocks. User error
- Phrase "Connect the blocks" really keyed in that they work
- Learner wanted to play the puzzle again after solving it

Evaluation 2 notes: 04/25

- When pulling out blocks, learner set them all out face-up
- Said hourglass was "a different shape"
- Spelled dragon unprompted

- Used picture to figure out what they say
- Facilitator tried tapping the app when trying to scan
 - Doesn't realize when it's reading the blocks
- Facilitator felt that "dragon sleeps" doesn't make sense to solve the zebra puzzle
- TTS voice was not very interesting
- Passing the iPad over all of the selected blocks helped grab the codes
- Learner had difficulty connecting hourglass blocks
- Facilitator would have liked being able to see the camera view
- Learner confused lowercase l with uppercase I on the blocks
- Learner used sandbox mode as a form of expression
- Treasure puzzle was unclear in setup and solutions