CollectiAR: Computer Vision-Based Word Hunt for Children with Dyslexia

Danlu Fei
dfeiaa@connect.ust.hk
IIP (Computational Media and Arts), The Hong Kong University of Science and Technology
Hong Kong SAR, China

Ze Gao
zgaoap@connect.ust.hk
IIP (Computational Media and Arts), The Hong Kong University of Science and Technology
Hong Kong SAR, China

Linping Yuan
lyuanaa@connect.ust.hk
Computer Science and Engineering, The Hong Kong University of Science and Technology
Hong Kong SAR, China

Zikai Alex Wen*
zikaiwen@ust.hk
Computational Media and Arts, The Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China

ABSTRACT

Children with dyslexia face extra challenges in reading and writing words. They need more learning exercises than children with typical development to acquire vocabulary, which is often repetitive and daunting. Research has shown that combining visuospatial information in practices helped children with dyslexia memorize words, especially the real-world physical context. Nevertheless, the existing word recognition and spelling training games for children with dyslexia were not able to leverage children's immediate vicinity. Therefore, we designed an augmented reality mobile game, *CollectiAR*, that uses computer vision to identify objects in the player's immediate vicinity and direct the player to learn words for these objects. Our formative study with two elementary school teachers and a first-grade pupil found that *CollectiAR* has the potential to be an integral part of teachers' instructional design and an engaging way for pupils to practice vocabulary exercises. Our teacher participants suggested that *CollectiAR* provide interfaces for teachers to participate in the game content design and computer vision model correction.

CCS CONCEPTS

• **Applied computing → Computer games**; • **Social and professional topics → People with disabilities**.

KEYWORDS

Educational Games; Special Education; Dyslexia; Computer Vision; Augmented Reality

ACM Reference Format:
Danlu Fei, Ze Gao, Linping Yuan, and Zikai Alex Wen. 2022. CollectiAR: Computer Vision-Based Word Hunt for Children with Dyslexia. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '22 EA), November 2–5, 2022, Bremen, Germany*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3505270.3558318

*Zikai Alex Wen is the corresponding author.

CHI PLAY '22 EA, November 2–5, 2022, Bremen, Germany

ACM ISBN 978-1-4503-9211-2/22/11.
https://doi.org/10.1145/3505270.3558318

1 INTRODUCTION

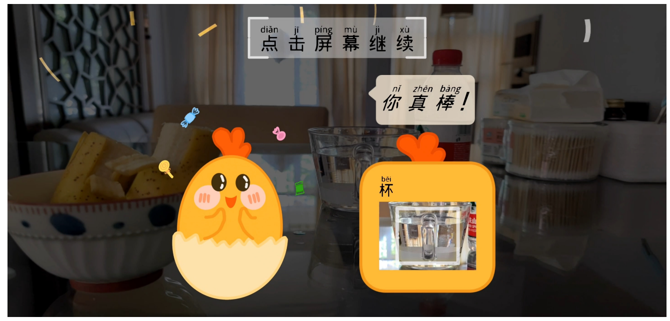
Approximate 10% of the world population experience dyslexia [23, 26]. *Dyslexia* is a specific learning disability that brings a challenge for children to establish a connection between the pronunciation and the letter symbol that represents it. As a result, children with dyslexia have trouble learning words, including short words commonly used in daily life. To overcome this challenge, children with dyslexia need to practice recognizing and spelling words frequently. Ideally, the practices should be productive and fun, but it is often repetitive and daunting.

With an aim of increasing engagement, prior work designed digital games for children with dyslexia (e.g., *Dybuster Orthograph* [13], *Dyseggxia* [19], and *DytectiveU* [20]). Existing games provide cartoon interfaces and scoring systems for quizzes about the meaning and spelling of words. *Dybuster Orthograph* helps children with dyslexia distinguish letter symbols by annotating every letter symbol with a unique visual pattern and sound. To improve the effectiveness of game-based exercises, *Dyseggxia* and *DytectiveU* customize words in the game according to the analysis of players' spelling errors.

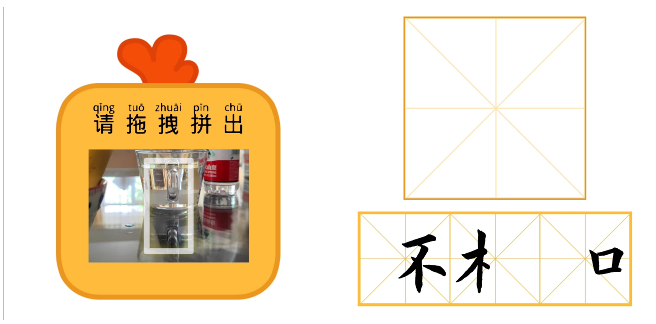
Although the existing games for children with dyslexia were reported to help children learn word recognition and spelling [13, 19, 20], the games did not provide sufficient context to tell children where the thing described by the word would appear. In other words, children cannot naturally infer the meaning of words by analyzing their surroundings in the games. We envision children learning vocabulary in their familiar physical context. For example, they can learn to spell the word "teacup" using the teacup from which they drink every day.

The idea of sending players off on a scavenger hunt to discover vocabulary words, also known as a word hunt, has been an exciting way to engage children with dyslexia in learning [3, 5]. Fujimoto et al. [7] also found that participants have a more accurate recall of information if it is visually associated with specific locations. In the past, word hunts depended on teachers or parents to design the list of vocabulary words based on children's living environment.

(a) At the hunting phase, the Chinese character of the target object is displayed at the bottom of the screen. The selected object is highlighted by a cyan box and the rest of the detectable objects are highlighted by white boxes. The player can listen to the pronunciation of the word of the selected object.



(b) At the success phase, the Red Bird tutor on the left applauds the player, "you are awesome!" Colorful ribbons fall in the background for celebration. The flashcard on the right displays the photo of the hunted object. The message at the top prompts the player to tap the screen to continue the game.



(c) At the traditional training phase, players need to use the correct radicals of the Chinese character to spell the word that describes the hunted object, which is displayed on the flashcard. Players can drag and drop a radical of the Chinese characters from the bottom slots to the big placeholder at the top.

(d) Players can review the words that they collected from their immediate vicinity in their inventory. Players can collect many objects in multiple rounds of word hunting. Their collection is displayed in the form of a flashcard that contains a picture taken by the player and the word that describes the object.

Figure 1: This diagram shows the flow of *CollectiAR*'s gameplay. After players enter the game, the words hunting game begins as shown in Fig. 1a. When players find the target object, the Red Bird tutor will appear, and congrats to players, as shown in Fig. 1b. Players will then proceed to the traditional spelling training session as shown in Fig. 1c. After completing the spelling training, the players will successfully collect the target object and find their collections in the inventory as shown in Fig. 1d.

The advances in augmented reality (AR) and computer vision provide new opportunities for the children to engage in vocabulary exercises and contextualize the vocabulary by exploring their surroundings. That being said, there is a critical challenge to solve as we design the word hunt game on AR mobile devices: How can AR be used together with computer vision to engage players in finding a target object in the player's immediate vicinity? Computer vision alone cannot solve this challenge.

With this in mind, we present *CollectiAR* as shown in Fig. 1. The version of *CollectiAR* in this paper is for children who struggle in learning Chinese words. It is simple to change the game to a different language. In *CollectiAR*, players start with a word hunt puzzle in their immediate vicinity where they can explore and collect highlighted objects (Fig. 1a). Each detected object is highlighted by a white box, and the selected object is highlighted by a cyan box. Each time players select an object, the game plays the Chinese pronunciation of that object as a hint. After players find the target object, the Red Bird tutor will compliment players and show the image of the correctly found object (Fig. 1b). After that, players will be navigated to a traditional spelling training (Fig. 1c). The game displays the image of the previously found object for players to tell which character to spell. Players can practice character spelling by dragging the Chinese radicals to form the target character that corresponds to the image. Besides, players can review the collections of word hunt items in the form of flashcards (Fig. 1d) in their inventory. When players complete the spelling training, they will automatically be directed to their inventory.

To study whether the design of *CollectiAR* meets the needs of pupils and their teachers. We conducted a formative study [9] with one first-grade pupil and two teachers for pupils with dyslexia in Chinese. The study showed that *CollectiAR* was fun and engaging and could potentially help pupils with dyslexia learn Chinese words. The teachers further suggested we add interfaces in *CollectiAR* that allow teachers to decide game scenes and to correct the computer vision model's mistakes.

To summarize, our work-in-progress shows that *CollectiAR* has the potential to engage children with dyslexia in learning Chinese words. In the future, we will continue polishing our game based on the formative study results. We will conduct a summative study to test the game's level of engagement and effectiveness in training word recognition and spelling. To dive deeper into understanding the children's gaming experiences, we will also conduct a long-term qualitative user experience evaluation.

2 RELATED WORK

2.1 Digital Games for Children with Dyslexia

The existing digital games developed for children with dyslexia primarily focus on word exercises. The exercises may involve recalling the meaning of words, spelling words, listening practices, etc. The majority of existing games gamify word exercises by designing cartoon interfaces and scoring systems. *DytectiveU* [20] takes a step further to adopt the whack-a-mole game mechanism, which requires players to pick the right answer that appears randomly from "holes". To improve the training effectiveness, *Dybuster Orthograph* [13] helps children distinguish letter symbols by annotating every letter symbol with a unique color, shape, structure and sound. *Dyseggxia* [19] and *DytectiveU* [20] reduce players' training time by customizing the word exercises according to players' spelling errors in the game. Nonetheless, the existing games for children with dyslexia cannot effectively leverage real-world physical context yet. Children still practice word exercises repetitively with little contextual information. Meanwhile, research [1] suggested that children with dyslexia have better learning performance in tasks that need to associate words with real-world physical contexts, including objects, locations, etc. Therefore, the primary design goal of our game is to let children with dyslexia enjoy learning vocabulary from their real-life environment.

Besides, Feng et al. [24] reported the unique challenge in learning Chinese words because of its logographic system. And we only found a few numbers of educational games research [6, 24] that targeted a logographic language like Chinese. *Multimedia Word* and *Drumming Strokes* are two multiplayer digital games that assist pupils in learning Chinese characters as a group [24]. Fan et al. [6] proposed using tangible shapes to teach Chinese characters. Unlike the previous work, our work attempts to leverage computer vision so that the objects in a player's immediate vicinity can be used as materials for a language learning game.

2.2 AR Games for Vocabulary Learning

The affordances of AR in learning have attracted a growing number of researchers to introduce AR into pedagogical scenarios, especially vocabulary learning. For example, Electric Agents [21] teaches vocabulary by interacting with a TV show through a specialized AR device. Hsu [10] designed an AR English spelling game that uses photos of objects to teach. Arbis Pictus [12] is an AR vocabulary learning application on HoloLens that allows a pupil to view the word of physical objects attached with pre-processed markers.

The limitation of these AR games is that they rely on marker-based tracking or position tracking, which may not always work well in outdoor environments or rooms that do not allow tagging. This approach also requires initial setup and maintenance effort, making it hard to deploy.

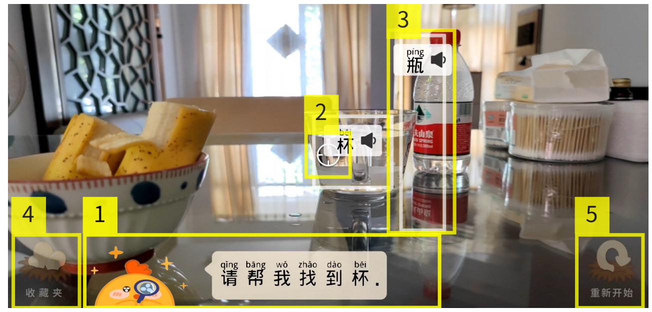


Figure 2: This screenshot shows that the Red Bird tutor is asking the player to "hunt" a teacup in the dining room. The player can use their camera to focus on objects bounded by white boxes. Highlight 1 shows the Red Bird tutor saying "please help me find the teacup." Highlight 2 is the telescopic sight. Players need to point the telescopic sight at the object they want to hunt. Highlight 3 is a selectable object detected by the computer vision model. Highlight 4 is the inventory mode button. Highlight 5 is the button that allows the player to skip and search for a new target.

3 GAMEPLAY DESIGN

CollectiAR is inspired by the design of a scavenger hunt [27] to discover vocabulary words, also known as word hunt. The goal of *CollectiAR* is to combine the advantages of both incentive structures and computer vision to transform the player's presence in the gaming environment. To implement *CollectiAR* that makes vocabulary learning fun, we need to: (1) detect and generate clues for word hunt items in the real-life environment, (2) provide reliable feedback in the game, (3) implement a collection system for long-term engagement, and (4) connect the word hunt game with traditional spelling training.

3.1 Generate Clues for Word Hunt Items

Before the game starts, *CollectiAR* requires the player to scan their surroundings to use computer vision to recognize several objects available for word hunt. Upon completing the environmental analysis, *CollectiAR* chooses a target from the list of recognized objects according to the player's historical correctness rate in the game. As shown in Fig. 2, *CollectiAR* chooses the teacup as the target.

After *CollectiAR* chooses a target word, the Red Bird tutor tells the player to find the target object by pronouncing its associated word (Highlight 1 in Fig. 2). For example, the Red Bird tutor tells the player to point at the object they think the teacup refers to. The player turns the mobile phone to aim at the object with the telescopic sight (Highlight 2 in Fig. 2). To tell the player which objects are interactable, *CollectiAR* highlights a teacup and a bottle on the dining table in front of the player by using white bounding boxes (Highlight 3 in Fig. 2).

3.2 Provide Reliable Feedback for Word Hunt

We can train computer vision models to recognize any objects, but their correctness rate is not 100% reliable. Nor can it filter out irrelevant detection information for a game design. Therefore, we must address the following challenges: (1) we need to limit the set of objects that the player needs to search through to find the target; (2) we need to verify whether the player finds the correct object and communicate the information to the player clearly.

When the player is surrounded by hundreds of objects, it becomes tedious for the player to go through all objects one by one. Instead, one of the design goals that *CollectiAR* must meet is to limit the set of objects that the player needs to search through to find the target object. The upper bound of the number of interactable objects is set to three in *CollectiAR*. First, the game filters out detected objects whose confidence level of computer vision detection is lower than 80%. Second, the game sorts the list of candidates according to the player's historical performance. Finally, the game picks the top three objects that the player makes the most errors or the game detects for the first time.

The second challenge is to provide the mechanism by which the player can indicate which object they think is the target. A strawman solution is to immediately reveal the result to the player when it detects an object being selected by the player. However, players may be uncertain of their decisions. Thus, they would hope for more time and hints to help them recall and confirm. Therefore, after the player selects an object, *CollectiAR* immediately changes the color of the bounding box into cyan and pronounces the selected object. After two seconds, *CollectiAR* confirms the result of the object that the player finds. Fig. 1a shows a particular scenario where the teacup is selected by the player. The cyan arc that grows around the open sight is the progression bar. *CollectiAR* will reveal the result when the cyan arc grows into a circle.

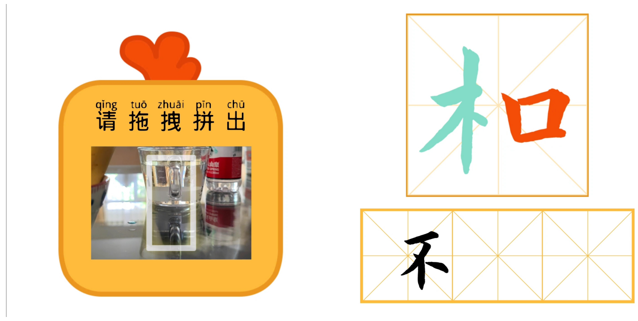
For the current game prototype, we used the computer vision model, *SSD MobileNet V1* [11], for objects segmentation and recognition. The top-1 accuracy rate was reported to be 70.6% [22], which suggests that on average, three out of 10 times of recognizing an object provide a wrong answer. The inconsistency of recognition results may cause players' frustration and negatively affect the gaming experience.

This technical constraint is noticeably annoying while the player chooses an object and waits for confirmation. To reduce the chances of missing frames or changing labels during the confirmation phase, we used the temporal smoothing technique to average recognition outcomes during the confirmation phase. We also designed a skip button to let the player skip and report a detection error that prevents them from proceeding (Highlight 5 in Fig. 2).

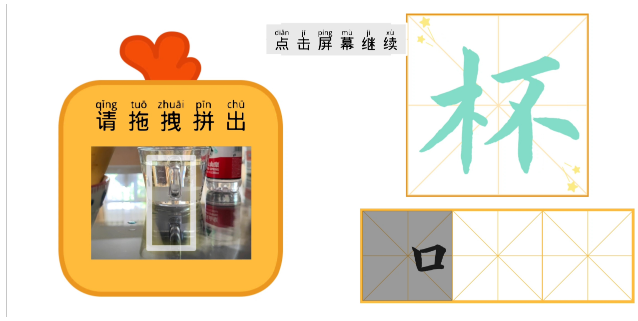
3.3 Collect Items for Long-Term Engagement

Because one of the game design goals is to keep players playing for an extended number of times, the additional challenge is how to encourage players to return and play *CollectiAR* again. In the Bartle taxonomy of player types [2], he also discusses the achiever who will devote themselves to achieving rewards that confer them little or no gameplay benefit simply for the prestige of having it. *Pokémon Go* [18] is one of the most successful AR games that attract this type of player.

Our long-term incentive design drives inspiration from the Pokémon Go collection system. Pokémon Go players do not necessarily need to use all of their collected Pokémon in gym combat. Still, the players are satisfied that they are filling up the Pokémon collection, especially when they catch rare Pokémon. Driven by a sense of



(a) The player made a mistake by dragging an incorrect Chinese radical, 口, to form the Chinese word, 杯. In this case, the game would mark the incorrect radical in red color and then shake the incorrect radical before sending it back to its default slot.



(b) The player successfully completed the spelling of the Chinese character, 杯, by placing radicals 木 and 不 in the right places. In this case, the game would show twinkling stars and then instruct the player to tap the screen to continue.

Figure 3: The figures show what players will experience in the spelling training game session when they make a mistake dragging and dropping an incorrect Chinese radical to form the target Chinese character (as shown in 3a) and when they correctly solve the puzzle (as shown in 3b).

satisfaction, Pokémon Go players constantly return to the game when they get to a new place to search for new or rare Pokémon.

Similarly, *CollectiAR* keeps track of the collection status and allows the player to check whether they complete any collection types: home living, electronics, kitchen appliances, etc. Fig. 1d shows the Collection mode that allows the player to review the items and new words found in their dining room.

3.4 Connect to Traditional Spelling Training

To solidify the knowledge of words in the heads of children with dyslexia, we also incorporate traditional spelling training into *CollectiAR*. We chose to implement spelling training that helps children learn better about the orthographic information of Chinese characters because children with dyslexia have difficulties processing multiple radicals that form one character [23]. Existing digital games rarely support such training for children with dyslexia. It is feasible to add other spelling training to *CollectiAR* in the future.

As shown in Fig. 1c, we designed three sections for this level, the "radical selection" section, the "spelling" selection, and the "question" section. The "question" section displays the image of the target object that the player correctly found. Players must drag the radicals from the "radical selection" section to the "spelling" section to piece together the Chinese character that describes the collected object. According to the characteristics of Chinese character structure, we classify Chinese characters into four categories: (1) characters that can be split left and right, (2) characters that can be split up and down, (3) characters that are nested inside and outside, and (4) characters that cannot be split. Players must select the correct radicals and drag them to the correct positions to complete the level.

As indicated in Fig. 3, When the player drags the wrong radical to the "spelling" section, the radical will vibrate and turn red to remind the player that they chose the wrong one (Fig. 3a). After that, the wrong radical will return to the "radical selection" section and turn gray to indicate that this wrong radical can no longer be selected and dragged. When the player drags the correct radical but does not place it in the correct position, the radical will return to the "radical selection" section for the player to try again. When the player drags the correct radical to the correct position, the radical will stay in the "spelling" section and turn green. When all the radicals are placed correctly, there will be a twinkling-star animation to celebrate the player's success (Fig. 3b).

4 STUDY RESULTS AND DISCUSSION

In this section, we report the results of a preliminary study on whether *CollectiAR* has the potential to be (1) an integral part of teachers' instructional design for children with dyslexia in Chinese and (2) an engaging way for children to practice vocabulary exercises. Then we discuss how we shall improve the game design based on the study results.

4.1 Method

We conducted a formative study [9] with two elementary school teachers and a first-grade pupil. We recruited the study participants through social media and emails.

We conducted the study at home with the child. We received consent from the child and their mother. The child's mother stayed with the child during the study. We let the child play the game for about 15 minutes. During the play, we noted all computer vision errors and how the child responded to the errors. After playing the game, we asked the child how they liked it, why they liked or disliked it, and how to improve its gameplay design. The study lasted about 30 minutes.

We conducted the study with two teachers for children with dyslexia in Chinese because they are experts at developing homework and recommending educational games for special-education pupils. We need to communicate with teachers to study how to integrate our game into their teaching. We interviewed our teachers remotely on Zoom. The study was remote because our teachers were working from home due to the COVID pandemic. We confirmed with the teachers that they only needed to watch how the pupils played the game to have adequate information to evaluate. We received consent from our teacher participants. We started by showcasing to the teacher how a pupil would play the game to practice the vocabulary exercises. The teacher could interrupt us, comment, or ask questions during this process. After the presentation, we asked them to comment on the usability of our game, how they would use the game in teaching, and how to improve the game design. The study lasted about an hour with each teacher.

4.2 Findings

The critical issue we found from the child's playtest was that the computer vision lost the recognition of the in-focus object in the middle, which made the child unable to aim at the object. The child was not surprised and shook the phone. They explained that they thought the game recognition was not good enough. It was interesting to find that the child already had some understanding of artificial intelligence, but we did not know whether this is a common occurrence for other children.

In addition, we found that the child was keen to find new objects. However, due to the limitation of the computer vision training model, some objects could not be detected (such as various types of fruits). Therefore, the child wanted this game to detect more items. They thought the game was pretty fun, mainly because it was not just about doing quizzes on a screen but walking around looking for answers.

The teachers thought the game was innovative and engaging because it was the first time they had seen an educational game that could teach children about real-life daily objects without needing teachers or parents. One teacher said they hoped to use this game in class to motivate the pupils once or twice. The two teachers suggested having pupils play the game at home as part of the homework. They were eager to know what words the pupils would collect in different scenes at home and how the speed of finding objects compares with the speed of spelling practice. They speculated that the game would help pupils connect the spelling of words with real-life use cases better than the spelling games they were using.

They also made two suggestions for game design improvement. First, our game may enable teachers to use their prepared videos or photos as game scenes, which helps teachers compare pupils' performance. Second, our game should provide a user-friendly interface for teachers or parents to check whether the computer vision model makes mistakes in judging the pupil's answers.

4.3 Discussion and Future Work

The teachers' suggestion shed lights on improving the computer vision model for games: to apply continual learning to the model through an expert-in-the-loop approach [4, 17]. Nonetheless, we must overcome the challenge of catastrophic forgetting [14], which means that artificial neural networks suddenly forget previously learned information when learning new information. Another challenge is to design the feedback loop without undermining player engagement. To expand the game learning content to teach phrases or sentences, we may replace the model with a photo caption generator [16, 25] or an AR-based optical character recognition [8, 15] and then modify the word hunt interactions accordingly.

5 CONCLUSION

Facilitating vocabulary learning through games has always been a challenging research topic. An ideal vocabulary learning app should be engaging, effective, and convenient. To achieve these goals, especially engagement, we designed a word scavenger hunt mobile game, *CollectiAR*, that uses computer vision and AR.

Our formative study with teachers and a pupil has shown that *CollectiAR* has the potential to be a new type of supplementary vocabulary learning game for training children with dyslexia. All of our study participants put forward the requirement for a higher and more stable computer vision model. In addition, the teachers hoped our game would provide interfaces for teachers to participate in the content design and computer vision model correction. We will continue improving the game based on the preliminary study results. We will also release *CollectiAR* to the app stores and conduct a summative study to test its level of engagement and effectiveness in helping children who struggle to acquire vocabulary knowledge. Besides, we will conduct a long-term qualitative user experience evaluation to understand the children's gaming experiences, especially after the novelty effect fades away.

ACKNOWLEDGMENTS

We gratefully acknowledge Erik Andersen and Bingxin Weng for their contributions to the early prototype design of *CollectiAR*. We would also like to thank the anonymous reviewers and Min Fan for their valuable inputs and comments.

REFERENCES

[1] Mary Alt, Tiffany Hogan, Samuel Green, Shelley Gray, Kathryn Cabbage, and Nelson Cowan. 2017. Word learning deficits in children with dyslexia. *Journal of Speech, Language, and Hearing Research* 60, 4 (April 2017), 1012–1028. https://doi.org/10.1044/2016_JSLHR-L-16-0036

[2] Richard Bartle. 1996. Hearts, clubs, diamonds, spades: Players who suit MUDs. (1996).

[3] Sarah Brown. 2011. *Dyslexia Games - Word Hunt 1*. CreateSpace.

[4] Samuel Budd, Emma C Robinson, and Bernhard Kainz. 2021. A survey on active learning and human-in-the-loop deep learning for medical image analysis. *Medical Image Analysis* 71 (2021), 102062.

[5] Liz Burton. 2016. Helping your student with dyslexia: Learn 5 strategies to rely on. https://www.dyslexic.com/helping-your-student-with-dyslexia-learn-5-strategies-to-rely-on/

[6] Min Fan, Jianyu Fan, Alissa N Antle, Sheng Jin, Dongxu Yin, and Philippe Pasquier. 2019. Character alive: A tangible reading and writing system for Chinese children at-risk for dyslexia. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.

[7] Yuichiro Fujimoto, Goshiro Yamamoto, Takafumi Taketomi, Jun Miyazaki, and Hirokazu Kato. 2012. Relationship between features of augmented reality and user memorization. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. https://doi.org/10.1109/ismar.2012.6402573

[8] Tushar Gupta, Leila Aflatoony, and Lynette Leonard. 2021. Augmenta11y: A reading assistant application for children with dyslexia. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–3.

[9] Marc Hassenzahl. 2008. User experience: (UX) towards an experiential perspective on product quality. In *Proceedings of the 20th Conference on l'Interaction Homme-Machine*. 11–15.

[10] Ting-Chia Hsu. 2017. Learning English with augmented reality: Do learning styles matter? *Computers & Education* 106 (2017), 137–149. https://doi.org/10.1016/j.compedu.2016.12.007

[11] Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, and Kevin Murphy. 2017. Speed/accuracy trade-offs for modern convolutional object detectors. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Honolulu, HI, 3296–3297. https://doi.org/10.1109/CVPR.2017.351

[12] Adam Ibrahim, Brandon Huynh, Jonathan Downey, Tobias Höllerer, Dorothy Chun, and John O'donovan. 2018. Arbis pictus: A study of vocabulary learning with augmented reality. *IEEE transactions on visualization and computer graphics* 24, 11 (2018), 2867–2874.

[13] Monika Kast, Gian-Marco Baschera, Markus Gross, Lutz Jäncke, and Martin Meyer. 2011. Computer-based learning of spelling skills in children with and without dyslexia. *Annals of Dyslexia* 61, 2 (2011), 177–200. https://doi.org/10.1007/s11881-011-0052-2

[14] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences* 114, 13 (2017), 3521–3526.

[15] Chen-Yu Lee and Simon Osindero. 2016. Recursive recurrent nets with attention modeling for ocr in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2231–2239.

[16] Pranay Mathur, Aman Gill, Aayush Yadav, Anurag Mishra, and Nand Kumar Bansode. 2017. Camera2Caption: A real-time image caption generator. In *2017 international conference on computational intelligence in data science (ICCIDS)*. IEEE, 1–6.

[17] Pascal Meier and Frank Teuteberg. 2021. Augmenting humans in the loop: Towards an augmented reality object labeling application for crowdsourcing Communities. *Innovation Through Information Systems* 2 (2021), 198.

[18] Niantic. 2016. Pokémon GO game. https://pokemongolive.com/en/

[19] Luz Rello, Clara Bayarri, and Azuki Gorriz. 2012. What is wrong with this word? Dyseggxia: A game for children with dyslexia. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*. ACM Press, Boulder, Colorado, USA, 219. https://doi.org/10.1145/2384916.2384962

[20] Luz Rello, Arturo Macias, Mariía Herrera, Camila de Ros, Enrique Romero, and Jeffrey P. Bigham. 2017. DytectiveU: A game to train the difficulties and the strengths of children with dyslexia. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Baltimore Maryland USA, 319–320. https://doi.org/10.1145/3132525.3134773

[21] Glenda Revelle, Emily Reardon, Kristin Cook, Lori Takeuchi, Rafael Ballagas, Koichi Mori, Hiroshi Horii, Hayes Raffle, Maria Sandberg, and Mirjana Spasojevic. 2014. Electric agents: Combining collaborative mobile augmented reality and web-based video to reinvent interactive television. *Computers in Entertainment* 12, 3 (2014), 1–21. https://doi.org/10.1145/2702109.2633413

[22] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. 2018. MobileNetV2: Inverted residuals and linear bottlenecks. (2018). https://doi.org/10.48550/ARXIV.1801.04381

[23] Liliane Sprenger-Charolles, Linda S. Siegel, Juan E. Jiménez, and Johannes C. Ziegler. 2011. Prevalence and reliability of phonological, surface, and mixed profiles in dyslexia: A review of studies conducted in languages varying in orthographic depth. *Scientific Studies of Reading* 15, 6 (2011), 498–521. https://doi.org/10.1080/10888438.2010.524463

[24] Feng Tian, Fei Lv, Jingtao Wang, Hongan Wang, Wencan Luo, Matthew Kam, Vidya Setlur, Guozhong Dai, and John Canny. 2010. Let's play chinese characters: Mobile learning approaches via culturally inspired group games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1603–1612.

[25] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. 2015. Show and tell: A neural image caption generator. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 3156–3164.

[26] Zikai Alex Wen, Erica Silverstein, Yuhang Zhao, Anjelika Lynne Amog, Katherine Garnett, and Shiri Azenkot. 2020. Teacher views of math e-learning tools for students with specific learning disabilities. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–13.

[27] Debra Wise and Sandra Forrest. 2003. *Great big book of children's games: Over 450 indoor and outdoor games for kids*. McGraw-Hill.