“Don’t judge a book by its cover”: Exploring book traits children favor

ASHLEE MILTON[1], LEVESSON BATISTA[2], GARRETT ALLEN[1], SOPHIE GAO[2], YIU-KAI NG[2], AND MARIA SOLEDAD PERA[1], [1] People and Information Research Team - Boise State University and [2] Computer Science Department - Brigham Young University

We present the preliminary exploration we conducted to identify traits that can influence children’s preferences in books. Findings offer insights for the design of recommender algorithms that would look beyond patterns inferred from traditional user-system interactions (e.g., ratings) for recommendation purposes, since when it comes to children such data is rarely, if at all, available.

CCS Concepts: • **Social and professional topics** → **Children**; • **Information systems** → *Recommender systems.*

Additional Key Words and Phrases: children, preferences, books, recommender systems, metadata

ACM Reference Format:
Ashlee Milton[1], Levesson Batista[2], Garrett Allen[1], Sophie Gao[2], Yiu-Kai Ng[2], and Maria Soledad Pera[1]. 2020. “Don’t judge a book by its cover”: Exploring book traits children favor. In *Late-Breaking Results ’20: ACM Recommender Systems Conference, September 22–26, 2020, Online.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

There are a plethora of well-known recommender systems (**RS**) that offer *traditional users*, i.e., adults, appealing items pertaining to well-known *domains* like movies, restaurants, books, and songs. State-of-the-art, popular recommender algorithms are often based on collaborative filtering and deep neural network architectures [3, 23] that depend upon the existence of user-system interactions in the form of likes, reviews, and ratings. Availability of this type of data is essential in making personalized suggestions, as well as in comparing items and users to identify those that share similar traits and preferences, respectively, which in turn can inform recommendations [2]. Subsequently, ratings, reviews, and users’ profiles are at the core of more complex RS design that can better serve traditional users. This prompted us to question what type of data can be relevant and useful for RS to consider when the main stakeholders for the recommendations are not adults?

There are many niche user groups, each with its own singularities. We focus our study on *children* and their journey to find appealing *books*, due to the importance of developing literacy skills at an early age [11, 17]. RS are an ideal medium to help match reading resources with the right readers. Unfortunately, research efforts allocated for the design of RS for children are limited [4]. Like adult users, children have diverse tastes [1]. Alas, adult’s user-system interactions are available in large quantities, allowing for identification of their tastes, whereas as a protected population explicit children user-system interactions are seldom (if at all) available, thus restricting pattern analysis. More so, children’s ratings are less frequent than adults’ and their ratings tend to skew towards 4’s and 5’s on the Likert scale [6]. This

Manuscript submitted to ACM

causes us to wonder which traits impact children's decision-making, causing them to favor certain books, and how can these traits be used to inform RS design when historical data is not available?

As a way to identify book traits that can serve as a step towards (i) tailoring RS for children, even if personalization is not always possible [5], and (ii) addressing cold-start based on children's trends—as opposed to those extrapolated from traditional RS users—we conduct a preliminary empirical exploration. We use several lenses to examine books favored by children of different ages and distinguish traits that most prominently arise. The main contribution of our preliminary exploration, as discussed in the rest of this manuscript, is the delineation of children's preferences across age groups, which have implications that can inform the design of book RS for children.

2 DATA DESCRIPTION

For exploration purposes, we turn to two different children-related data sources. **OnlineBookApp** is comprised of books read by children across the USA and bookmarked on BiblioNasium.com, a site dedicated to encouraging reading development. While relatively small, **PublicLibrary** includes book ratings/reviews explicitly provided by children at two local public libraries in the US, one in Indiana and one in California. Statistics related to these data sources are presented in Table 1. To enable analysis, we enriched **PublicLibrary** and **OnlineBookApp** with corresponding book metadata (description, cover, number of pages, title, author, and ISBN), which we extracted from the GoodReads dataset [20, 21], when available, as well as book-related APIs[1].

Table 1. Descriptive statistics of the data sources considered in our analysis.

Data Source	# of Unique users - # of (user, book) pairs					
	≤ 5	6-8	9-11	12-13	≥ 14	Overall
PublicLibrary	10–10	51–74	115–154	34–48	9–57	219–343
OnlineBookApp	33–607	492–7234	913–9642	23–224	N.A.	1461 – 17707

3 DISCUSSION: RESULT ANALYSIS AND POSSIBLE IMPLICATIONS

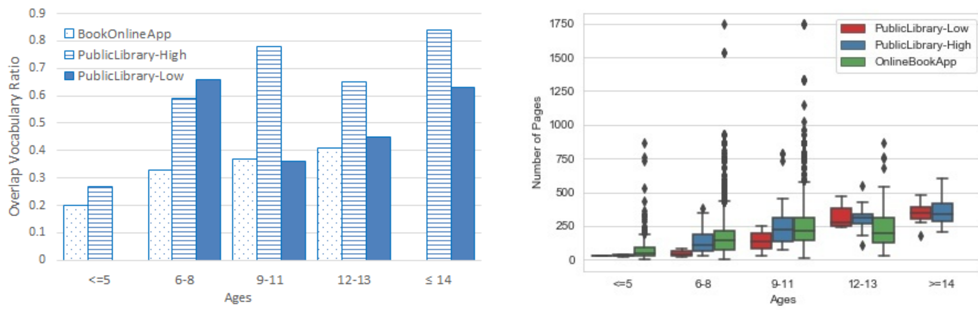
We discuss below, patterns and favored book traits that emerged from examining resources discussed in Section 2 from multiple perspectives. Along the way, we present implications of our findings, in terms of how they can inform and influence the design of book recommender algorithms tailored to the needs and expectations of children. Interests and preferences are known to evolve as children mature. Thus, whenever pertinent, we discuss our findings across different age groups: from emergent readers to high-school students. Further, to take advantage of the explicit preference samples available in the **PublicLibrary**, we compare and contrast results between **PublicLibrary-High** and **PublicLibrary-Low**, which capture **high-rated** (ratings above 3) and **low-rated** books, respectively. Unless stated otherwise, **statistical significance** of reported results is based on t-test, $p \leq 0.01$.

3.1 Book Title

Titles are a determining factor in children's preference for a book [15]. With that in mind, we look at vocabulary often used in titles, along with the length of titles with(out) stop words for books in **PublicLibrary** and **OnlineBookApp**. We first identify individual terms in titles to investigate if they also appear in the Age of Acquisition (AoA) data[2] for a given age group [9]. This enables us to compute the proportion of known words found among titles for each age group.

[1] https://openlibrary.org/dev/docs/api/books; https://libraryofcongress.github.io/data-exploration/; https://pypi.org/project/isbnlib/

[2] Each of the 51,715 terms in this dataset is assigned a score that reflects the age at which an individual should be able to comprehend that term.



(a) Title vocabulary

(b) Book length

Fig. 1. Book title and length analysis for books in **OnlineBookApp** and **PublicLibrary** among different age groups.

The growth in ratio between the terms in titles that are also present in the AoA data across ages serves as evidence that children tend to favor books with titles reflecting vocabulary that they can comprehend. For both data sources examined, the ratio for children age 5 or lower is approximately in the 20^{th} percentile (Figure 1(a)). We attribute this to parents reading to and with children enabling discussion of new terminology, as well as younger children having a more limited vocabulary. Another interesting anecdote that emerges from this analysis refers to the noticeably smaller rations when comparing **PublicLibrary-High** and **PublicLibrary-Low**. This leads us to believe that children tend to pass by books with titles that they cannot comprehend.

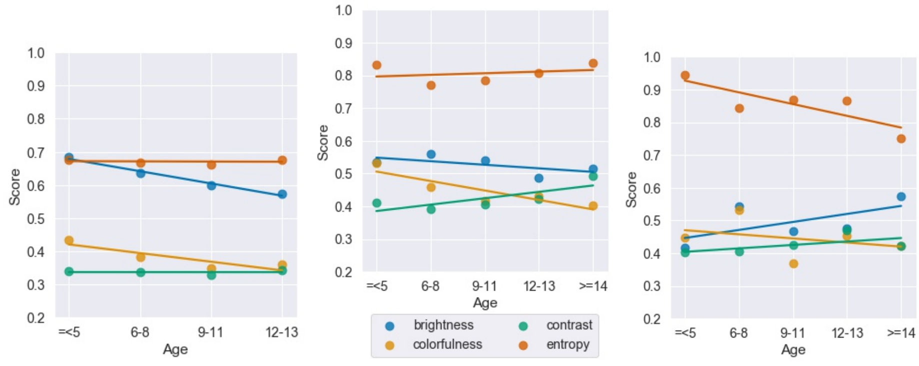
From title length analysis, with(out) stop words, we find several statistically significant preference changes based on age. On **PublicLibrary-High**, the title length gets shorter between the age groups 9-11 and 12-13 ($p \leq 0.02$). On **OnlineBookApp**, the title length increases between ages ≤ 5 to 6-8 ($p \leq 0.02$) and 6-8 to 9-11. Surprisingly, although not statistically significant, title length for low-rated books in age group 12-13 is longer than for high-rated ones. Results reveal that vocabulary and number of terms in book titles can sway children's book preferences, which RS can leverage to prioritize suggestions that best suit children at different stages.

3.2 Book Length

The length of a children's book can have a significant impact on children's comprehension of it and affect their preference of the book [7]. To explore whether children exhibit a proclivity towards specific book lengths at different ages, we examine the number of pages for books in **PublicLibrary** and **OnlineBookApp** (summarized in Figure 1(b)). For **PublicLibrary** there is a statistically significant uptrend in book length as children age. Unlike age group 12-13 in **PublicLibrary**, who view longer books unfavorably, the remaining age groups often assign lower ratings to shorter books. Furthermore, longer books are preferred by children of age 9-11 (visibly from page length spanning from 250 to 1750 pages) for both **PublicLibrary-High** and **OnlineBookApp**.

3.3 Book Covers

A well-known idiom advises to not judge a book by its cover. Yet, aesthetic factors have been shown to play a role in capturing traditional users' preferences on RS in the movie domain [16]. As reported in [10, 15], aesthetic influences which books children select for themselves to read, but the specific visual features that draw children to books has yet to be explored. This motivates us to examine book covers. Inspired by [19], who analyzed recipe images for



(a) **OnlineBookApp** (b) **PublicLibrary-High** (c) **PublicLibrary-Low**

Fig. 2. Average brightness, colorfulness, contrast, and entropy by age groups.

recommendation purposes, we consider a number of lenses in our own exploration of book covers: *dominant color*, *brightness*, *colorfulness*, *contrast* and *entropy* (defined as the randomness in the intensities of an image). For finding the dominant color of an image, we use kMeans to cluster each pixel in an image based on its RGB (red, green, blue) numerical representation[3]. From the generated clusters, we find the cluster with the most pixels and compute its centroid. The resulting RGB value is then used to find the family color name utilizing the CSS color module level 4. For the remaining scores (each normalized to [0,1]), we use the Python Image Library and OpenCV. Further details on how to compute these scores can be found in [19].

As shown in Figure 2(a), both colorfulness and brightness trend down as users' age increases. This decrease in colorfulness and brightness is statistically significant for ages ≤ 5 compared to 6-8, 9-11, and 12-13, as well as ages 6-8 regarding ages 9-11 and 12-13 for brightness, but only ages 9-11 for colorfulness. Contrast begins to trend upwards with ages 6-8 being statistically significant over ages 9-11. Shifting focus to the book covers in **PublicLibrary** the main change is related to contrast. From Figure 2(b), we see a significant increase in contrast for ages ≥ 14 over 6-8, 9-11, and 12-13 ($p \leq 0.03$). We began to see this trend in covers in **OnlineBookApp**, i.e., significant increase in contrast between ages 6-8 and 9-11, but there is no information for ages ≥ 14 in **OnlineBookApp**, which is where the significance is in **PublicLibrary**. Although no other significance was found in the **PublicLibrary** covers, there is a similar change seen in **OnlineBookApp** for brightness and colorfulness of high-rated books (Figure 2(b)). We also see a downward turn in entropy over age from Figure 2(c), but there is no statistical significance related to the books in **PublicLibrary**-LowRating, likely due to the small sample size. No patterns emerged from the dominant colors of the book covers. Findings demonstrate that aesthetic aspects do play a role in the books that children choose at different ages. This presents the need for recommenders designed to support children reading to consider visual features in the recommendation process regarding users age, specifically brightness, colorfulness, and contrast.

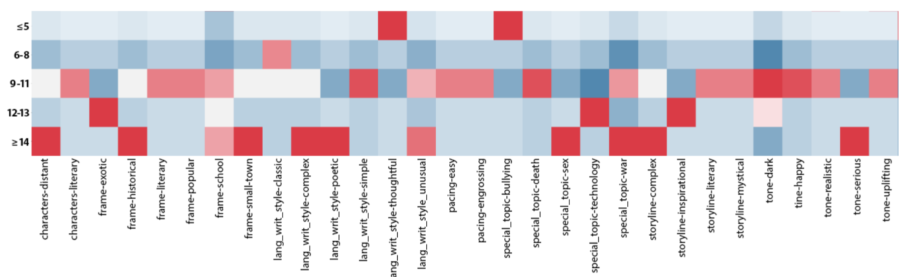
3.4 Literary Elements

Books are often described via a set of structural components, literary elements, which serve as a template for authors to construct their stories and include *characters*, *frame*, *language and writing style*, *pacing*, *storyline*, *tone*, and *special topics*. We investigate these elements in children's books to see if any preferences emerge across different age groups. To the

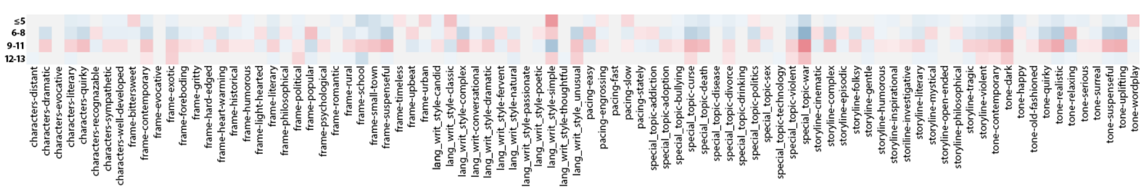
[3]For the kMeans we set k=8, to account for the 6 colors of the rainbow plus black and white.

best of our knowledge, book metadata pertaining to literary elements is not available via well-known book-related APIs or the GoodReads dataset. For this reason, we rely on the collection of literary elements and their descriptive terms found in [14], e.g., *frame-historical* or *pacing-slow*. We identify the presence of any of the 118 element-term pairs (in [14]) in each book description in **OnlineBookApp** and **PublicLibrary**. We then use this information to calculate Pearson correlations of the different element-term pairs across different age groups.

As illustrated in Figure 3, readers ages 9-11 prefer a larger variety of element-term pairs. In contrast, the remaining age groups have preferences for a smaller variety of element-term pairs. For example, readers aged 5 and under have preference for *language_and_writing_style-simple*, *language_and_writing_style-classic* and *pacing-slow*, whereas readers older than 14 have more affinity with complex elements, such as *special_topic-war* and *language_and_writing_style-passionate*. Additionally, readers ages 9-11 have strong preference for books about war (*special_topic-war* yields $r = 0.05$ in **OnlineBookApp** and $r = 0.06$ in **PublicLibrary-High** for this age group, as illustrated in Figures 3(b) and 3(a), resp.). Overall, correlation trends across element-term pairs and age groups are similar in both data sources considered in our analysis. We did see a number of very dissimilar correlations for *special_topic-war*, *special_topic-technology*, and *frame-timeless* when looking at ages 12-13 on the different data sources (i.e., positive correlation in **PublicLibrary-High** and negative in **OnlineBookApp** for the first two, and negative correlation in **PublicLibrary-High** and positive in **OnlineBookApp** for the last). We believe these differences could be a result of the much smaller number of observations in **PublicLibrary-High**.



(a) **PublicLibrary-High**



(b) **OnlineBookApp**

Fig. 3. Preference on literary elements across age groups, based on Pearson correlation coefficients.

Results analysis highlight that as children get older, their book preferences change to include more complex element-term pairs. Emergent readers start out preferring simple writing with a slow pace, moving towards books with more serious tones, e.g., battles, journeys, or quests at ages 9-11 and by age 14 they further progress to passionate, poetic writing styles. While this trend was expected, it showcases that recommender algorithms tailored to the population under study should not adopt a one-size-fits-all approach to children. Instead, they should explicitly consider the inclinations that are more common for different age groups.

3.5 Emotions

As defined in [18] RS users receive "various stimuli (e.g., visual, auditory, etc.) that induce emotive states", which can influence their decision-making process. Indeed, books can make a reader feel a plethora of emotions. Consider Figure 4, which portrays the intensity of emotions found on three books. For "How do dinosaurs say happy birthday" (readers age 5 and below), we see a high proportion of positive emotions like joy, trust, surprise, and anticipation, whereas a book for more mature children, like "The summer house", exhibits more negative emotions like fear, anger, and disgust. This prompted us to examine emotions present in books children favor at different ages. We analyzed the emotions exhibited in the description of each book in **OnlineBookApp** and **PublicLibrary** and created a vector that captures the intensities for eight emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. The intensity of the emotions of non-stop, lemmatized terms from book descriptions is determined using the Emotion Intensity Lexicon (NRC-EIL) [12]. For each term, NRC-EIL assigns a [0,1] score for each emotion; terms not found in the lexicon were treated as "objective". This results in a book emotion vector representation that is derived as the element-wise average of the emotion vector representation of its corresponding terms.

We illustrate in Figure 5 the emotion distribution inferred from books favored by children of different age groups. Since the differences regarding **PublicLibrary-Low** were not statistically significant (as anticipated due to the small sample size of low-rated books), we focus on the emotions extracted from **PublicLibrary-High**. From **PublicLibrary-High**, we notice a pattern where *joy* decreases and *sadness, fear, anger*, and *disgust* increase as children get older. This pattern becomes more evident when analyzing the **OnlineBookApp** set which presents significant results in the emotion distributions of *joy, sadness, fear, anger*, and *disgust* between ages, especially between the age groups under 5, 6-8 and 9-11, as seen in Figure 5(b). For **OnlineBookApp** there is a significant decline in the intensity of *objective* terms for age groups ≤ 5 vs 6-8, and 6-8 vs 9-11 ($p < 0.001$). A similar trend (although not significant) is observed for **PublicLibrary**. We attribute this to the more prominent and varied display of emotion on books intended for older children.

We notice an emotional shift in preference between adjacent age groups (Figure 5). Upon closer inspections of book descriptions, we find that young children are inclined towards joy, favoring books including terms like *celebration, excitement, gratitude, cheerful*, and *smile*, which are terms for which NRC-EIL assigns high intensity scores for the emotion joy. This indicates that young children prefer books with happy subjects as contrasted with books that tailor towards complex themes. As children get older, the intensity of the emotion *joy* decreases, while *sadness, fear, anger*, and *disgust* increase, that are the result of children rating higher books including terms such as *bully, compassion, crime, battle*, and *mystery*, for which NRC-EIL assigns high intensity scores to the aforementioned emotions. This would imply

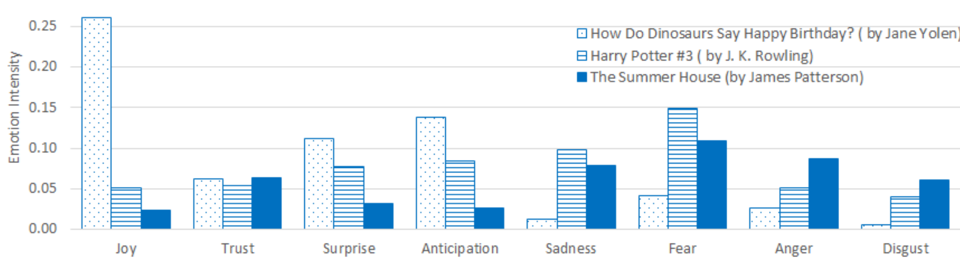
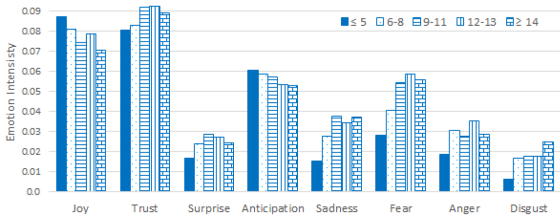
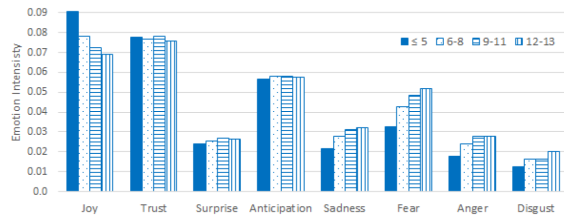


Fig. 4. Emotion intensity distribution for three different children's books, generated using NRC-EIL.



(a) PublicLibrary-High



(b) OnlineBookApp

Fig. 5. Emotion distributions among different age groups.

that older children favor books with more elaborated themes such as *adventures, good versus evil*, or *romances*, which aligns with our findings reported in Section 3.4.

4 CONCLUSIONS AND NEXT STEPS

Book RS can serve to promote children's love of reading, as well as overall learning and development. Yet, for RS to be effective, in terms of matching the right reading material with the right child, they need to "work in conjunction with diversification mechanisms to challenge and widen children's thinking and diversification should not be conflated with randomization" [8]. In light of the fact that (i) personalization of RS for children is non-trivial–restricted by lack of recorded interactions among young users and RS, and (ii) interests do change over time as children mature, in this paper we have discussed the results of the initial empirical exploration we conducted in order to identify patterns that RS can leverage to better support children's selection of reading materials.

Preliminary results reveal that younger children prefer bright colorful covers, emotions with a positive connotation (like surprise), and shorter books with simple writing styles. As complex topics cannot really be explored in only a handful of pages, it is natural for books targeting this age group to be about a single topic, explaining fewer literary elements that yield noticeable correlations for this age group. As they get older, children prefer less colorful/bright covers, darker, gloomier themes in books, more varied emotions, and more complex topics. The increase over time of children's preference for contrast shows that although young children are attracted to colorful images, as they get older, they gravitate towards images that "pop", rather than just being colorful. We also saw that the number of simultaneous topics mentioned in a single book tended to increase with age, demonstrated by more literary elements being correlated with older age groups. Both title vocabulary and length can be used by RS as a proxy for children's interest. Further, titles including known vocabulary can draw them to a book; at the same time, RS prioritizing books with titles including new vocabulary can foster learning though exposure. While a number of our findings were anticipated, this exploratory analysis serves as groundwork that can inform the design of RS that give children what they want, in their quest for appealing reading materials, rather then what other stakeholders (e.g., parents and educators) believe they should get.

For future work, we will explore how cover-related feature preferences change when children are affected by conditions like visual impairment or dyslexia[4]. We also plan to examine other perspectives inspired by outcomes reported in child-psychology and education literature [13, 22]—book genres, topics, media types, purpose (for learning vs. for leisure), and user gender—as they can reveal other traits recommender algorithms should consider. We strive to incorporate these findings on the design of a RS for emergent readers, one that explicitly relies on traits that are prominent among this age group as a step towards personalization when historical data is not available.

[4]https://www.w3.org/TR/low-vision-needs/#brightness-and-color

ACKNOWLEDGMENTS

Research partially supported by NSF Award #1763649.

REFERENCES

[1] Richard E Ahl and Yarrow Dunham. 2020. Have your cake, and your asparagus, too: Young children expect variety-seeking behavior from agents with diverse desires. *Cognitive Development* 54 (2020), 100882.

[2] Enrique Amigó, Fernando Giner, Julio Gonzalo, and Felisa Verdejo. 2020. On the foundations of similarity in information access. *Inf. Retr. J.* 23, 3 (2020), 216–254.

[3] Chong Chen, Min Zhang, Chenyang Wang, Weizhi Ma, Minming Li, Yiqun Liu, and Shaoping Ma. 2019. An efficient adaptive transfer neural network for social-aware recommendation. In *Proc. of International ACM SIGIR Conference on Research and Development in Information Retrieval*. 225–234.

[4] Jerry Alan Fails, Maria Soledad Pera, Franca Garzotto, and Mirko Gelsomini. 2017. KidRec: Children & recommender systems: Workshop co-located with ACM conference on recommender systems (recsys 2017). In *Proc. of the Eleventh ACM Conference on Recommender Systems*. 376–377.

[5] Cristina Gena, Pierluigi Grillo, Antonio Lieto, Claudio Mattutino, and Fabiana Vernero. 2019. When Personalization Is Not an Option: An In-The-Wild Study on Persuasive News Recommendation. *Information* 10, 10 (2019), 300.

[6] Michael Green, Oghenemaro Anuyah, Devan Karsann, and MS Pera. 2019. Evaluating prediction-based recommenders for kids. In *3rd KidRec Workshop co-located with ACM IDC*.

[7] Maryam Jaliehvand. 2012. The Effects of Text Length and Picture on Reading Comprehension of Iranian EFL Students. *Asian Social Science* 8, 3 (2012), 329.

[8] Natalia Kucirkova. 2019. The Learning Value of Personalization in Children's Reading Recommendation Systems: What Can We Learn From Constructionism? *International Journal of Mobile and Blended Learning (IJMBL)* 11, 4 (2019), 80–95.

[9] Victor Kuperman, Hans Stadthagen-Gonzalez, and Marc Brysbaert. 2012. Age-of-acquisition ratings for 30,000 English words. *Behavior research methods* 44, 4 (2012), 978–990.

[10] Monica Landoni and Elisa Rubegni. 2017. Aesthetic relevance when selecting multimedia stories. In *International Workshop on Children & Recommender Systems. Available at: https://drive. google. com/open.*

[11] Margaret Kristin Merga and Saiyidi Mat Roni. 2018. Children's perceptions of the importance and value of reading. *Australian Journal of Education* 62, 2 (2018), 135–153.

[12] Saif Mohammad. 2018. Word Affect Intensities. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

[13] Kathryn J O'Toole and Kathleen N Kannass. 2018. Emergent literacy in print and electronic contexts: The influence of book type, narration source, and attention. *Journal of Experimental Child Psychology* 173 (2018), 100–115.

[14] Maria Soledad Pera and Yiu-Kai Ng. 2014. Automating readers' advisory to make book recommendations for k-12 readers. In *Proc. of the 8th ACM Conference on Recommender systems*. 9–16.

[15] Kara Reuter. 2007. Assessing aesthetic relevance: Children's book selection in a digital library. *Journal of the American Society for Information Science and Technology* 58, 12 (2007), 1745–1763.

[16] Mohammad Hossein Rimaz, Mehdi Elahi, Farshad Bakhshandegan Moghadam, Christoph Trattner, Reza Hosseini, and Marko Tkalčič. 2019. Exploring the power of visual features for the recommendation of movies. In *Proc. of the ACM Conference on User Modeling, Adaptation and Personalization*. 303–308.

[17] Diana Patricia Sukhram and Amy Hsu. 2012. Developing reading partnerships between parents and children: A reflection on the reading together program. *Early Childhood Education Journal* 40, 2 (2012), 115–121.

[18] Marko Tkalcic, Andrej Kosir, and Jurij Tasic. 2011. Affective recommender systems: the role of emotions in recommender systems. CEUR-WS. org.

[19] Christoph Trattner and Dietmar Jannach. 2020. Learning to recommend similar items from human judgments. *User Modeling and User-Adapted Interaction* 30, 1 (2020), 1–49.

[20] Mengting Wan and Julian J. McAuley. 2018. Item recommendation on monotonic behavior chains. In *Proceedings of the 12th ACM Conference on Recommender Systems, RecSys 2018, Vancouver, BC, Canada, October 2-7, 2018*, Sole Pera, Michael D. Ekstrand, Xavier Amatriain, and John O'Donovan (Eds.). ACM, 86–94. https://doi.org/10.1145/3240323.3240369

[21] Mengting Wan, Rishabh Misra, Ndapa Nakashole, and Julian J. McAuley. 2019. Fine-Grained Spoiler Detection from Large-Scale Review Corpora. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, Anna Korhonen, David R. Traum, and Lluís Màrquez (Eds.). Association for Computational Linguistics, 2605–2610. https://doi.org/10.18653/v1/p19-1248

[22] Rabia M Yilmaz, Sevda Kucuk, and Yuksel Goktas. 2017. Are augmented reality picture books magic or real for preschool children aged five to six? *British Journal of Educational Technology* 48, 3 (2017), 824–841.

[23] Shuai Zhang, Lina Yao, Aixin Sun, and Yi Tay. 2019. Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys (CSUR)* 52, 1 (2019), 1–38.