A review of digital storytelling in language learning in children: methods, design and reliability

Clare Ong [1] and Vahid Aryadoust [2] *

*Correspondence:
vahid.aryadoust@nie.edu.sg
National Institute of Education,
Nanyang Technological
University, 1 Nanyang Walk,
Singapore, 637616, Singapore
Full list of author information is
available at the end of the article

Abstract

Digital storytelling (DS) is an innovative approach to language learning and teaching. Generally, DS refers to the form of storytelling that utilizes digital technology for expression. Scholars have established the value of DS in both traditional and non-traditional (online) classrooms as a tool to teach and learn languages. However, the research methods and standards of such studies continue to be overlooked even though the robustness of these studies needs to be established for the implementation of DS into the language curriculum for children. Thus, the present research conducted a systematic investigation of research methods, design and reliability in DS studies on children's language learning. We identified and extracted 50 documents from the Scopus database that satisfied the criteria of inclusion. In the initial evaluation, we coded every paper for (a) the research method applied, (b) research design and (c) reliability investigation of the instruments. We observed that most studies in the dataset used qualitative methods (n = 24, 48%) and most examined the effect of DS on children's written abilities (n = 25, 50%). The abilities of children to speak (n = 15, 30%) and read a specific language (n = 10, 20%) were investigated to a lesser extent. Yet, none of these studies investigated listening skills. Notably, more than 92% of DS studies on children language learning provided no evidence of reliability investigation. While we coded for eight reliability statistics in the DS dataset, only two of the indexes were identified. Among these methods, Cronbach's α was most often used to examine internal reliability, whereas correlation coefficient was applied to establish external reliability. Based on these findings, we offer some suggestions and guidelines for future DS research.

Keywords: Children, Digital storytelling, Educational technology, Language learning

Introduction

As natural storytellers, humans have relied upon storytelling to impart knowledge, beliefs and traditions (Suwardy et al., 2013). Since it is well-accepted that stories facilitate a

listener's understanding of complex concepts and ideas (Chung, 2006), schools have encouraged students to nurture their storytelling abilities, using both written and oral assessments, while educators have depended on stories to deliver their curriculum (Ballast et al., 2008). Furthermore, it has been suggested that the integration of storytelling in school curriculums creates an immersive environment for learners thereby improving learning outcomes (Coulter et al., 2007). When applied to the educational context, good storytelling allows students to easily recall earlier lessons, helping learners build a stronger foundation in that particular subject (Schank, 1990; Zull, 2002). Likewise, such stories may motivate students to pose questions and contribute to a livelier classroom (Bruner, 1996). Educators have utilised traditional storytelling methods that relied on language and verbal communication, and sometimes used basic visual aids to support the storytelling processes (Baim, 2015). However, modern efforts to digitalise storytelling have progressed significantly, especially with technological advancements. These advancements have increased the availability and affordances of new media devices, such as digital cameras, smartphones and software. Thus, a novel, multimedia form of digital storytelling (DS) has emerged. Through multimedia presentations and other computer-driven enhancements, the storytelling experience is said to have been substantially enhanced for both educators and students alike (Hung et al., 2012; Xu et al., 2011).

Notably, DS has been made popular with the initiatives by the Center for Digital Storytelling (or StoryCenter). The StoryCenter has identified the seven elements of DS (see Appendix A). Scholars have, in contrast, offered different definitions of DS. For instance, Xu et al. (2011) defined DS as "storytelling that is conducted using digital technology as the medium or method of expression, in particular using digital media in a computer-network environment" (p. 181). From these definitions, the common understanding of DS involves the use of multimedia, such as soundtracks to complement storylines. In addition, the immersive nature of DS differentiates it from the traditional methods of storytelling. Ever since DS has gained considerable traction, it has been examined from various lenses, such as mental and emotional health (Lim et al., 2022; Wexler et al., 2013), and from the perspectives of educators (Kildan & Incikabi, 2015; Yee et al., 2018) and students alike (Chen & Liu, 2019). For instance, Wexler et al. (2013) demonstrated that the DS process enabled youths to form more certain and positive identities, which are associated with positive youth health outcomes. On the other hand, Yee et al. (2018) studied the effects of DS on pre-service teachers who taught children literature. They observed that DS improved the pedagogical content knowledge, professional development, and teaching methods of the teachers (Yee et al., 2018). Furthermore, the educators became familiarized with DS and were more comfortable with the implementation of DS in their classrooms (Yee et al., 2018). Similarly, Diaz (2016) identified that the use of DS to train foreign language

teachers encourages them to implement DS in their lessons and, therefore, facilitate better knowledge transfer from educators to students (Diaz, 2016).

The present study will focus on the literature on DS in educational contexts from the perspective of students, especially in language learning. For student education, DS can be used in different forms, such as instructor-created stories, narrated documentaries of historical events and student-led and produced videos (Castañeda, 2013; Liang, 2019; Oppermann, 2007). In addition, DS has been noted to enhance educational outcomes and the attitudes of learners since it encourages student engagement (Chen & Liu, 2019; Hung et al., 2012). For instance, Hung et al. (2012) observed that DS technologies enabled elementary school students to adopt an active role in the storytelling process through participation in the creation and documentation of their personal digital stories. They demonstrated that the use of digital technologies in storytelling assisted learners in organising their own observations and story elements in a coherent manner, increased interaction, and collaboration among students to achieve better educational outcomes (Hung et al., 2012).

Likewise, Chen and Liu (2019) tested the effects of DS on language learning outcomes in an elementary school. They aimed to improve the writing of students by focusing on story structures through a six-stage structure that was integrated with digital books. The approach adopted includes six elements namely, "setting", "theme", "attempt", "consequence", "climax" and "resolution". These six elements support the understanding of story structures for students. Following their DS intervention, Chen and Liu (2019) reported a significant improvement in the written abilities of students, as compared to students in the control group who were assigned to a paper-based storybook. Furthermore, DS has improved classroom engagement, thereby encouraging positive attitudes towards writing, as reflected in the increased levels of affection and helpfulness towards their peers with regard to writing among the young student participants (Chen & Liu, 2019). This observation was supported by their results which revealed significant statistical differences for the interest dimensions measured, namely "triggering", "immersing" and "extending interests".

While the studies reviewed above explored the use of DS in traditional (language learning) classrooms, it should be noted that DS is not restricted to formal, and/or traditional educational venues (Baim, 2015). For instance, students enrolled in online courses can also learn novel content and benefit from instructional videos created based on DS (Baim, 2015). Baim (2015) observed that videos, multimedia presentations, audio recordings, and other similar tools are some of the most effective methods for learners who are engaged in remote learning to connect with their online instructors. Other studies have also recorded the success of DS in online contexts, regardless of whether they took place in a unidirectional method of instructor-generated content, or in a bidirectional format where learners generate

digital content as part of their coursework (Jenkins & Healey, 2012; Palacios, 2012; Rigney, 2010; Rossiter & Garcia, 2010). Furthermore, Lindgren and McDaniel (2012) identified that the implementation of advanced DS technologies and use of relevant computer applications can advance learning outcomes as compared to lecture-based classrooms. They observed that students better comprehended course concepts and reported increased satisfaction with their coursework following the DS intervention. DS features such as personalised interactive instructions help to take online learning in a different direction as students have control over their learning progress as opposed to classroom-based learning (Lindgren & McDaniel, 2012).

As demonstrated in the above research, DS improves online learning outcomes. While this educational technology is suitable for the current generation of students that are familiar with technology and experiments with new technological tools to keep themselves abreast with knowledge (Suwardy et al., 2013), DS has to be utilised by educators in a suitable manner to unlock its potential as a means to transform the language learning processes of children into one that focuses on production, collaboration, project management, teamwork and critical thinking (Moradi & Chen, 2019).

Research gap

Most of the studies surveyed above investigated DS and established its utility as a tool for language learning. Here, the utility of DS refers to the use of DS to nurture the language abilities of young learners. A review of DS could examine the utility, robustness and trustworthiness of previous research all of which are tied to the replicability of research. Current studies often fail to demonstrate the robustness of their results, as they seem to lack reliability or validity reports. Low reliability means that the results of a study could be confounded with errors and are not highly trustworthy (Field, 2018). Consequently, there are increased chances of Type I and II errors when unreliable data is used in follow-up influential statistics such as t-test and ANOVA. In addition, if the result of studies cannot be replicated, then they could be biased or not generalizable (Marsden et al., 2018).

In this study, we first attempt to address gaps in understanding through a systematic investigation of research methods in DS studies on language learning of children, with the aim to facilitate its integration with current education curriculums. We will consider the research features of these studies to evaluate their research standards. Second, we examine how the DS studies are designed to seek evidence about the trustworthiness and replicability of the studies. The replicability of a study is influenced by factors such as the theoretical framework employed, the design of the study – whether it was an experimental or cross-sectional study – as well as the test instruments used, data analysis method and the type of data analyses. Replicability in research establishes how reliable the findings of studies are and indicates whether or not their results can be generalized (Marsden et al.,

2018). Particularly, the demand for replication research has surged as language acquisition studies are increasingly challenged to prove their validity and reliability (Marsden et al., 2018).

Finally, previous research has highlighted that reliability reports of instruments and coding practices have been neglected in previous language learning or assessment research (e.g., Hou & Aryadoust, 2021). Therefore, we focus on the reliability reports of each research study. Generally, this factor is closely associated with the standards of language research (Johnson & Saville-Troike, 1992), and describes the extent to which a test measures what it claims to evaluate consistently and accurately (Chiedu & Omenogor, 2014; Nunally, 1982). This factor is also essential to ensure research standards, especially if the aim is to generate better research and knowledge on DS for educators to consult in their implementation of DS methods for the language learning of children.

The research questions of the study are, therefore, as follows:

1. What are the research methods adopted in DS studies of language learning of children?
2. How were DS studies on language learning of children designed?
3. How reliable are the instruments and coding practices of DS studies on language learning of children?

Methodology

Literature search

We carried out a broad literature search to find published research relevant to DS. To collect articles that explored DS, a literature search of DS studies on child language learning was conducted on the Scopus database. Scopus is recognised as the "largest single abstract and indexing database ever built" (Burnham, 2006, p.1), and was chosen because of the broader coverage of peer-reviewed journals it offers. Scopus is also intuitive to use as compared to other databases such as the Web of Science. Subsequently, we decided on the key search terms, "digital storytelling" and "child" to define the search for DS studies. These two phrases were integrated with other terms to achieve a comprehensive focus on language studies.

Adapted from Lado's (1961) skills and elements of language proficiency, which continues to be one of the most dominant approaches to language in learning and assessment, we targeted the four language skills namely, "listening", "reading", "speaking" and "writing". Lastly, to ensure that all relevant DS studies on child language learning were included, the general terms "language" and "oral" were added as search terms. No limitations on the year of publications were applied for the search protocol (see Appendix B for the search protocol). The year 2004 is the lower limit as it is the earliest year of

coverage by Scopus and the final date of publication inclusion was the end of September 2020 when the data was extracted.

Initial results

Overall, 117 documents were found with the search protocol. The most common type of document published were articles (49.6%), with conference papers being the second most common type of documents found (29.1%). Appendix C presents the total number of articles published by year in the Scopus database. An increased number of DS studies is recorded between 2008 and 2011. Specifically, the number of DS studies in 2011 were six-fold of that in 2008.

   The top three journals or publishers were Lecture Notes in Computer Science (LNCS), including the subseries Lecture Notes in Artificial Intelligence (LNAI) and Lecture Notes in Bioinformatics (LNBI), Digital Education Review and ACM International Conference Proceedings Series, with a total of five, four and three papers published between 2004 and 2019 respectively (for documents per year by source, see Appendix D). It should be noted that a high number of publications from LNCS, LNAI, and LNBI are recorded. This is because the three journals tend to feature studies in computer science and information technology research, and teaching – areas which are relevant to DS.

   The top five countries or territories producing the greatest number of articles were the United States (US), the United Kingdom (UK), Italy, Canada and Australia (see Appendix E for list of countries or territories). The top three academic institutes were Università della Svizzera italiana (n = 5), Universidad de Oviedo (n = 4) and Bournemouth University (n = 4). However, two of the top three academic institutes publishing articles were not from these top five countries or territories; only Bournemouth University is located in the UK. Based on the Scopus dataset, DS studies tend to be Eurocentric.

Inclusion and exclusion criteria

In this analysis, the inclusion criteria for research consists of four elements: target language, technology used, publication source and research design. Table 1 details the inclusion and

Table 1 Inclusion and exclusion criteria

Inclusion criteria: The paper…	Exclusion criteria: The paper…
1. investigated language learning of young children.	1. investigated language learning of groups other than young children.
2. used at least one digital device and/or software for language learning.	2. did not use a digital device and/or software for language study, which do not lend themselves to digital storytelling.
3. used qualitative or quantitative methods for data collection and/or analysis.	

Fig. 1 Graph represents publication dates of DS studies included in the final dataset
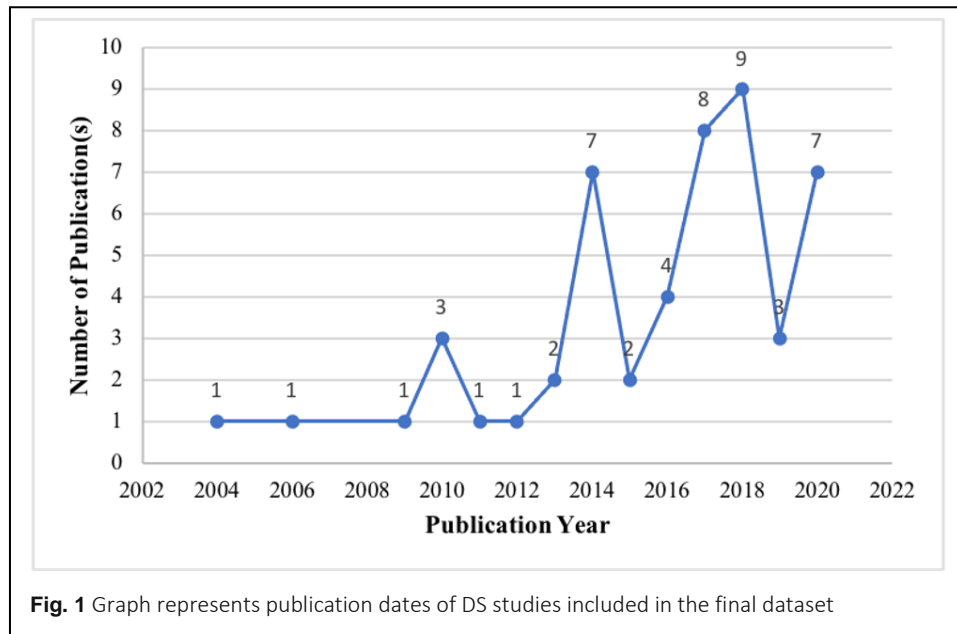
exclusion criteria, which was designed following Grotjahn's (1987) framework. According to Grotjahn (1987), research design is characterised by three components namely: (1) data collection methods, (2) data analysis, and (3) characteristics of the data.

In the initial review, a total of 55 publications were irrelevant and therefore excluded, resulting in 62 papers remaining. Thereafter, the inclusion and exclusion criteria in Table 1 were applied to assess the remaining 62 papers and determine their eligibility. Finally, 12 papers were excluded since they did not investigate the language learning of young children. The remaining 50 studies were included for coding and analysis. The final dataset which consists of DS studies published from 2004 to 2020 is illustrated in Figure 1.

In the dataset, an exponential growth in papers was recorded between 2012 and 2014. Specifically, an upward trend is observed in the publication of DS studies from 2015, with a record number of nine studies published in 2018. From the dataset, we observed that the publication of DS studies peaked in the years 2010, 2014 and 2018. It must be noted that in the dataset, the number of publications in 2020 is not representative of the total publications for that year, as the search period in Scopus concluded in September 2020. The years of publication indicated here represent the online publication years, as several papers may have received an issue number after they have been presented online.

Coding scheme and reliability

We adapted a coding scheme from various sources (e.g., see Appendix F) to record the research methods adopted, research design, and reliability analysis, which respectively pertain to research questions 1, 2, and 3. The scheme includes multiple variables which

were classified and presented in Table 2, which offers a detailed overview of the variables used in the study. The table is organized based on the research questions.

To ensure that all admissible publications were included in the review, a second coder was invited to review and code these articles for target language, research design and the use of digital devices and/or software. The second coder was a researcher from the university where the study was conducted. The inter-coder agreement was 93.2%, which suggests high reliability in the exclusion of irrelevant publications. Thereafter, the disagreement was resolved between the coders in an online meeting conducted over Zoom.

Data analysis

In response to the first question, data were organised according to the research methods. For example, we were able to recognise the frequency and percentage of the application of different research methods such as quantitative, qualitative, and mixed methods. To answer

Table 2 Overview of criteria used in the study

Category	Research question(s)	Criteria	Description of criteria
Research method	RQ1: What are the research methods adopted in DS studies of language learning of children?	Method of the study	What research method was adopted? (i.e., true and quasi experiment, cross-sectional studies)
		Data analysis	How were the data analysed in the research? (i.e., quantitative, qualitative, mixed-method, review or synthesis of the literature)
		Types of data analysis	What type of data analysis method is applied to the research? (i.e., phenomenological method, case studies)
Research design	RQ2: How were DS studies on language learning of children designed?	Target language skills	Which language skills are examined in the studies? (i.e., Writing, Reading, Listening, Speaking)
		Language of research	Which language is the research conducted in? (i.e., English, Non-English)
		Participants' average age	What is the average age of participants in the published research?
Reliability analysis	RQ3: How reliable are the instruments and coding practices of DS studies on language learning of children?	Reliability index used	What is the reliability index used in the research? (i.e., external reliability, internal reliability)
		Type of reliability index used	What type of reliability index is reported in the research? (i.e., Cronbach's α, correlation coefficient)

the second research question, data were coded in terms of specific research designs. We analysed the descriptive features of studies for targeted language skills as well as the languages that were investigated in the studies. Finally, the frequency of the methods applied for reliability assessment was computed, with consideration of the target language skills. An in-depth examination of the findings of the studies is synthesized and presented in the Discussion.

Results

Research method

The research methods adopted in the dataset of the DS studies are summarised in Table 3. Appendix F describes in detail each of these research methods. The quantitative methods included descriptive design (n = 7, 14%), correlational analysis, quasi-experimental design, and experimental design (n = 2, 4%). The qualitative methods include case studies (n = 8, 16%), grounded theory studies (n = 3, 6%), phenomenology (n = 9, 18%), ethnography (n = 4, 8%), and historical method, which altogether accounted for 48% (n = 24) of total research methods used. Among the qualitative methods, phenomenology (n = 9, 18%) was the most frequently used method followed by case studies (n = 8, 16%), ethnography (n = 4, 8%), and grounded theory studies (n = 3, 6%). We note that the experiences of participants in language studies are often interpreted and understood using the

Table 3 Breakdown of research methods used

Research methods	Research types		# of studies using the type	%
Quantitative	Descriptive		7	14
	Correlational		0	0
	Quasi-Experimental		0	0
	Experimental		2	4
		Subtotal	9	18
Qualitative	Case Study		8	16
	Grounded Theory		3	6
	Phenomenology		9	18
	Ethnography		4	8
	Historical		0	0
		Subtotal	24	48
Mixed Method	Sequential Explanatory Design		2	4
	Sequential Exploratory Design		0	0
	Sequential Transformative Design		1	2
	Concurrent Triangulation Design		8	16
		Subtotal	11	22
Review or Synthesis of the literature	Not Applicable		6	12
		Subtotal	6	12
Total Number			50	100

phenomenological method (Tuohy et al., 2013; Wilson, 2014). In contrast, the historical method was the least used qualitative research method since none of the studies utilised the method in the dataset.

Investigation of research design

To explore the research design of the studies, the descriptive features of the participants of the studies were investigated. Table 4 presents the average age of participants investigated in the studies in the DS dataset. Among the 50 studies examined, 21 (42%) studies investigated participants of ages six to 11. Of these 21 studies, speaking (n = 10, 20%) was the most commonly investigated language skill, followed by writing (n = 8, 16%) and reading (n = 3, 6%).

For other descriptive features relevant to research design such as target skills investigated and the language of the studies, the same reporting technique as in Table 4 was used (see Table 5, for further details). Firstly, writing was the most investigated language component for both English and non-English studies, regardless of the average age of participants (n = 25, 50%). This is followed by speaking (n = 15, 30%) and reading (n = 10, 20%). Notably, none of these studies investigated the listening skill.

Table 4 Breakdown of average participant age

Targeted language skills	Average age of participants		# of studies investigating participant age group	%
Listening	Ages 0-5		0	0
	Ages 6-11		0	0
	Ages 12-17		0	0
	Ages 18-23		0	0
	Age 24 and above		0	0
		Subtotal	0	0
Reading	Ages 0-5		6	12
	Ages 6-11		3	6
	Ages 12-17		0	0
	Ages 18-23		1	2
	Age 24 and above		0	0
		Subtotal	10	20
Speaking	Ages 0-5		5	10
	Ages 6-11		10	20
	Ages 12-17		0	0
	Ages 18-23		0	0
	Age 24 and above		0	0
		Subtotal	15	30
Writing	Ages 0-5		12	24
	Ages 6-11		8	16
	Ages 12-17		2	4
	Ages 18-23		1	2
	Age 24 and above		2	4
		Subtotal	25	50
Total Number			50	100

Table 5 Breakdown of the languages investigated

Targeted skills		# of studies investigating the skill	%
English			
Listening		0	0
Reading		8	16
Speaking		11	22
Writing		20	40
	Subtotal	39	78
Non-English			
Listening		0	0
Reading		2	4
Speaking		4	8
Writing		5	10
	Subtotal	11	22
Total Number		50	100

Reporting reliability

Table 6 demonstrates the reliability of research instruments and data coding practices in the studies in the DS dataset. We note that few papers carried out reliability analysis (n = 4, 8%) and the majority of the studies did not report reliability (n = 46, 92%). Among the studies, writing (n = 25, 50%) was the most frequently investigated language skill, followed by speaking (n = 15, 30%) and reading (n = 10, 20%). Some of the studies investigated internal reliability (n = 3, 6%). External reliability of the instruments used was significantly under-researched, with only one study on writing (n = 1, 2%) that investigates the external reliability of the test instrument and no studies on other language components.

   We used our coding scheme to determine which one of the eight types of reliability statistics were applied in the DS dataset namely, Cronbach's α (n = 3, 6%), Cohen's κ, Fleiss's κ, Kuder-Richardson reliability coefficients, Spearman-Brown prophecy formula, correlation coefficient (n = 1, 2%), Kendall's W, and coder/rater agreement percentage (see Table 7). However, we identified only two out of the eight reliability statistics in the DS dataset. The most commonly used index to assess internal consistency in the DS studies was Cronbach's α, while only correlation coefficient was applied for external reliability investigation.

Table 6 Investigation of reliability

Target skills	# of studies reporting internal reliability	%	# of studies reporting external reliability	%
Listening	0	0	0	0
Reading	1	2	0	0
Speaking	1	2	0	0
Writing	1	2	1	2
Total	3	6	1	2

Notes: Studies that reported multiple reliability indices were counted several times.

Table 7 Reliability indices used in the DS dataset

Types of reliability	Reliability indices	# of studies using the index	%
Internal reliability	Cronbach's α	3	6
	Cohen's κ	0	0
	Fleiss's κ	0	0
	K-R reliability coefficient	0	0
	S-B prophecy formula	0	0
Subtotal		3	6
External reliability	Correlation coefficient	1	2
	Kendall's *W*	0	0
	Coder/rater agreement percentage	0	0
Subtotal		1	2
Total Number		4	8

Notes: K-R reliability denotes Kuder-Richardson reliability coefficient. S-B prophecy formula refers to Spearman-Brown prophecy formula.

Discussion

This study set out to investigate the methodological quality of DS research on the language learning of children. In the following sections, we discuss the findings of the analysis of the three research questions of the study.

Research question one: research methods in DS research

Overall, the findings revealed that most studies in the DS dataset used a qualitative research method (n = 24). For language research, there are some benefits of qualitative methods (Chalhoub-Deville & Deville, 2008; Denzin, 1989). For instance, Denzin (1989) identified that they enable researchers to form a rich, vivid account of the emotions, ideas and experiences of the participants based on interpretation. Likewise, Chalhoub-Deville and Deville (2008) reported that qualitative methods allow for the collection of in-depth perspectives into issues relevant to the design, administration and interpretation of language assessment. While we acknowledge these benefits, it must be noted that such methods are dependent on an "interpretive and naturalistic approach to its subject matter" (Flick, 2014, p. 542). Therefore, qualitative methods are not appropriate for research investigations that are intended to be generalizable to a larger population of language learners.

  Phenomenological studies were most common in the dataset. The advantage of phenomenology research is that it allows the researcher to investigate the perception and attitude of DS users towards its utility, thus broadening "our understanding of the complex phenomena involved in learning, behaviour, and communication that are germane to our field" (Neubauer et al., 2019, p. 95). The limitation of phenomenology-based research in DS was the weak alignment between the research processes adopted and the theoretical roots that underlie the DS phenomenon investigated (Neubauer et al., 2019). In addition, a

well-known limitation of phenomenology in general has been highlighted by a number of scholars who argued the phenomenological approach tends to avoid contextual sensitivities due to its emphasis on meanings and experiences (Silverman, 2010; Tuhoy et al., 2014).

In all, the examination of storytelling processes is centred on the learners' experiences, which makes qualitative methods a natural choice for researchers, because of the interpretive and naturalistic approach. However, researchers must consider the aims of their studies in the selection of research methods. For instance, a study that aims to prove the effects, as opposed to processes, of DS on children's language learning would have to adopt some means of quantification. In this case, the application of qualitative methods would detract from the robustness of the study. The availability of cutting-edge technologies such as eye tracking and neuroimaging for language and education research make it possible for researchers to quantify the cognitive, emotional, and behavioural processes underlying language learning in DS research. Thus, we stress that for research claims to be extrapolated beyond the sample of participants in the study, it would behoove researchers to use quantitative designs founded upon theoretically justifiable models of DS, language learning, and educational psychology.

Research question two: investigation of research design

Research question two examined the research design of studies in the DS dataset. In comparison to the other language competencies, most DS studies tend to focus on the written abilities of children (n = 50). Coincidentally, there has been a wealth of literature beyond this comprehensive review that established the value of DS in improving the written work of children (see Chen & Liu, 2019).

We found that previous research investigated and supported the use of DS to facilitate the speech development of children (Hwang et al., 2016; Lestariyana & Widodo, 2018), and enhance their abilities to read and comprehend texts (Hamdy, 2017; Liu et al., 2019). For instance, Hwang et al. (2016) examined the effects of DS on the speaking skills of children who were non-native speakers of English. They identified that the process of recording composed stories orally allowed learners to practice speaking in the target language, thereby enhancing their English-speaking performance, as well as their general learning achievement in English (Hwang et al., 2016). Similarly, Lestariyana and Widodo (2018) tested the effects of DS on the speech development of English-as-a-foreign-language speakers in an Indonesian primary school. They implemented DS as a pedagogical innovation with the aim to boost the confidence of students in speaking English. For their experiment, students were encouraged to narrate their personal digital stories. Eventually, Lestariyana and Widodo (2018) observed that the process of reviewing and editing their voice recording allowed the students to become more confident in speaking English. Furthermore, students were also observed to make use of the learned

vocabularies, which include difficult and technical terms (i.e., "propagation", "pruning", and "weeding") to better express their ideas verbally in English (Lestariyana & Widodo, 2018).

In contrast, Liu et al. (2019) studied the effects of DS on the reading skills of elementary school students. They observed that the students became proficient in oral reading through participation in the DS intervention (Liu et al., 2019). Concurrently, the DS intervention fostered a collaborative language learning environment, which encouraged students' engagement and sustained their learning progress (Liu et al., 2019). Likewise, Hamdy (2017) identified that the use of DS to teach students reading comprehension achieved better outcomes than conventional modes of instruction. Specifically, he noted that DS approaches combine visual images with written text to enhance and accelerate the students' abilities to read (Hamdy, 2017).

However, there are some limitations to the preceding studies with respect to research design. First, these studies have not utilised any theoretical framework to examine the issue rigorously and systematically. Thus, their findings appear to lack theoretical basis and explanatory power. In addition, we observed that some studies attempted to demonstrate the effects of DS on the children's motivation to learn languages or other intangible factors. However, it is not certain how these researchers delineated and operationalized the target constructs to assess the participants' levels of motivation, since no robust framework was employed in the first place. This limitation calls for further investigation of the validity and accuracy of the findings of previous DS research. Ioannidis's (2005) seminal research showed that the probability that findings of a stream of research are reliable depend on a variety of factors including "study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field" (p. 696). According to Ioannidis (2005, p. 696), the findings of previous research would be less likely to be reliable when "the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance." Lamentably, much of the information underscored by Ioannidis (2005) are missing in the DS published literature. Therefore, future DS researchers should consider elaborating on these factors in drawing conclusions from the data they collect and examine.

Next, the DS studies explored a wide-ranging age group of "children". From this, we infer a lack of consensus among studies on the age range that best characterises "children". As a consequence, the proponents of DS programmes may struggle to demonstrate the language learning outcomes for this particular demographic. Instead, DS research should be informed by educational psychology research where the age group of students are

thoroughly investigated. Lenneberg (1967) first established that language acquisition is limited to a "critical period" which extends from early infancy until puberty. This suggests that young children are better language learners than adults. Thereafter, Johnson and Newport (1989) tested Lenneberg's (1967) assumption and compared the English proficiency of Asians (Koreans and Chinese) who arrived in the United States (US) between 3 and 39 years old to those who lived in the United States between 3 and 26 years old. Their findings corresponded with Lenneberg's (1967) as they identified a strong correlation between the age at which participants arrived in the US and their second language acquisition and performance. In particular, the participants who were taught English before the age of 10, were expected to reach native proficiency in the language (Johnson & Newport, 1989). Recently, Hartshorne et al. (2018) have found support for the existence of a "critical period" for second language learning. They also estimate that the rate of language learning declines at 17.4 years old (Hartshorne et al., 2018). From this, we observe the influence of the participants' age on language learning outcomes. Thus, it is critical to establish the age group of language learners to ensure that results of future studies can be replicated.

Lastly, we identified a lack of studies in the dataset which investigated the listening skills of children. Listening is an important language skill and serves as the most crucial skill for receiving language input, which is essential in language learning. Thus, future research should examine whether the digitization of input via DS can benefit young language learners.

We recognize that it is difficult to differentiate the key language competencies examined in a research. This is because DS methods create an immersive environment that evokes the various senses of the learner, such as visual and auditory senses through vibrant sights, sounds, and even encourages speech. Furthermore, DS can be considered as a tool that simultaneously increases the digital literacy of language learners. For instance, Churchill et al. (2008) reported that interaction with advanced technologies used for DS enhanced the abilities of students to make sense of and represent multimodal texts. Specifically, the multimodal characteristic of digital texts nurtures digital literacy (Churchill et al., 2008). All of these suggest that DS is capable of developing various language competencies simultaneously, which would make it an appropriate tool for engaging young language learners in authentic language learning where different language components are not disaggregated. Thus, more focus should be placed on the processes that enable the simultaneous enrichment of various language competencies, instead of a narrow focus on individual language competencies. We recognize the difficulties of conducting broader studies, and, thus, attempt to offer some suggestions to guide for future researchers in the next section.

Research question three: reporting of reliability

An overwhelming majority of the studies in the DS dataset (92%) have not reported the reliability of test instruments and the results achieved. As mentioned earlier, the exclusion of reliability analysis is another factor that affects the robustness of DS studies. Concurrently, it hinders the accurate inference of research outcomes (Grabowski & Oh, 2018). Moreover, DS researchers were more likely to focus on reporting internal reliability of instruments used, whereas only one research paper analysed external reliability. Coincidentally, internal reliability investigation has been popular in applied linguistics studies (Grabowski & Oh, 2018). We also observed that the most used index in DS studies on the language learning of children was Cronbach's α, followed by the correlation-coefficient. Nonetheless, this observation raises the question of the precision of the instruments used, as well the results obtained. Crucially, it echoes the earlier concerns of researchers over the neglect of reliability reports of instruments and coding practices (Hou & Aryadoust, 2021).

  When the precision of the instruments used is low in quantitative research or when there is no or minimal agreement between coders in a qualitative research design, the results of the study are convoluted with error and consequently cannot be trusted. This results in multiple drawbacks such as undermining the replicability of the results across other contexts and erroneous conclusions. Reliability, therefore, becomes an essential factor to be taken heed of in the future DS research.

Implications of the study

Following this review, we make several recommendations for future research. First, future DS research should attempt to adopt quantitative methods, while considering the aforementioned research guidelines expounded by Ioannidis (2005). This shift is crucial to encourage future implementation of DS into the curriculum of children as a novel approach for them to learn languages. We make this suggestion in view of the insufficiencies of qualitative studies mentioned above, as well as the fact that policymakers and/or stakeholders tend to ascribe low credibility to the results of qualitative studies (Rahman, 2016; Sallee & Flood, 2012), likely because qualitative research, despite its utility, is not appropriate for drawing generalizable conclusions. In fact, Ravitch (2010) found that stakeholders in the US educational field deemed quantitative studies as more credible, and relied upon them to account for the performances of students and educators alike.

  As reflected in the Scopus dataset, DS research tends to be Eurocentric. Most of these studies reported the outcomes of DS interventions in European countries. In comparison, the experiences of using DS in Asian or African classrooms for language learning are overlooked. Hence, DS researchers should conduct and compare DS interventions across other geographic areas to consider whether cultural differences may influence the

educational outcomes of DS and whether research outcomes are reproducible across different cultures.

Next, DS researchers should examine and indicate the reliability and possibly the validity of the instruments and/or coding schemes utilised in their studies to increase the reliability of DS studies. Both reliability indices are overlooked, with only one paper in the dataset reporting external reliability. Therefore, we advocate for an increased focus on external reliability as it ensures the replicability of the test results. Finally, DS researchers should consider the exploration of multiple language competencies in their studies, especially with consideration of the immersive nature of DS methods that simultaneously engages the various senses of the learners. Additionally, increased calls for the adoption of multimodality in language learning are observed (Perniss, 2018). Some research on the use of multimodal teaching methods have also highlighted its relevance in higher education institutions (Reid et al., 2016). This strand of research stresses the integration of different language skills such as listening, reading and viewing, which are increasingly emphasized in language learning and assessment. DS, with its integration of several forms of media, thus offers exciting opportunities to operationalize multimodality in language learning and assessment.

Finally, DS researchers should investigate language learning in a more holistic manner to focus on the processes that enable the simultaneous enrichment of various language competencies, rather than a narrow focus on individual language competencies.

Conclusion

The present study investigated the methodological quality of DS studies on children's language learning. From a comprehensive literature search in Scopus, we identified 117 DS studies relevant to children language learning, of which 50 that met our criteria were subsequently coded and examined to address our three research questions.

With regard to the first question, most studies in the DS dataset utilised qualitative research methods, as compared to quantitative methods, mixed-methods or literature reviews. However, the reliance on qualitative methods may be problematic, especially if the aim is to convince international stakeholders in the educational field to implement DS methods into their existing curriculums as a novel approach to improve language learning outcomes for children – and even more so as a viable alternative for remote learning that continues to take place under the COVID-19 global pandemic. Proponents of such programmes must be able to demonstrate learning outcomes.

The second research question examined the research design (descriptive features) of studies in the DS dataset. While most DS studies tend to focus on the written abilities of learners, we opined that it is difficult to differentiate the key language competencies

examined in a research. Instead, DS should be investigated as a tool that can be used to nurture different language competencies at the same time.

Finally, the third research question addressed the reliability reports of instruments and coding practices. We found that more than 92% of DS studies on children language learning provided no evidence of reliability investigation. This observation exemplifies existing concerns of researchers over the neglect of reliability reports of instruments and coding practices. We hope the results of this study will be useful as inspiration to researchers in the field of applied linguistics and encourage the implementation of DS into the language curriculum for children around the world.

Appendices

Appendix A

| The Seven Elements of Digital Storytelling	
Element of DS	Description
Point of View	What is the main point of the story and what is the perspective of the author?
A Dramatic Question	A key question that keeps viewer's attention and will be answered by the end of the story.
Emotional Context	Serious issues that come alive in a personal and powerful way and connects the story to the audience.
Gift of Voice	A way to personalise the story to help the audience understand the context.
Power of Soundtrack	Music or other sounds that support and embellish the storyline.
Economy	Using just enough content to tell the story without overloading the viewer.
Pacing	The rhythm of the story and how slowly or quickly it progresses.

Source: Adapted from the University of Houston

Appendix B

The following search protocol was used in Scopus to generate datasets:

((TITLE-ABS-KEY ("digital storytelling") AND TITLE-ABS-KEY (child*))) AND (language OR reading OR listening OR speaking OR writing OR oral)

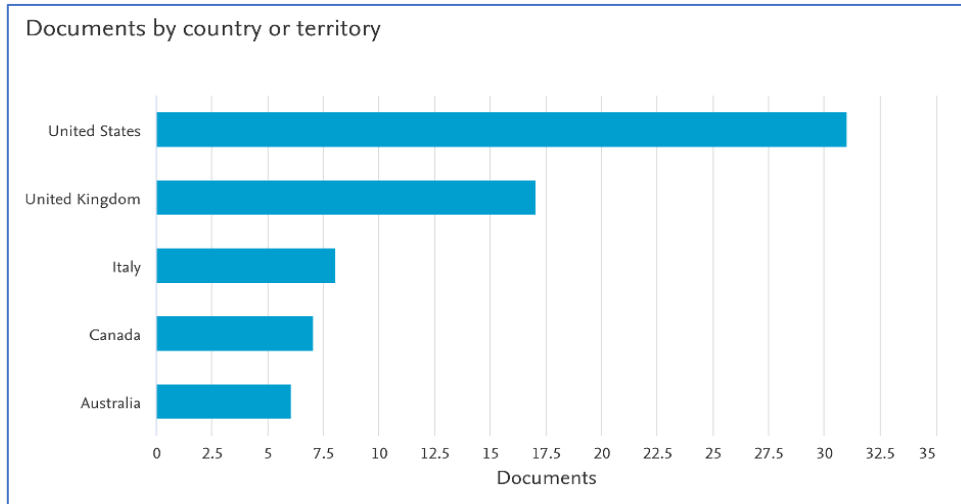## Appendix C

Documents by year
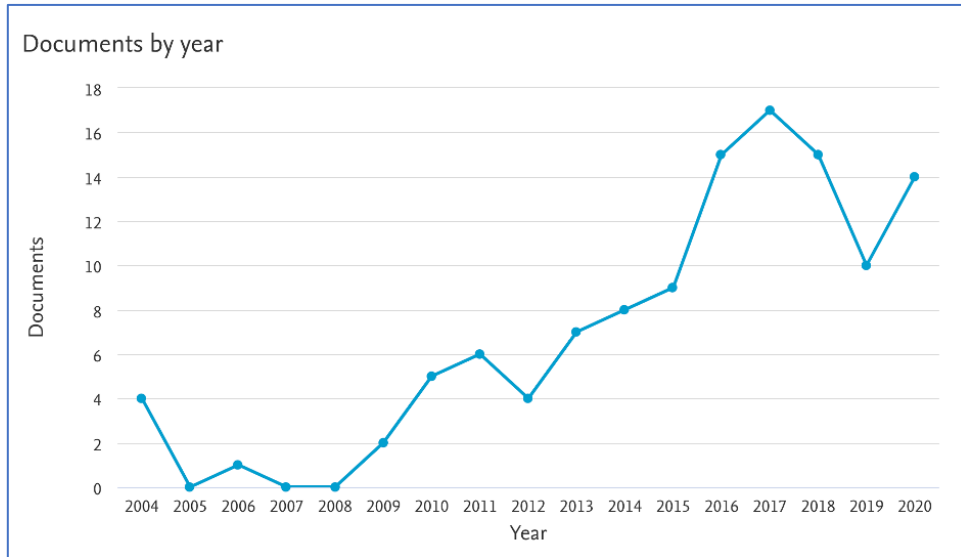
Notes: The graph above indicates the total number of documents released per year in the Scopus database.

Appendix D

Documents per year by source
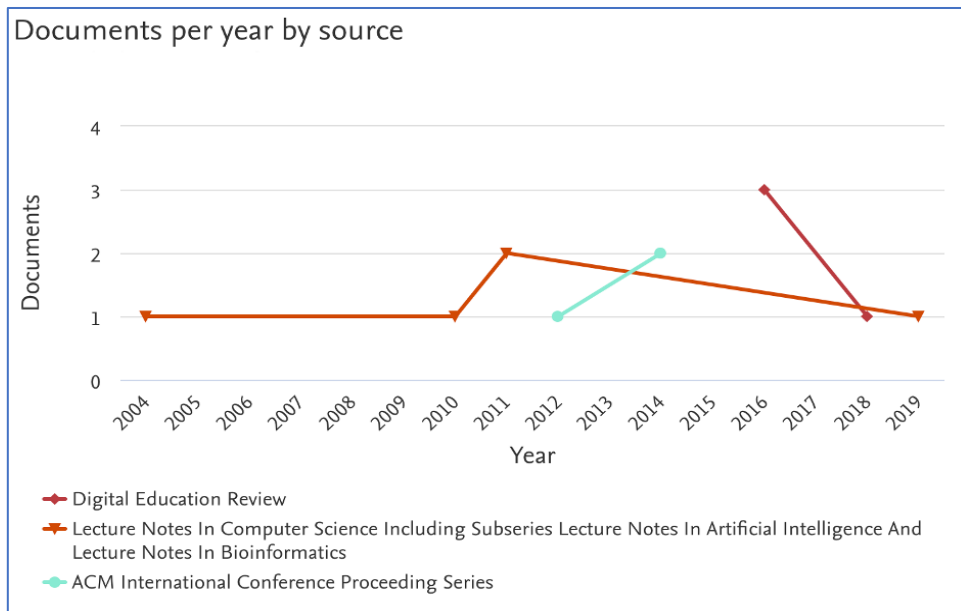
Notes: Graph above indicates the total number of documents released per year by source.

Appendix E

Documents by country or territory

Notes: The graph above indicates the total number of documents released per year by countries or territories.

Appendix F

Quantitative Research Methods*	
Research Type	Description
Descriptive	Describes the current status of an identified variable.
Correlational	Seeks to determine the extent of a relationship between two or more variables using statistical data.
Quasi-Experimental	Seeks to establish cause-effect relationships among the variables.
Experimental	Uses the scientific method to establish the cause-effect relationship among a group of variables that make up a study.
Qualitative Research Methods**	
Research Type	Description
Case Study	Reveals a phenomenon by studying in depth a single case example of the phenomenon.
Grounded Theory	Comprehends the social and psychological processes that characterize an event or situation.
Phenomenology	Describes the structures of experience as they present themselves to consciousness, without recourse to theory, deduction, or assumptions from other disciplines.
Ethnography	Focuses on the sociology of meaning through close field observation of sociocultural phenomena.
Historical	Systematic collection and objective evaluation of data related to past occurrences in order to test hypotheses regarding causes, effects, or trends of these events that help to explain present and future events.
Mixed Methods***	
Research Type	Description
Sequential Explanatory Design	A two-phase project that is more focused on quantitative data and consists of: (1) Quantitative Data Collection and Analysis; and (2) Qualitative Data Collection and Analysis. Phase (1) informs Phase (2). For Example: Survey data informs interviews.
Sequential Exploratory Design	A two-phase project that is more focused on qualitative data and consists of: (1) Qualitative Data Collection and Analysis; and (2) Quantitative Data Collection and Analysis. Phase (1) informs Phase (2). For Example: Focus group shapes survey questions.
Sequential Transformative Design	A four-phase project that consists of two data collection series: (1) Qualitative Data Collection and Analysis; and (2) Quantitative Data Collection and Analysis. Phase (1) informs Phase (2) and vice versa. Process is repeated twice.
Concurrent Triangulation Design	A one-phase project that collects quantitative and qualitative data at the same time. Results from qualitative data are compared to those from quantitative data. Analysis for qualitative and quantitative data is conducted separately.
Literature Review****	
Research Type	Description
NA	A search and evaluation of the available literature in the given subject or chosen topic area.

Notes: Description of research methods are adapted from various sources, indicated with asterisks.

Source/Legend: *Winston Salem State University ; **University of Wisconsin-Madison ; ***Georgia State University Library ; ****Royal Literary Fund .

Abbreviations
DS: Digital storytelling; LNAI: Lecture Notes in Artificial Intelligence; LNBI: Lecture Notes in Bioinformatics; LNCS: Lecture Notes in Computer Science; StoryCenter: Center for Digital Storytelling; US: The United States; UK: The United Kingdom.

Authors' contributions
Clare Ong and Vahid Aryadoust conceived the study, created the coding scheme, and discussed the content of the paper. Clare Ong wrote the paper, and Vahid Aryadoust reviewed and revised it.

Authors' information
Clare Ong is a Final Year Public Policy & Global Affairs Undergraduate at the School of Social Sciences of Nanyang Technological University (NTU), Singapore. She is a student researcher under NTU's Undergraduate Research Programme (URECA). As a student researcher, she has (co)authored research in language learning and public administration.

Vahid Aryadoust is Associate Professor of language assessment at the National Institute of Education of Nanyang Technological University, Singapore. His areas of interest include language assessment, quantitative methods, eye tracking, and brain imaging in language assessment. Vahid has published his research in *Computer Assisted Language Learning*, *Language Testing*, *System*, *Current Psychology*, *Language Assessment Quarterly*, *Assessing Writing*, *Educational Assessment*, *Educational Psychology*, etc. He has also (co)authored a number of book chapters and books that are published by Routledge, Cambridge University Press, Springer, Cambridge Scholar Publishing, Wiley Blackwell, etc. Vahid has also led a number of assessment research projects supported by educational fund-providers in Singapore, USA, UK, and Canada. He is a member of the Advisory Board of multiple international journals and has been awarded the Intercontinental Academia Fellowship (2018–2019). His YouTube channel, *Statistics and Theory*, has been awarded the John Cheung Social Media Award, 2020, which recognizes exemplary and innovative use of social media. The channel is available from: https://www.youtube.com/user/vahidaryadoust .

Availability of data and materials
Not applicable.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1] School of Social Sciences, Nanyang Technological University, Singapore. [2] National Institute of Education, Nanyang Technological University, Singapore.

References
Baim, S. A. (2015). Digital storytelling: Conveying the essence of a face-to-face lecture in an online learning environment. *Journal of Effective Teaching, 15*(1), 47–58. https://eric.ed.gov/?id=EJ1060431
Ballast, K., Stephens, L., & Radcliffe, R. (2008). The effects of digital storytelling on sixth grade students' writing and their attitudes about writing. In K. McFerrin et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2008* (pp. 875–879). Association for the Advancement of Computing in Education.
Burnham, J. F. (2006). Scopus database: A Review. *Biomedical Digital Libraries, 3*(1), 1. https://doi.org/10.1186/1742-5581-3-1
Bruner, J. (1996). *The culture of education*. Harvard University Press.
Castañeda, M. E. (2013). "I am proud that I did it and it's a piece of me": Digital storytelling in the foreign language classroom. *CALICO, 30*(1). https://www.equinoxpub.com/CALICO/issue/view/1920
Chalhoub-Deville, M., & Deville, C. (2008). Utilizing psychometric methods in assessment. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopaedia of language and education* (2nd ed., Vol. 7, pp. 211–224). Springer Science + Business Media LLC.

Chen, Z. H., & Liu, W. Y. (2019). A six-stage story structure approach for elementary students' story production: Quality, interest, and attitude. *Computer Assisted Learning Language, 34*(1-2), 184–207. https://doi.org.remotexs.ntu.edu.sg/10.1080/09588221.2019.1688354

Chiedu, R. E., & Omenogor, H. D. (2014). The concept of reliability in language testing: Issues and solutions. *Journal of Resourcefulness and Distinction, 8*(1), 1–9. https://globalacademicgroup.com/journals/resourcefulness/The%20Concept%20of%20Reli ability%20in.pdf

Chung, S. K. (2006). Digital storytelling in integrated arts education. *Journal of Arts Education, 4*(1), 33–50.

Churchill, N., Lim, C. P., Oakley, G., & Churchill, D. (2008). Digital storytelling and digital literacy learning. *Readings in Education and Technology: Proceedings of ICICTE 2008* (pp. 418–430). https://www.academia.edu/831867/Digital_storytelling_and_digital_literacy_learning

Coulter, C., Michael, C., & Poyner, L. (2007). Storytelling as pedagogy: An unexpected outcome of narrative inquiry. *Curriculum Inquiry, 37*(2), 103–122. https://doi.org/10.1111/j.1467- 873X.2007.00375.x

Denzin, N. K. (1989). *Interpretive interactionism*. Sage.

Diaz, M. A. (2016). Digital storytelling with pre-service teachers: Rising awareness for refugees through ICTS in ESL primary classes. *Digital Education Review, 30*, 1–16. https://files.eric.ed.gov/fulltext/EJ1125503.pdf

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage.

Flick, U. (2014). *An introduction to qualitative research* (5th ed.). Sage Publications Ltd.

Georgia State University Library. (2022). *Mixed methods*. https://research.library.gsu.edu/c.php?g=1050115&p=7622501

Grabowski, K. C., & Oh, S. (2018). Reliability analysis of instruments and data coding. In A. Phakiti, P. De Costa, L. Plonsky & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 541–565). Palgrave Macmillan.

Grotjahn, R. (1987). On the methodological basis of introspective methods. In C. Færch & G. Kasper (Eds.), *Introspection in second language research* (pp. 54–81). Multilingual Matters.

Hamdy, M. F. (2017). The effect of using digital storytelling on students' reading comprehension and listening comprehension. *Journal of English and Arabic Language Teaching, 8*(2), 112–123. https://www.researchgate.net/publication/342364293_The_Effect_of_Using_Digital_Story telling_on_Students'_Reading_Comprehension_and_Listening_Comprehension

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition, 177*, 263–277. https://doi.org/10.1016/j.cognition.2018.04.007

Hou, Z., & Aryadoust, V. (2021). A review of the methodological quality of quantitative mobileassisted language learning research. *System, 100*, 102568. https://doi.org/10.1016/j.system.2021.102568

Hung, C. M., Hwang, G. J., & Huang, I. (2012). A project-based digital storytelling approach for improving students' learning motivation, problem-solving competence and learning achievement. *Educational Technology and Society, 15*(4), 368–379.

Hwang, W.-Y., Shadiev, R., Hsu, J.-L., Huang, Y.-M., Hsu, G.-L., & Lin, Y.-C. (2016). Effects of storytelling to facilitate EFL speaking using web-based multimedia system. *Computer Assisted Language Learning, 29*(2), 215–241. https://doi.org/10.1080/09588221.2014.927367

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Med, 2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Jenkins, M., & Healey M. (2012). Developing and embedding inquiry-guided learning across an institution. *New Directions for Teaching and Learning, 129*, 27–37. https://doi.org/10.1002/tl.20004

Johnson, D. M., & Saville-Troike, M. (1992). Validity and reliability in qualitative research on second language acquisition and teaching: Two researchers comment. *TESOL Quarterly, 26*(3), 602–605. https://doi.org/10.2307/3587189

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Kildan, A. O., & Incikabi, L. (2015). Effects on the technological pedagogical content knowledge of early childhood teacher candidates using digital storytelling to teach mathematics. *International Journal of Primary, Elementary and Early Years Education, 43*(3), 238–248. https://doi.org/10.1080/03004279.2013.804852

Lado, R. (1961). *Language testing: The construction and use of foreign language tests. A Teacher's Book*. McGraw-Hill.

Lenneberg, E. (1967). *Biological foundations of language*. Wiley.

Lestariyana, R. P. D., & Widodo, H. P. (2018). Engaging young learners of English with digital stories: Learning to mean. *Indonesian Journal of Applied Linguistics, 8*(2), 489–495. https://doi.org/10.17509/ijal.v8i2.13314

Liang, M.-Y. (2019). Beyond elocution: Multimodal narrative discourse analysis of L2 storytelling. *ReCALL, 31*(1), 56–74. https://doi.org/10.1017/S0958344018000095

Lim, N. Z. L., Zakaria, A., & Aryadoust, V. (2022). A systematic review of digital storytelling in language learning in adolescents and adults. *Education and Information Technologies, 27*, 6125–6155. https://doi.org/10.1007/s10639-021-10861-0

Lindgren, R., & McDaniel, R. (2012). Transforming online learning through narrative and student agency. *Educational Technology and Society, 15*(4), 344–355.

Liu, C. C., Yang, C. Y., & Chao, P. Y. (2019). A longitudinal analysis of student participation in a digital collaborative story. *Educational Technology Research, 67*(4), 907–929. https://doi.org.remotexs.ntu.edu.sg/10.1007/s11423-019-09666-3

Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning, 68*(2), 321–391. https://doi.org/10.1111/lang.12286

Moradi, H., & Chen, H. (2019). Digital storytelling in language education. *Behavioural Sciences, 9*(12), 147. https://doi.org/10.3390/bs9120147

Neubauer, B. E., Witkop, C. T., & Varpio, L. (2019). How phenomenology can help us learn from the experiences of others. *Perspectives on Medical Education, 8*, 90–97. https://doi.org/10.1007/s40037-019- 0509-2

Nunally, C. (1982). Reliability of measurement. *Encyclopaedia of Educational Research, 4*, 15–16.

Oppermann, M. (2007). Writing in "that other space": Digital storytelling and the scholarship of teaching in American studies. *Teaching American Studies in The Twenty-First Century, 52*(3), 22–28. https://www.jstor.org/stable/41158317

Palacios, J. (2012). Traditional storytelling in the digital era. *Fourth World Journal, 11*(2), 41–56. https://search.informit.org/doi/abs/10.3316/informit.021672701588162

Perniss, P. (2018). Why we should study multimodal language. *Fronters in Psychology, 9*(1109), 1–5. https://doi.org/10.3389/fpsyg.2018.01109

Rahman, S. M. (2016). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "testing and assessment" research: A literature review. *Journal of Education and Learning, 6*(1). https://files.eric.ed.gov/fulltext/EJ1120221.pdf

Ravitch, S. M. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. Basic Books.

Reid, G., Snead, R. C., Pettiway, K. M., & Simoneaux, B. A. (2016). Multimodal communication in the university: Surveying faculty across disciplines. *Across the Discipline, 13*. http://wac.colostate.edu/atd/articles/reidetal2016.cfm

Rigney, A. (2010). When the monograph is no longer the medium: Historical narrative in the online age. *History and Theme, Theme Issue 49*, 100–117.

Rossiter, M., & Garcia, P. A. (2010). Digital storytelling: A new player on the narrative field. *New Directions for Adult and Continuing Education, 126*, 37–38. https://doi.org/10.1002/ace.370

Royal Literary Fund. (n.d.). *What is a literature review*. https://www.rlf.org.uk/resources/whatis-a-literature-review/

Sallee, M. W., & Flood, J. T. (2012). Using qualitative research to bridge research, policy, and practice. *Theory into Practice, 51*(2), 137–144. https://doi.org/10.1080/00405841.2012.662873

Schank, R. (1990). *Tell me a story: Narrative and intelligence*. North-western University Press.

Silverman, D. (2010). *Qualitative research*. Sage.

Suwardy, T., Pan, G., & Seow, P. S. (2013). Using digital storytelling to engage student learning. *Accounting Education, 22*(2), 109–124. https://ink.library.smu.edu.sg/soa_research/921/

Tuohy, D., Cooney, A., Dowling, M., Murphy, K., & Sixsmith, J. (2013). An overview of interpretive phenomenology as a research methodology. *Nurse Researcher, 20*(6), 17–20. https://pubmed.ncbi.nlm.nih.gov/23909107/

University of Houston. (n.d.). *The seven elements of digital storytelling*. https://digitalstorytelling.coe.uh.edu/page.cfm?id=27&cid=27&sublinkid=31

University of Wisconsin-Madison. (n.d.). *Nursing resources: Types of research within qualitative and quantitative*. https://researchguides.library.wisc.edu/c.php?g=861013&p=6170080

Wexler, L., Gubrium, A., Griffin, M., & DiFulvio, G. (2013). Promoting positive youth development and highlighting reasons for living in Northwest Alaska through digital storytelling. *Health Promotion Practice, 14*(4), 617–623. https://doi.org/10.1177/1524839912462390

Winston Salem State University. (n.d.). *Key elements of a research proposal: Quantitative design*. https://www.wssu.edu/about/offices-and-departments/office-of-sponsoredprograms/pre-award/_Files/documents/develop-quantitative.pdf

Xu, Y., Park, H., & Baek, Y. (2011). A new approach toward digital storytelling: An activity focus on writing self-efficacy in a virtual learning environment. *Educational Technology and Society, 14*(4), 181–191.

Yee, B. C., Abdullah, T., & Nawi, M. (2018). A reflective practice of using digital storytelling during teaching practicum. *Lecture Notes on Data Engineering and Communications Technologies, 5*, 708–715. http://dx.doi.org/10.1007/978-3-319-59427-9_73

Zull, J. (2002). *The art of changing the brain: Enriching the practice of exploring the biology of learning*. Stylus Publishing.

Publisher's Note

The Asia-Pacific Society for Computers in Education (APSCE) remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Research and Practice in Technology Enhanced Learning (RPTEL) **is an open-access journal and free of publication fee.**