# NEW EDUCATIONAL TECHNOLOGIES AND THEIR IMPACT ON STUDENTS' WELL-BEING AND INCLUSION PROCESS

EDITED BY: Maria Luisa Lorusso, Daniela Traficante, Mirta Vernice, Barbara Carretti and Daniela Maria Sarti

PUBLISHED IN: Frontiers in Psychology

frontiers Research Topics

frontiers



ISSN 1664-8714
ISBN 978-2-88974-065-9
DOI 10.3389/978-2-88974-065-9

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

# NEW EDUCATIONAL TECHNOLOGIES AND THEIR IMPACT ON STUDENTS' WELL-BEING AND INCLUSION PROCESS

Topic Editors:
**Maria Luisa Lorusso,** Eugenio Medea (IRCCS), Italy
**Daniela Traficante,** Catholic University of the Sacred Heart, Italy
**Mirta Vernice,** University of Urbino Carlo Bo, Italy
**Barbara Carretti,** University of Padua, Italy
**Daniela Maria Sarti,** IRCCS Carlo Besta Neurological Institute Foundation, Italy

**Citation:** Lorusso, M. L., Traficante, D., Vernice, M., Carretti, B., Sarti, D. M., eds. (2022). New Educational Technologies and Their Impact on Students' Well-being and Inclusion Process. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-065-9

# Table of Contents

frontiers in Psychology

EDITORIAL
published: 25 November 2021
doi: 10.3389/fpsyg.2021.753471

# Editorial: New Educational Technologies and Their Impact on Students' Well-Being and Inclusion Process

*Mirta Vernice*[1]*, *Barbara Carretti*[2], *Daniela Sarti*[3], *Daniela Traficante*[4] and *Maria Luisa Lorusso*[5]

[1] Department of Humanities, University of Urbino Carlo Bo, Urbino, Italy, [2] Department of General Psychology, University of Padua, Padua, Italy, [3] Fondazione Istituto Ricovero Cura a Carattere Scientifico (IRCCS) Istituto Neurologico Carlo Besta, Milan, Italy, [4] Department of Psychology, Catholic University of the Sacred Heart, Milan, Italy, [5] Eugenio Medea (Istituto Ricovero Cura a Carattere Scientifico), Bosisio Parini, Italy

**Keywords: innovative technology in education, inclusion, students well-being, technology and learning, teaching efficacy**

## Editorial on the Research Topic

## New Educational Technologies and Their Impact on Students' Well-Being and Inclusion Process

The Research Topic collects contributions that allow to rethink educational goals in the digital era from new angles of exploration. Moving from the assumption that introducing innovative technology paradigms in education requires unraveling aspects related to the impact of digital technologies on students' well-being, teaching efficacy and learning success, the Research Topic addressed these points by proposing novel tools and didactic methodologies to be implemented in the educational practices. These aspects appear even more relevant after the pandemic breakout that forced the educational systems of most countries to rely on digital education mainly.

The first point addressed in the Research Topic highlights the impact new technologies play on students' well-being. In this vein, Mascia et al. shed light on the relationship between dysfunctional use of smartphones and adolescent well-being, identifying two dimensions that can affect their quality of life (QoL), emotional intelligence and self-regulation. The authors observed a potential moderating effect of smartphone addiction on the relation between self-regulation and well-being and between emotional intelligence and well-being. Using a performance-based test and self-reports, Sarti et al. explored whether general well-being and school engagement of students with Specific Learning Disabilities (SLD) and Typical Development (TD) could be related to emotional and socio-cognitive functioning. The authors included a Smartphone addiction scale to examine its link with emotional and social functioning. Crucially, students with SLD represent a population vulnerable to problematic smartphone use, as they reported more internalizing and externalizing problems and more difficulties in reaching satisfying and supportive relationships. Importantly, digital well-being concerns the supposed benefits of digital engagement and some of its possible risks; in this view, the digital well-being of students may differ substantially from that of educational staff. The work by Panesi et al. involved the use of the SELFIE platform (a self-assessment tool) that offered a holistic view of how students, teachers, and school leaders perceive the digital *status quo* of their policies and practices, encompassing key aspects in education contexts like students' inclusion.

OPEN ACCESS

***Edited by:***
*Ehsan Namaziandost, University of Applied Science and Technology (UAST), Iran*

***Reviewed by:***
*Rui Xu, Henan University, China*

***Correspondence:**
*Mirta Vernice*
*mirta.vernice@uniurb.it*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *04 August 2021*
***Accepted:*** *29 October 2021*
***Published:*** *25 November 2021*

***Citation:***
*Vernice M, Carretti B, Sarti D, Traficante D and Lorusso ML (2021) Editorial: New Educational Technologies and Their Impact on Students' Well-Being and Inclusion Process. Front. Psychol. 12:753471. doi: 10.3389/fpsyg.2021.753471*

Second, other studies addressed the potential contribution of ICT to cognitive and neuropsychological assessment by presenting online tools developed to test online verbal comprehension (Caccia et al.), sentence comprehension (Vernice et al.), text reading comprehension (Capodieci et al.), and executive functions (Berg et al.). The need for such tools has been highlighted in the recent pandemic, where the use of online testing and training tools has become a necessity more than a choice. Clinicians and teachers have explored all the possible adjustments of traditional activities to make them fit online use, and a sort of huge natural experiment has taken place highlighting the advantages and disadvantages of the different delivery modalities. Some of the difficulties that children and students may face when reading and processing online information have been analyzed in the study by Caccia et al. in this study, online comprehension is at the same time the object and the means of investigation. On the other hand, the study by Vernice et al. shows how the transition from traditional testing to online testing allows for valuable, large-scale information to be collected that can provide not only data on individual skills but also in-depth analysis of the relationships existing among environmental factors and neuropsychological functioning. Finally, as shown by Capodieci et al.'s study (Capodieci et al.), the growing diffusion of technology in many fields of school life, along with the use of an increasing number of digital reading devices, offers new opportunities to improve traditional reading comprehension and learning skills, extending to more general cognitive abilities such as, for instance, inference generation. In this perspective, internet-based rehabilitation activities may constitute a valid alternative to more traditional, in-presence treatment for learning disorders. Berg et al.'s work (Berg et al.) complete this overview of applications and ICT tools, showing that neuropsychological assessment, in this case specifically of executive functions, in the form of games, can successfully be transferred to non-clinical settings and young, primary school children. This study highlights the importance of finding an ideal (and often hard to determine) level of difficulty that may keep the proposed digital activities challenging but still enjoyable, and proposing a sufficiently varied set of items and trials, but carefully avoiding the risk for the test to become too long and demanding.

A third point emphasized to what extent technology may promote a committed learning environment. In this regard, Feraco et al. explored the effect of the use of interactive teaching practices on academic performance in large classes where attendance is not compulsory. In particular, they focused their attention on the so-called Student Response Systems (SRS) which require the student to answer quizzes during university lessons. The use of quizzes has been integrated with extra-curricular activities that require an in-depth study of the course contents, such as laboratory experiences and writing reports. Both activities were successful in improving the students' final exam. However, these extra-curricular experiences had a positive effect not only on the students' academic performance but also on their motivation. This was not the case with quizzes. To better understand the relationship between technology, motivation, and learning outcomes, An et al. studied this issue in the context of language learning. They found that students with a higher level of motivation (in terms of self-efficacy) experienced greater involvement in activities that promote effective language learning, such as those equipped with technological tools, with better results on English learning tests. Overall these findings suggest that technological tools are not motivating in themselves, but that they can provide a more motivating context in the service of learning. Along this line, Ritella et al. highlighted the importance of considering new technologies as an innovative teaching tool to be used, however, in a theoretical framework. They analyzed students' perceptions and lasting memories of a course delivered in a mixed modality and studied the transfer of skills and knowledge over 10 years. The course was based on the model of constructive and collaborative participation (CCP) (Cucchiara et al., 2014) and included activities involving each student individually or in interaction with other students. The students reported vivid memories of the teaching methods and contents of the course, but of most interest, they stated that they use the soft skills acquired during the course in their current work activities. These findings reaffirm the importance of designing educational experiences that allow students to build knowledge through individual and collaborative activities that can take advantage of technological innovations.

The last issue referred to novel technologies in teaching, offering insights about teachers' cognitions of online educational paradigms. Chen et al.'s work (Chen et al.), through semi-structured interviews with teachers, provided evidence that using interactive spherical video-based virtual reality (ISV-VR) might be effective to improve descriptive composition writing in L1, in secondary school. The authors underlined that applying VR had an impact not only on students' skills and motivation but also on teachers' perspectives, which from "teacher-centered" became "student-centered." This contribution highlighted the need for teachers to increase their skills in integrating new technologies in teaching-learning processes. Rüth and Kaspar's work (Rüth and Kaspar) offers interesting clues about the use of commercial videogames in high-school courses. They proposed the use of a videogame on the evolution by natural selection in a 10th grade biology course and the use of a videogame on the First World War in a 12th grade course on history. Such activities increased motivation and contributed to share experiences not only on the subject contents but also on coding skills, promoting discussions on the way to improve the videogames. The authors provided evidence that videogames at school can be used as "objects-to-think-with." Gao and Zhang offered an interesting framework to integrate teachers' representations of the use of ICT in teaching English as a Foreign Language (EFL) in Chinese universities during the pandemic breakout. Moving from Koehler and Mishra (2005) model (TPACK) that integrates the perception of technology use within teachers' cognition (Borg, 2015), technological knowledge (TK) interacts with pedagogical knowledge (PK), and content knowledge (CK) in teachers' representations of on-line teaching EFL. Qualitative analysis of in-depth interviews with three EFL teachers showed that teachers are aware of the positive and negative aspects of the use of ICT in teaching English; they agree that the digital skills, acquired during the pandemic emergency, might be used to integrate traditional classroom teaching with online activities.

In conclusion, the contributions point out new research directions to inform educational practices and bridge the gap between technology innovation and educational methodologies, offering new perspectives of development for researchers and stakeholders.

## AUTHOR CONTRIBUTIONS

MV worked on the initial alignment of ideas and worked on subsequent editing. BC, DS, ML, and DT worked on specific sections. ML completed final editing. All authors contributed to the article and approved the submitted version.

## FUNDING

This publication was supported by the CARIPLO Foundation [Grant 2017-NAZ-0131, Tecnologie Innovative per il Benessere e l'Integrazione Scolastica (IBIS) (New technologies for education and their impact on students' well-being and inclusion)].

## REFERENCES

Borg, S. (2015). *Teacher Cognition and Language Education: Research and Practice*. London: Bloomsbury.

Cucchiara, S., Ligorio, M. B., and Fujita, N. (2014). "Understanding online discourse strategies for knowledge building through social network analysis," in *Innovative Methods and Technologies for Electronic Discourse Analysis*, eds H. Lim and F. Sudweeks (Hershey, PA: Information Science Reference), 42–62. doi: 10.4018/978-1-4666-4426-7.ch003

Koehler, M., and Mishra, P. (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *J. Educ. Comput. Res.* 32, 131–152. doi: 10.2190/0ew7-01wb-bkhl-qdyv

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 01 November 2019
doi: 10.3389/fpsyg.2019.02433

# ORCA.IT: A New Web-Based Tool for Assessing Online Reading, Search and Comprehension Abilities in Students Reveals Effects of Gender, School Type and Reading Ability

*Martina Caccia[1], Marisa Giorgetti[1], Alessio Toraldo[2,3], Massimo Molteni[1], Daniela Sarti[4], Mirta Vernice[5] and Maria Luisa Lorusso[1]**

[1] Scientific Institute IRCCS Eugenio Medea, Bosisio Parini, Italy, [2] Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy, [3] Milan Center for Neuroscience (NeuroMI), Milan, Italy, [4] Carlo Besta Neurological Institute (IRCCS), Milan, Italy, [5] Department of Psychology, University of Milano-Bicocca, Milan, Italy

OPEN ACCESS

**Edited by:**
*Meryem Yilmaz Soylu, University of Nebraska–Lincoln, United States*

**Reviewed by:**
*Angeliki Mouzaki, University of Crete, Greece*
*Elena E. Forzani, Boston University, United States*

***Correspondence:**
*Maria Luisa Lorusso*
*marialuisa.lorusso@bp.lnf.it*

**Specialty section:**
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

**Received:** *30 June 2019*
**Accepted:** *14 October 2019*
**Published:** *01 November 2019*

**Citation:**
*Caccia M, Giorgetti M, Toraldo A, Molteni M, Sarti D, Vernice M and Lorusso ML (2019) ORCA.IT: A New Web-Based Tool for Assessing Online Reading, Search and Comprehension Abilities in Students Reveals Effects of Gender, School Type and Reading Ability. Front. Psychol. 10:2433. doi: 10.3389/fpsyg.2019.02433*

ORCA.IT, a new online test of online research and comprehension was developed for the Italian population. A group of 183 students attending various types of upper secondary schools in Northern Italy were tested with the new tool and underwent further cognitive and neuropsychological assessment. The different school types involved in the study are representative of the school population in the Italian system, but can also be easily compared with the educational systems of other countries. The new test turned out to have good psychometric properties after accurate item construction and final selection. In particular, Version 1 showed better characteristics than Version 2. Subsequently, comparison with one-way ANOVAs were performed to test whether differences exist between different school types, between groups with and without reading difficulties, and between males and females. Such differences are sometimes reported in the literature, but many remain controversial. Further, Pearson's bivariate correlations were calculated to analyze associations between scores on the ORCA.IT and cognitive/neuropsychological variables. Finally, a stepwise regression analysis was performed on aggregated scores to identify the predictors of performance on each of the two versions. The test, especially in the most complete version (Version 1), appears to accurately and reliably capture students' web searching abilities and online reading comprehension. The tool could highlight differences in online search and comprehension ability between students with and without reading difficulties, not penalizing overall performance but allowing very specific weaknesses to be pointed out. Further, it seems to be able to capture differences due to both educational pathways (different school types) and social attitudes (differences between males and females). Most interestingly, it shows to be clearly resting on specific cognitive and neuropsychological abilities, including language, memory, and attentional skills, which explain a large portion of the total variance. Offline text reading comprehension is a crucial predictor of online reading

performance, while decoding ability is not. Prior knowledge also influences the results, as expected. The new tool turns out to be rather independent of previous Internet experience and to measure more cognitively grounded processes related to information gathering, processing, and communicating.

**Keywords: online reading comprehension, assessment, gender effects, education effects, reading difficulties**

# INTRODUCTION

## ICT, Internet and Literacy

In the last years, there has been an increasing interest in how young people use the Internet and other new technologies in their everyday lives and how this use may enhance informal and formal learning opportunities (Becta, 2008). Indeed, previous research has shown that there is great variability in the ways they access and use Internet sources (Facer and Furlong, 2001; Livingstone and Helsper, 2007; Hargittai and Hinnant, 2008). In light of this, there is a need to better understand the complex factors determining this heterogeneity and what this may mean for the educational system (Selwyn, 2009). This constitutes one of the challenges for researchers, policy makers and teachers in order to use new technologies more effectively for formal education and develop more targeted initiatives that better support young people in their uses of the Internet and other information and communication technologies (ICTs) (Leu et al., 2011).

Nowadays, there is a tendency to promote ICTs adoption in many fields, including schools. However, many teachers make a poor and rare use of technology during their lessons and, usually, they do not fulfill the expected results (Korte and Hüsing, 2006; Shewbridge et al., 2006; Law et al., 2008; Leu et al., 2011; Davies and West, 2014; Fraillon et al., 2014). Leu et al. (2011) supported the idea that ICTs for educational purposes should take into account that the notion of literacy has now changed: it is now deictic; its nature and meaning continuously changes. Moreover, effective use of online information requires additional online reading comprehension practices. Public policy, assessment, and education should be able to prepare students for such challenges.

The arrival and spread of electronic resources and digital libraries have changed and extended the notion of literacy beyond its original application to the medium of writing. Nowadays, many researchers converged on the concept of "multiliteracies," to define both the social diversity of contemporary forms of literacy, and new communications media and communicative competence (Cope and Kalantzis, 2000). "Digital literacy" perpetually changes because new technologies for information and communication constantly appear online and require new social practices of literacy.

The United Kingdom media regulator (Office of Communications [Ofcom], 2004) defined "Digital literacy" as the "ability to *access*, *understand,* and *create communications* in a variety of contexts." *Access* refers to the skills needed to locate media content, using the available technologies (Buckingham, 2007). *Understand* takes into account the ability to decode or interpret media and it also involves knowledge of production processes and an ability to evaluate the specific media, for example, in terms of the accuracy or reliability of the web sources (Buckingham, 2007). Lastly, *create* consists of the ability to use the media to produce and communicate one's own messages (Buckingham, 2007).

## Offline and Online Reading

As shown above, reading online implies high levels of critical skills; indeed Fabos (2004) stressed the importance of promoting more critical analysis of online content. Actually, online reading comprehension is a process which requires to analyze many different sources of online information, using several recursive reading practices (Coiro, 2003; Henry, 2006), following what is now known as the LESC model (Leu et al., 2013): (i) reading to locate information (L); (ii) reading to evaluate information (E); (iii) reading to synthesize information (S); and (iv) reading and writing to communicate information (C). Specifically, online reading requires both new online and traditional offline reading comprehension skills (Leu et al., 2011).

Reading and searching online information usually implies that a question has been formulated. Taboada and Guthrie (2006) identified differences, within traditional texts, between reading that was, or was not, initiated by a question. Moreover, online reading is a multi-componential process and requires, for example, the generation of effective keyword search strategies (Bilal, 2000; Kuiper et al., 2008), inference as to which link may be most useful within a set of search engine results (Henry, 2006), and efficient evaluation of relevant information within websites (McDonald and Stevenson, 1996; Rouet, 2006). Successful online reading requires also to tell reliable information from fake news (Sanchez et al., 2006; Graesser et al., 2007). However, such practices present challenges that are quite different from those regarding traditional print and media sources, because the content of online information is more assorted and commercially biased than that of print sources (Fabos, 2008; Leu et al., 2011). Online Reading Comprehension also requires the ability to synthesize information from multiple sources (Jenkins, 2006) and communicate and discuss it via the Internet (Britt and Gabrys, 2001; Kiili et al., 2012).

Online and offline reading skills are organized in complex ways (Leu et al., 2015) and they share some similarities (Coiro, 2011; Hahnel et al., 2016). During reading comprehension, prior knowledge plays an important role in building a logical representation of the text (Kintsch, 1998). An additional factor that may play a role, especially when one reads about controversial issues (as it often happens during online searching), is a reader's prior beliefs on an issue. This may influence text interpretation (Nickerson, 1998) or website evaluation (van Strien et al., 2016).

Kiili et al. (2018), in their study on the students' abilities to critically evaluate online sources, found that offline reading skills were necessary but not sufficient for a successful evaluation of online resources; they hypothesized that offline reading does not involve the same amount of critical evaluation skills that is required in online reading. Moreover, to the authors, poorer offline readers seem to show more difficulties in online evaluation, suggesting that offline reading skills are necessary for a successful online critical evaluation.

However, other studies found a lack of a strong relationship between offline reading practices and online reading evaluation. It appears that online and offline reading comprehension are different processes (Coiro and Dobler, 2007; Afflerbach and Cho, 2010) actually, a study by Bråten et al. (2009) showed that online critical evaluation skills appear to be separate and independent of offline reading skills.

## Factors Underlying Individual Differences in Online Reading

Among the context-dependent and experiential factors that may influence online reading abilities, special attention has been devoted to familiarity with the topics. As mentioned above, prior knowledge of the topic plays an important role in the comprehension of the different types of texts (Cromley et al., 2010; Tarchi, 2010) and hypertexts (Amadieu et al., 2009). However, Coiro (2011) found that even though prior topic knowledge played an important role in online research and comprehension performance of students with low online reading skills, it did not influence the performance of students with high online reading skills. Moreover, Kanniainen et al. (2019) found that the relationship between prior knowledge and Online Reading Comprehension was not significant.

Indeed, students may feel inadequate in assessing sites when they are unfamiliar with their topics. They largely fail to apply reliable criteria; rather, they emphasize (e.g.) speedy access to information and appealing visual design (Buckingham, 2010). In their research, Eynon and Malmberg (2011) investigated five uses of the Internet: communicating, information seeking, entertainment, participating, and creating, among young people. They found four different Internet user profiles: the peripherals, the normatives, the all-rounders, and the active participants. The first group was the least frequent users of the Internet, tending to do less of all Internet activities than the other groups and showing a lack of skills to use it. The normatives showed average uses of three types of Internet activity – communicating, entertaining, and information seeking – and were engaged less in more proactive uses of the Internet such as creating and sharing actively contents. The all-rounders used the Internet for all five types of Internet use more frequently than the average. Active participants are those who use the Internet most frequently, for all five activities, and tend to engage in online participatory behaviors (like contributing to a blog or wiki page) more frequently.

Familiarity with specific topics, as well as with specific approaches to information collection and analysis may further depend on the students' educational paths. In the Italian system there are three main types of secondary schools: Lyceum, Technical Schools, and Vocational Schools. These types of schools differ with respect to specific curricula, as well as to a more theory-oriented or more practice-oriented approach, and these differences may entail different habits with respect to information search (how and where to look for information) and processing (how to handle it). To our knowledge, no previous study focused specifically on the influence by school type on students' Internet usage. In their study on Internet inequality, Zhao et al. (2010) found that students with available digital devices at home tend to have the strongest perception of Internet skill. Availability of social support from school has a greater effects than that from home on Internet self efficacy (ISE), which describes learners' confidence in their general ability to operate Internet functions or applications in Internet-based learning conditions (Tsai et al., 2011). A higher ISE has been shown to be associated with better information search strategies, better learning skills, and better learning performance, while more controversial results have emerged on the relationship between ISE and actual Internet usage or navigational paths (Tsai et al., 2011). Zhao et al. (2010) further found that school Internet accessibility seems not to be significantly related to ISE. According to the authors, one reason might be that in most high schools of China, the Internet-related resources, such as computers with access to the Internet, are inadequate. As to the Italian context, in spite of the strong pressure from the Ministry of Education to update and upgrade ICT resources in the schools (e.g., a wider use of MIW – Multimedia Interactive Whiteboards, and Classrooms 2.0, as fostered by the Piano Nazionale Scuola Digitale – National Plan for Digital Education in response to OECD requirements, Miur, 2015), equipments are very often old or poorly maintained and teachers are often not familiar with the new technologies (Gremigni, 2019); moreover, the students' abilities in searching information in the Internet and evaluating its reliability are still lower than the OECD average (Miur, 2015). Overall, it would thus be difficult to predict whether different school types differ with respect to the opportunities they offer to their students to reach a good mastery of online search and comprehension skills. Indeed, more theory-oriented schools probably require frequent comparison of different sources (but not necessarily through the Internet); at the same time, more practically oriented schools may use digital tools also for technical purposes (e.g., graphics software), which could improve students' familiarity and confidence with ICT tools.

Following the previous literature, other factors may concur to reading comprehension, specifically, Snowling (2013) suggested that students with low verbal and non-verbal reasoning skills are more likely to have comprehension difficulties. Non-verbal reasoning has been shown to have direct and indirect effects on reading comprehension (Swart et al., 2017) and, in line with this, Kanniainen et al. (2019) found that non-verbal reasoning contributed independently to the variance of Online Reading Comprehension performance.

Reading abilities are clearly good candidates to play a central role in online reading, as suggested by the relationships between offline and online reading skills. However, very little is known about the behaviors of people with dyslexia in web usage. This is probably due to the general focus on the consequences that

decoding difficulties have on school performance in younger readers, and on the belief that decoding ability only plays a major role in beginning reading acquisition, and is subsequently replaced by comprehension skills. Indeed, it is now evident that the distinction between decoding and comprehension is less clear-cut than previously thought (e.g., Gough and Tunmer, 1986), the two kinds of difficulty interacting with each other and being frequently present in the same individual, possibly at different time points during development (Bishop and Snowling, 2004; García and Cain, 2014). This may clearly apply to online reading too. Very little is known, though, about the behaviors of people with dyslexia in web usage. McCarthy and Swierenga (2010), in their review, argued that dyslexic-friendly practices may help overcome difficulties faced by all Internet users. In 2004, the United Kingdom's Digital Rights Commission conducted a task-oriented examination of 100 websites and a group of people with dyslexia took part in the study. Each user completed two tasks on ten different websites. Dyslexic users experienced a 17% failure rate (which was lower than the rate experienced by – for example – blind and partially sighted participants). The main issues experienced by dyslexic users were: confusing page layout, unclear navigation, poor color selection, difficulties in decoding graphics and complicated language.

Kurniawan and Conroy (2007) tested reading comprehension speed and accuracy during Internet gathering information of dyslexic and non-dyslexic students and they found that participants with dyslexia made increasingly frequent mistakes as reading material became more complex but allowing users to select their ideal color scheme increased reading speed for both groups. A general idea emerging from the literature is that dyslexia is highly variable; actually there is no "typical" dyslexic Internet user. Pollak (2001) interviewed college students with dyslexia and he found that they underline the potential strengths of multimodal documents with respect to unimodal ones.

It is well established that many students with dyslexia also experience language or visual-spatial difficulties (Giovagnoli et al., 2016; Snowling and Melby-Lervåg, 2016), and that their visual-verbal integration skills may also be poor (Hahn et al., 2014). This means that complex text comprehension may be problematic, even when the coding difficulties have been largely overcome or compensated, and this may be true not only for written text, but also for complex oral explanations that accompany online videos, or graph interpretation. Such difficulties may be more evident in high school, when texts are more often supported by non-textual materials, and may be especially relevant for the multi-media and multi-modal information that is found in the Internet.

A final interesting issue in literacy research is gender-related differences. Several studies showed an advantage for girls in reading fluency and reading comprehension (Logan and Johnston, 2009, 2010; Torppa et al., 2018) and similar patterns have also been observed in other studies on Online Reading Comprehension (Forzani, 2016; Salmerón et al., 2018; Kanniainen et al., 2019). McKenna et al. (2012) reported that middle school girls have more positive reading attitudes toward reading print texts for recreational as well as academic purposes, while their attitudes toward reading digital texts are better for academic purposes, but not for recreational purposes. A similar finding is reported by Lupo et al. (2017) with high school students. Meelissen and Drent (2008) propose that this may be due to female students showing less positive attitudes toward using computers as compared to male students. By contrast, no significant differences emerged between females and males in a Korean middle school sample (Jang and Ryoo, 2018) in academic-related Internet activities. The authors propose that this may result from the strong achievement-driven characteristics of Asian secondary schools.

## Effects of Web Usage on Cognitive Functions

The ready availability of information on the Internet may decrease the need to store and recall data. Sparrow et al. (2011) suggested that people may be becoming better at remembering where information is located than at recalling it; this has been defined as the "Google effect." It has been suggested that individuals born after 1993 (the so-called "Google generation") may show weaker working memory and be less confident about their answers as compared with older individuals, even if they retrieve information and make responses more rapidly (Nicholas et al., 2011). Moreover, Dong and Potenza (2015) showed in their study that even if Internet-based searching may have facilitated the information-acquisition process, this process may have been performed more rapidly and be more likely associated with difficulties in recollection. In addition, people appeared less confident in recalling information learned through Internet searching.

Online searching seems to have changed our attentional abilities. In their study, Ophir et al. (2009) explored the impact of the sustained media multi-tasking on cognitive skills. They found that frequent and extensive (called "heavy") media multi-tasking performed worse in task-switching tests than the normal multi-media users. It was suggested that the compromised ability in heavy media multi-tasking people was due to their increased proneness to distraction from irrelevant environmental inputs. However, literature on multi-media and Internet usage have produced conflicting findings (Firth et al., 2019). Nevertheless, on the whole, the literature seems to agree on the fact that those who engage in frequent and extensive media multi-tasking in their daily lives perform worse in various cognitive tasks than those who do not, particularly for sustained attention (Uncapher and Wagner, 2018). Moreover, a longitudinal study of media multi-tasking in young people has found that frequent multi-tasking behaviors predict the development of attentional deficits specifically in early adolescents but not in older ones (Baumgartner et al., 2017).

## The Present Study

The present study is outlined using an online research and comprehension framework (Leu et al., 2011), which focused on the four crucial component skills mentioned above (LESC, i.e., Locate, Evaluate, Synthesize, and Communicate) and on different types of media (texts, images, videos, and graphs). Specifically, students' online reading abilities were measured with "ORCA.IT,"

an Italian adaptation of the online research and comprehension assessment (ORCA) originally developed in the ORCA Project (Coiro and Kennedy, 2011; Leu et al., 2014). Several ORCA tools had been developed in this Project (Leu et al., 2014). ORCA.IT was inspired by ORCA-Multiple Choice. This was a performance-based assessment within a more restricted and limited simulation of the Internet. In this format, students were guided through a research task by a student avatar that contacted them through a social network, and another student avatar that contacted them through instant messaging, all within the ORCA space. Topics were: Energy drinks and heart health, Videogames and effects on eyes, Use of decorative lenses and effects on eyes, Safe volume levels for Ipods. Students used fully functional tools (a social network, text (chat), email, wikis, a search engine, and a notepad) to conduct their research. Italian ORCA was designed to emulate a natural online research process even if it did not use a fully functioning simulation of the Internet (such as social networks, wikis, avatars, and the possibility to write texts). Students were asked to use Internet tools (a search engine, email etc.) to conduct their research on a specific given topic within a simulation of the Internet. All questions required to choose among pre-constructed responses (either verbal responses or images). The topics included in the Italian version (Electromagnetic waves and health and Music and brain) were different from those addressed in the original version and had been chosen so as to be edge-cutting, likely interesting to the students and not part of any standard school science program and/or curriculum. The skill areas that were evaluated (LESC) did not appear in a strictly linear sequence in any of the two topics. Also, differently from the original version and due to specific hypotheses about the impact of different communication modalities on performance (especially for students with special needs and different reading abilities), the Italian version distinguished between items that belong to four different typologies: verbal texts, graphs, images, and videos. Separate scores could be calculated for each of these areas, and a profile of each student could be visualized at the end of the test, showing the areas of relative strength and weakness, along with graphs depicting the LESC/typology profile for the single students as compared to her/his classmates (class averages).

ORCA.IT, in a similar way to the original ORCA project, assesses and considers the student's pre-existing knowledge of each topic. In the Italian version only, however, questions on prior knowledge included also a whole section devoted to previous experience with ICT tools and habitual use of multimedia technology. Throughout the assessment of both prior knowledge and the two topics, the multiple choice format allowed all the scores to be calculated automatically through implemented algorithms. Thus, differently from the original test, the Italian test does not require any additional judgment or scoring by the teachers or examiners.

Finally, the Italian device includes an original feature that is meant to support reading by students with reading difficulties: Text To Speech with a natural Italian female voice, that can be activated by the student for any part of the texts to be read (instructions, questions, answers, embedded texts, and parts of the graphs).

Summing up, our research questions were:

1. What are the general psychometric characteristics of the new test? In other words, is the test able to capture the students' abilities in a valid and reliable manner?
2. Are online reading abilities related to offline reading skills?
3. What are the predictors and components of online reading abilities (digital experience, prior knowledge, offline reading comprehension, STM, WM, and executive skills)?
4. How do students' online reading abilities differ by school type, gender, and reading (dis)ability status?

Based on previous studies and specific hypotheses (concerning the nature of the teaching approaches, which is more theory-focused in lyceum studies, more practice-based in vocational schools, and with intermediate characteristics in technical schools), lyceum students were expected to score generally higher than technical school students, who in turn were expected to score higher than vocational school students. However, non-textual information (especially interpretation of graphs and still images: indeed, videos require good oral language comprehension abilities) was expected to show smaller differences between school types as compared to text-based information; students with reading difficulties were expected to score lower than students without reading difficulties – although, in this case too, differences could be expected to emerge for textual information (and, possibly, video) only; males were expected to score higher than females, as found in previous research; this advantage was expected to concern non-textual more than textual information. Finally, statistical analyses were expected to reveal contributions from offline text comprehension abilities, prior knowledge, decoding abilities, attentional, and executive functions as expressed by visual search/selection tasks and questionnaires on attention and concentration abilities.

# MATERIALS AND METHODS

## Participants

Participants were 183 students (53% boys and 47% girls) all attending the upper secondary school in four different provinces of Lombardy (northern Italy); schools were selected from diverse regions of Lombardy so as to cover a wide range of SES profiles. Students were recruited from three different school types/levels: Scientific Lyceum (35%), Technical School (43%), and Vocational School (22%). These curricula differed for gender distribution ($\chi^2 = 14.18$; $p < 0.001$), with girls preferring the Technical Schools (51%), and boys preferring Scientific Lyceum (48%). Age ranged from 14 to 17 years ($M = 15.84$, $SD = 0.72$). Inclusion criteria were: adequate socio-educational conditions and absence of neurosensory deficits or cognitive impairment (IQ $\geq$ 85 as assessed by either the Cattell Culture-fair test, Cattell and Cattell, 1981 or the Raven Standard Progressive Matrices, SPM, Raven, 2003). Fourteen students had a formal diagnosis of Specific Learning Disorders formulated by experienced clinicians, based on standard diagnostic criteria (ICD-10). Seven students were classified as having reading difficulties based on their actual performance on the Reading tests described below. Criteria

for this classification were 2 SDs below the mean on at least one parameter (speed or accuracy in either word on non-word reading), or 1.6 SDs below the mean in at least two parameters. The sample also included 12 bilingual students, who had complete mastery of the Italian language. Bilingual students with insufficient mastery of Italian were excluded.

After receiving the school-manager's approval to carry out the research, the caregivers and the students were informed on the aim and procedure of the study. Parents provided a written consent for their children's participation in the study and students gave informed written consent to the study, according to the General Data Protection Regulation (GDPR 2016/79, 25/05/2018). Students completed the questionnaires and the tests in two group sessions and their decoding ability was assessed in one individual session. The present study was approved by the Scientific and Ethics Committee of the Department of Psychology of the Catholic University of Milan, in accordance with the Helsinki Declaration.

## Standardized Tests

All participants were administered with the following standardized tests (individually or collectively), which are the commonly used tests for assessment of Specific Reading Disorders in Italy.

### Reading

Single Word/Non-word Reading: "Batteria per la Valutazione della Dislessia e della Disortografia Evolutiva, DDE-2" (Battery for the assessment of Developmental Reading and Spelling Disorders), by Sartori et al. (2007). This test assesses speed (in syllables per second) and accuracy in reading word lists (four lists of 24 words each) and non-word lists (three lists of 16 non-words each) and was standardized on high school students (Arina et al., 2013). Concurrent validity for the DDE-2 was assessed through correlations between word and non-word lists from this and another widely used test, on a sample of primary and middle school students: correlations between word lists is 0.96, between non-word lists it is 0.79. Reliability was assessed through test-retest procedures (0.77 for reading speed and 0.56 for errors) and through correlations among subtests (average correlation = 0.79).

### Written Spelling

Spelling accuracy was assessed with a write-to-dictation task. A short text was read aloud by one examiner, in a clear and neutral voice, without emphasizing the source of the spelling difficulty and without explanations about words or expressions (Cornoldi et al., 2017). Accuracy is expressed by the number of errors made by the student. Substantial errors like omissions, inversions and substitutions are assigned one point (maximum 1-point per word). Standard scores are then calculated based on age norms. Correlation between word and non-word written spelling is 0.685.

### Comprehension

This was evaluated using Italian texts appropriate for the student's age (Cornoldi et al., 2010). The task requires silent reading followed by answering ten multiple-choice questions. The ability to extract the exact meaning from the text and to examine the information contained in a sentence is assessed. Internal consistency of the test for the second class is adequate (Cronbach's alpha = 0.84).

### Memory

Verbal short-term/working memory was measured by means of the digit span task from the Wechsler batteries (WISC–IV Wechsler, 2003). The experimenter reads aloud lists of single digits, and the participant is asked to recall them immediately after the end of the examiner's reading, either in the same order (forward) or in reverse order (backward). One point is assigned for each sequence of digits correctly recalled; the number of digits in each sequence increases (by one digit) if the participant has correctly recalled at least one of two sequences of a given length. Administration is discontinued when both items from a given pair are failed. The sum of points for each subtest represents the total score for that subtest. Then, age-corrected weighted scores are calculated according to age norms. Internal consistency of the scales is satisfactory (for age 15, reported reliability is 0.87; for age 16, reliability is 0.85).

### Attention

(a) The Brown Adolescent attentional disorder (ADD) Scale (Brown, 1996) is a self-assessment tool whose overall score is an indicator of the likelihood that the individual has some ADD; ADD probability is stratified in three categories: (i) possible but unlikely (overall score $<40$), (ii) likely but uncertain (40–54); (iii) very likely ($>54$). The Brown ADD subscales target subclinical impairments of executive functioning that impact academic, social, emotional and behavioral functioning. The adolescent version features five clusters frequently associated with ADD:

- Organizing, Prioritizing, and Activating to Work;
- Focusing, Sustaining, and Shifting Attention to Tasks;
- Regulating Alertness, Sustaining Effort, and Processing Speed;
- Managing Frustration and Modulating Emotions;
- Utilizing Working Memory and Accessing Recall.

As to reliability, the item-total correlation in the non-clinical Italian adolescent sample (average of the five clusters) equals 0.86.

(b) The d2-R Test is a neuropsychological measure of selective and sustained attention and visual scanning speed (D2-R; Brickenkamp et al., 2010). It is a paper-and-pencil test on which the participant is asked to scan some lines of letters and cross out all occurrences of the "d" letter while ignoring all other letters. D2-R provides a variety of measures, including Processing Speed, Accuracy of visual scanning, Coordination of speed and accuracy. Reliability estimates for the target age group are very high (Correct answers = 0.89; attention Performance = 0.90; Error percent = 0.91).

### Language

(a) The Italian version of the Peabody Pictures Vocabulary Scale (Dunn and Dunn, 1997) was used to assess receptive lexicon. Additionally, (b) a newly developed online sentence comprehension test developed by one of the authors

(Vernice et al., 2019) was proposed to assess receptive syntax. The test is based on 20 multiple choice trials. Each trial involves a target sentence of varying syntactic complexity that has to be read silently. The four sentences from which the response has to be chosen, include: a paraphrase of the target sentence (correct choice), a sentence contradicting the meaning of the target sentence, a sentence compatible with the target sentence, but not equivalent to it, and a distractor with different content. The score is the total number of equivalent sentences detected. The sentence comprehension test was presented using Google forms. No reliability data are available at present.

## Materials

An online platform has been developed to test students' ability to gather, comprehend, evaluate, synthesize, and report on information, to conduct research in order to answer questions or solve problems through new media forms and Internet.

The online reading comprehension test was delivered through a specially designed web-app. The web-app was developed using Ruby on Rails + Vue.js, one of the JavaScript frameworks for web applications, used for the realization of graphical interfaces. The interface is designed for use on tablets and desktops.

The graphics is friendly and simple (see **Figure 1**). The computer screen is divided into two parts: the left side of the screen contains the multiple-choice questions, while the right part is the Internet simulation space.

The device includes *Text To Speech* with a natural Italian female voice. It was selected for the quality of the Italian voice[1]. The button off/on was integrated in the form, and by clicking on it, it is possible to listen to any text element on the page.

The device has been designed according to the theoretical framework described in the introduction which identifies four crucial component skills: (1) locating information, (2) evaluating information critically, (3) synthesizing information, and (4) communicating information. These component skills have been assessed through questions addressing two different topics: (a) Music, Brain, and Neurosciences; (b) Electromagnetic Pollution and Health. Such topics were considered to be interesting for adolescents and not straightly linked to school content.

The students are required to plan the correct steps and select the best elements within a fictional context where they are requested to build a report on a given topic exploring online resources (news web pages with different credibility and reliability) and to distinguish relevant online resources from irrelevant ones or from potential fake news. Component skills (LESC) were tested with 19 multiple-choice (four options) questions for topic (a) and topic (b).

The answer options included one correct option, two incorrect options, and a partially correct option. Two points were given for each correct choice, one point for a partially correct choice and zero points were given for incorrect options. See **Figure 2** for an example. A partially correct option is an answer that captures some, but not all, of the relevant aspects or elements; alternatively it may be a relatively good option, but clearly less appropriate than the full-score option (accordance between different judges had been assessed through blinded assignment of scores and subsequent comparison, with an inter-rater agreement of about 0.87). As an exception, questions where only one correct answer was possible did not envisage partially correct (1-point) options, and thus included three incorrect (0-point) and one correct option (2-points).

Questions were grouped around the four LESC components described for online reading comprehension, i.e., (1) Locating information, (2) Evaluating information critically, (3) Synthesizing information, and (4) Communicating information. This allowed subscale total scores to be automatically calculated for each component.

Further, all questions were classified according to the type of materials used to convey information, distinguishing among Texts, Images, Videos, and Graphs. Also for these four categories it was thus possible to automatically calculate total subscale scores. Even if watching a video or analyzing a graph certainly activate multiple processes that need to be integrated following complex pathways, we decided to classify also multi-media (or multi-modal) and non-verbal items within the LESC model. As a result, all questions concurred in producing the scores for both the LESC and the "type of media" subdivision.

Prior to the proper ORCA.IT test, three subtests for prior knowledge (one concerning Internet use and two concerning the topics of the ORCA.IT test) were delivered as parts of the total assessment session (and part of the software). The three subtests are described below.

## Digital Skills

Familiarity and competence/skills in different forms of digital literacy was subsequently assessed. The test covered three main dimensions of digital literacy: preference and frequency of use of digital instruments, operational skills (needed to operate computers and network hardware and software), and self-evaluation skills. The items from the three sections are inspired by DigComp 2.0 (The Digital Competence Framework for Citizens, Vuorikari et al., 2016) that identifies the key components of digital competence. The students responded using a 5-point Likert scale (from 0 = "not at all/never" to 5 = "very much/very often") for all items.

## Prior Knowledge

Then, students were required to answer questions assessing Prior Knowledge, with respect to the two topics: (a) Music, Brain, and Neurosciences; (b) Electromagnetic Pollution and Health. The first question investigated the level of familiarity with the research topic. The following questions concerned specific information on the subject. For each research topic, there were five further multiple-choice questions. A total score was automatically calculated for topic (a) and for topic (b).

## ORCA.IT

Two equivalent versions of the ORCA.IT test were designed, intended as two parallel versions of the test for repeated testing. The general structure of the tests was the same for both versions, but the specific questions varied. More specifically, the same questions that on one version were formulated using topic (a),

[1] https://responsivevoice.org/

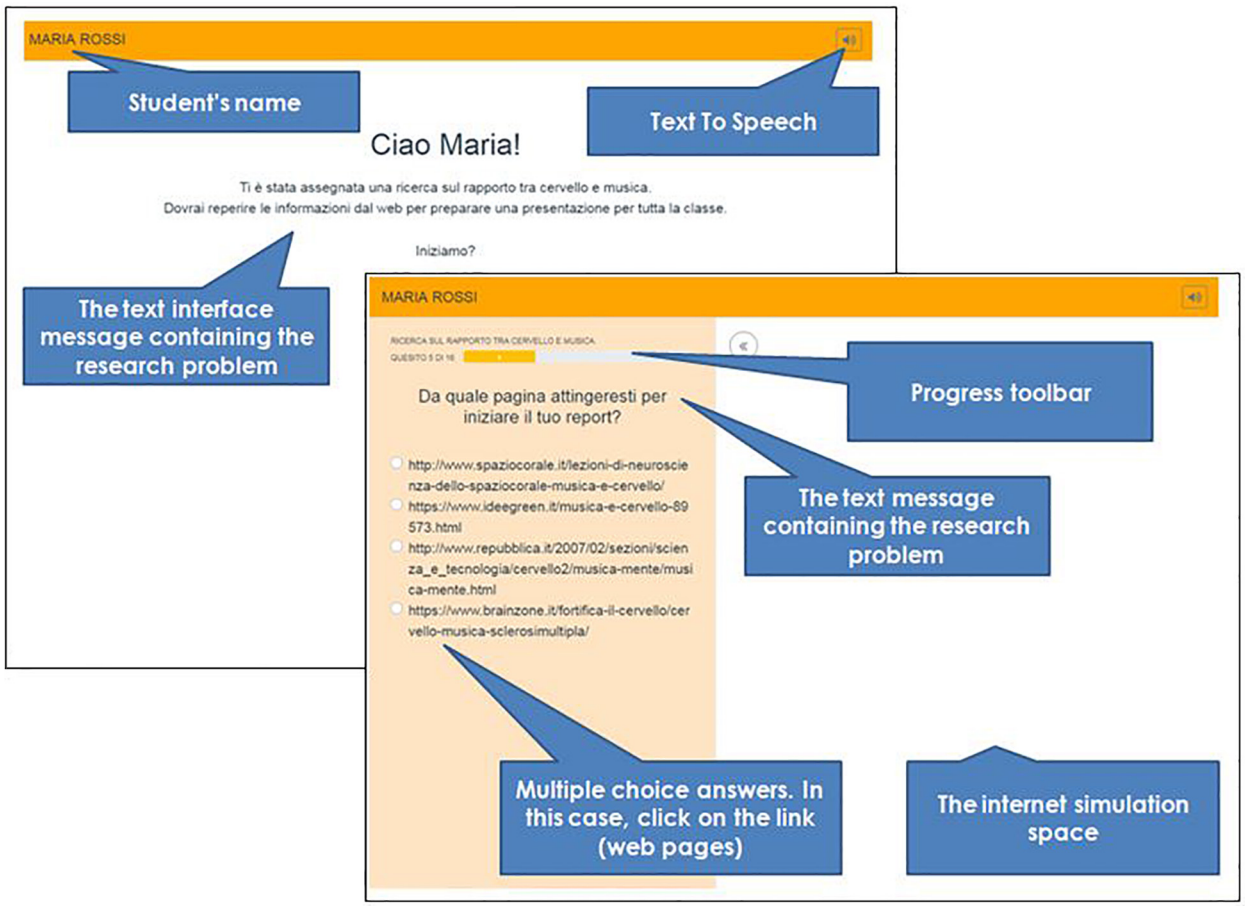


**FIGURE 1 |** Structure of the ORCA.IT online test.

were reformulated for the other version based on topic (b). The number of questions for each component of the model (Locate, Evaluate, Synthesize, and Communicate online information) and text-image-graph-video structure was also maintained equal for the two versions. Digital skills and Prior Knowledge questions are exactly the same in the two versions. For each version there were 38 items: 19 items on topic (a) and 19 items on topic (b).

The fourth LESC component skill – Communicating information – assessing ability to address a specific audience (teacher or schoolmates), in each version of ORCA.IT, was assessed by introducing the request to either prepare a report and write an email, or to prepare a slide presentation to show the results. These tasks were always accomplished by choosing among a series of four different formats; each of the questions tapped different aspects of communication skills and strategies. Both the "email communication" mode and the "report in Power Point" mode is available in Version 1 and Version 2, albeit on different topics.

The sequential composition of the two versions is described below:

**Version 1** – Digital skills; Prior Knowledge (Music and Brain); questions based on Topic (a) (first part); Prior Knowledge (Electromagnetic Pollution); questions based on Topic (b) (second part).

**Version 2** – Digital skills; Prior Knowledge (Music and Brain); questions based on Topic (a) (second part); Prior Knowledge (Electromagnetic Pollution); questions based on Topic (b) (first part).

## Procedure

The test was administered in the months of April and May 2019. On two separate days, students were administered all tests described above (reading, writing, text comprehension, memory, and cognitive tests). At least 2 months elapsed between the Standardized Tests and online reading comprehension assessment (ORCA.IT). The administration of ORCA.IT was done in the computer room and each student operated on computer (laptop or desktop computers). Completion of the online test required about 45 min. Cognitive and neuropsychological testing required about 1 h for collective testing and 20 min for individual testing. All testing was performed by licensed psychologists.

In a selection of schools, students were required to undergo testing with both Version 1 and Version 2. A total of 44 students

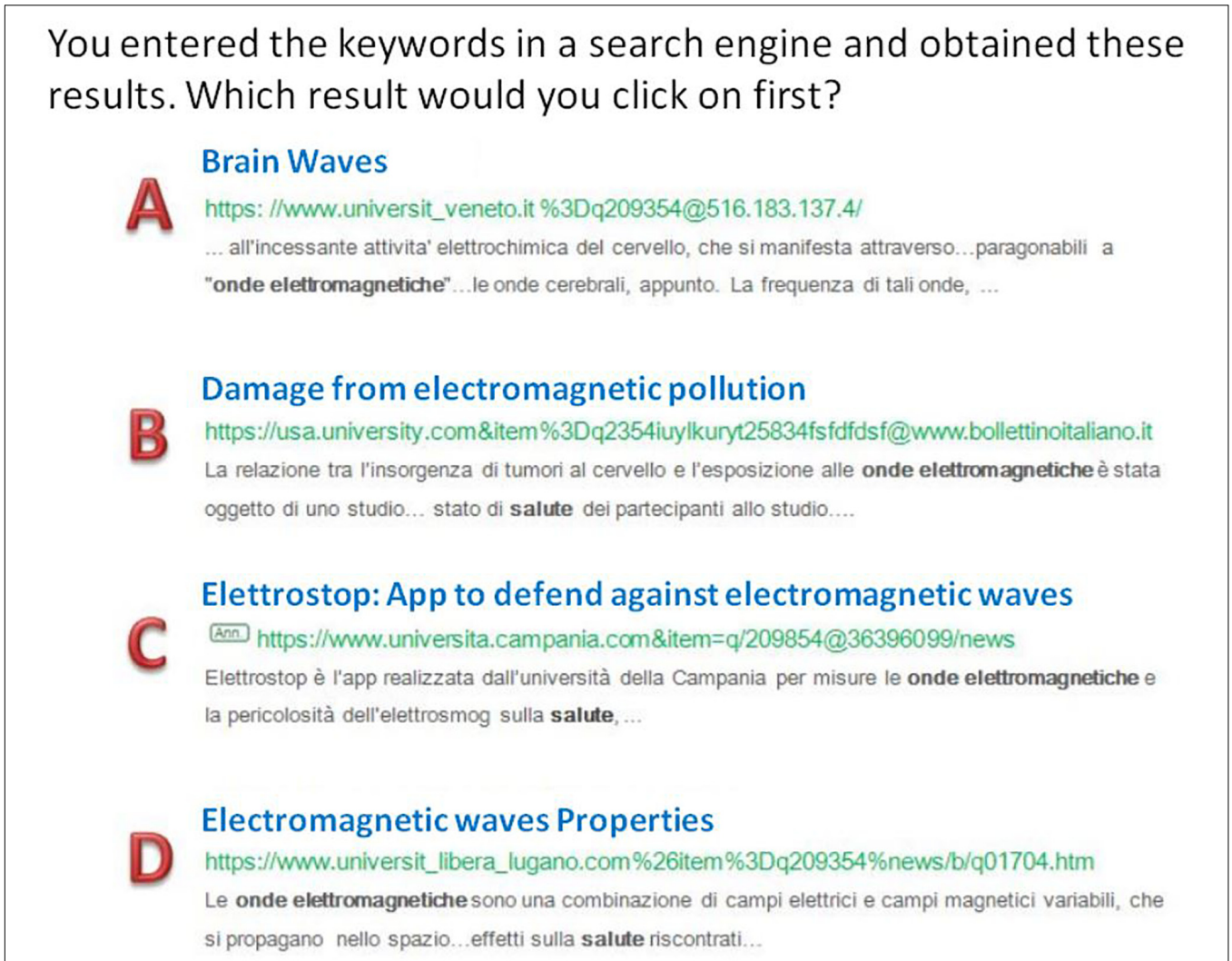


**FIGURE 2 |** Example of a question with the four multiple-choice answers. The question was preceded by an introductory item: "You have been requested to write a report on the relationship between electromagnetic waves and health. You will have to find information on the web and prepare a presentation for your classmates" and the student had chosen the keywords for the search. Answer A was assigned 0 points (unrelated topics) as well as answer C (advertisement); answer B was assigned 2 points, and answer D 1 point (correct but less complete topics).

belonged to this group. The aim was to assess the equivalence of the two versions. One half of the students took first Version 1 and then Version 2; for the other half, the opposite order was followed. Only the scores obtained in the version that was taken first were included in the data to be analyzed, so as to avoid learning or tiredness effects. The scores obtained on the version taken second were used for comparison purposes only.

# RESULTS

## Pilot Studies

The scores collected from the pilot study ($N = 30$) were analyzed item-by item, in order to identify ceiling and floor effects. As a result, about 20% of the 38 initial items were either reformulated so as to facilitate item comprehension, or eliminated when more than 90% of the students gave correct responses. Moreover, eight new items were added to the list, so as to have a larger range of items to choose from in a later selection.

A second pilot study was run ($N = 10$) in order to confirm that the new items had a better distribution. Following this step, six more items were replaced by new ones with clearer formulation and/or more challenging content.

## Final Tests and Subject Selection

The final test included 29 items (plus five items for prior knowledge assessment) for both Version 1 and Version 2. These were administered, respectively, to 79 and 89 subjects. Since each of the 29 items had a forced-choice four-alternative structure, in which the correct response was granted two points, the overall score ranged 0–58. Subjects' mean scores ($\pm SD$) were 35.1 $\pm$ 7.2 for Version 1 and 32.2 $\pm$ 6.8 for Version 2.

At this stage, we tackled the issue of chance level – that is, subjects should be excluded who, due to low motivation, did not perform the test properly and selected responses at random. A simple Monte Carlo simulation study ($N$ = 4,000) confirmed that chance level was 19/58 points (which can be mathematically derived from the scores granted by each item), and found a standard deviation of 4.562. Since the shape of this simulated score distribution was almost perfectly Gaussian, we learnt that 95% of scores obtained by selecting responses completely at random would be below 26.5/58. Hence, setting a cut-off of score at 26.5 should exclude the vast majority of subjects who did not follow the instructions, in any of the items. There were indeed some subjects who did so: the cut-off led to the exclusion of 9 out of 79 subjects (11.4%) who were tested with Version 1, and 19 out of 89 subjects (21.3%) who were tested with Version 2. Any further analysis was carried out without those subjects.

## Psychometric Analysis, Item Selection and New Scores

Before performing inferential analyses, the general psychometric characteristics of the test were analyzed. Test coherence as measured by Cronbach's Alpha turned out to be rather low for both versions: 0.514 for Version 1 and 0.361 for Version 2. Thus, we selected items in order to achieve a Cronbach's Alpha of at least 0.6. First, we excluded all items showing a negative correlation with total scale score. Then we performed a stepwise procedure: at each step, Alpha was re-computed after having left out one of the items; the item associated with the highest increase in Alpha was excluded, and the procedure repeated on a new step. If on a same step two or more items led to the same increase in Alpha, the choice was made relying on the quality of the distribution of the items. This led Version 1 to be reduced to 24 items (Cronbach's Alpha = 0.605). Excluded items were A1_3, A1_11, A1_15, B2_9 and B2_13. Inter-item correlations ranged from −0.25 to 0.471; mean correlation was 0.063. Mean value ($\pm SD$) for the new scale was 30.52 ± 6.27. Inter-class correlation was 0.060, $F(70,1610)$ against 0 = 2.532, $p < 0.001$. For Version 2, it was not possible to reach Alpha = 0.6 even after reducing items to as few as 15, so it was decided to keep the 17-item version. In it, no item had a negative correlation with the total scale score and Alpha (equaling 0.548) had reached a plateau – further exclusion of items produces negligible increase of Alpha. The total score for Version 2 had mean ($\pm SD$) = 19.97 ± 5.1.

Aggregate scores were computed for each version based on the new selection of items, Vers1TOT and Vers2TOT. Aggregate scores for the various subscales were computed for Version 1 only (the reduced Version 2 had too few items in some of the subscales). These sub-scale specific scores were computed both as raw sums, and as percent of maximum possible score. Thus, for the LESC components, we obtained percent scores Lperc1, Eperc1, Sperc1, and Cperc1. Finally, a score was computed for each of the modalities with which the information was conveyed, again both as raw sums and as percent of the maximum possible score (Textperc1, Graphperc1, Imageperc1, and Videoperc1).

All scores of Version 1, and the total aggregate score of Version 2 underwent the statistical analyses we had planned for the whole sample, which are reported below.

## Tests of Specific Predictions/Questions

First of all, a general analysis of the distributional properties of the various variables and scales was performed, revealing sufficiently close-to-normal distributions for all variables.

Subsequently, planned analyses were run, yielding the following results.

### Comparison Between Different School Types

Firstly, the three school types were compared with one-way ANOVAs with respect to the two pre-tests of ORCA.IT, i.e., the sections testing their digital and web-surfing abilities and their prior knowledge on the two topics addressed in the test. Students from the three school types obtained very similar scores on the questionnaires on digital competence and web-surfing habits and skills. However, their prior knowledge on the two topics differed, $F(2,175) = 8.397$, $p < 0.001$. Tukey's post-test highlighted that the difference was due to students of Vocational schools scoring lower than both students of Technical schools ($p < 0.001$) and Lyceum ($p = 0.002$). A further analysis with repeated-measures ANOVA comparing the two types of topics (intra-subject factor) and the three schools (inter-subject factor) revealed that a significant main effect of School type [$F(1,173) = 8.397$, $p < 0.001$, partial $\eta^2 = 0.088$] and of Topic [$F(1,173) = 35.376$, $p < 0.001$, partial $\eta^2 = 0.170$] but also a significant School type by Topic interaction [$F(1,173) = 16.014$, $p < 0.001$, partial $\eta^2 = 0.156$]. This was due to students of Lyceum yielding better scores compared to students of Vocational schools on the questionnaire on Topic a (brain and music) and students of Technical schools yielding better scores on topic b questionnaires (electromagnetic waves and health). These effects are shown in **Figure 3**.

One-way ANOVAs were carried out on Version 1's total score and its various subscales L1, E1, S1, C1 as well as Text1, Graph1, Image1, and Video1, with School type (three levels: Lyceum, Technical, and Vocational school) as a factor. No significant differences emerged between the different school types, with the exception of the aggregate scale Text1perc, part of Version 1, $F(2,68) = 3.96$, $p = 0.024$, partial $\eta^2 = 0.104$. Tukey's *post hoc* test revealed that the difference was due to students of Lyceum ($M = 68.10$, $SD = 18.96$) scoring higher than students of Vocational Schools ($M = 51.67$, $SD = 19.92$), $p = 0.035$ (see **Figure 4**).

### Comparison Between Groups With and Without Reading Difficulties (RD, no-RD)

Some significant differences emerged when comparing the two groups (non-parametric tests were preferred, because in this case there were only seven subjects with reading difficulties). These were found with the aggregated score of Version 1 (Mann–Whitney's $U = 10.5$, $Z = -2.29$, $p = 0.022$, mean rank 36.35 versus 18.36 for students without and with reading difficulties, respectively), and particularly in the Evaluate percent scale (Mann–Whitney's $U = 91$, $Z = -2.48$, $p = 0.013$, mean rank

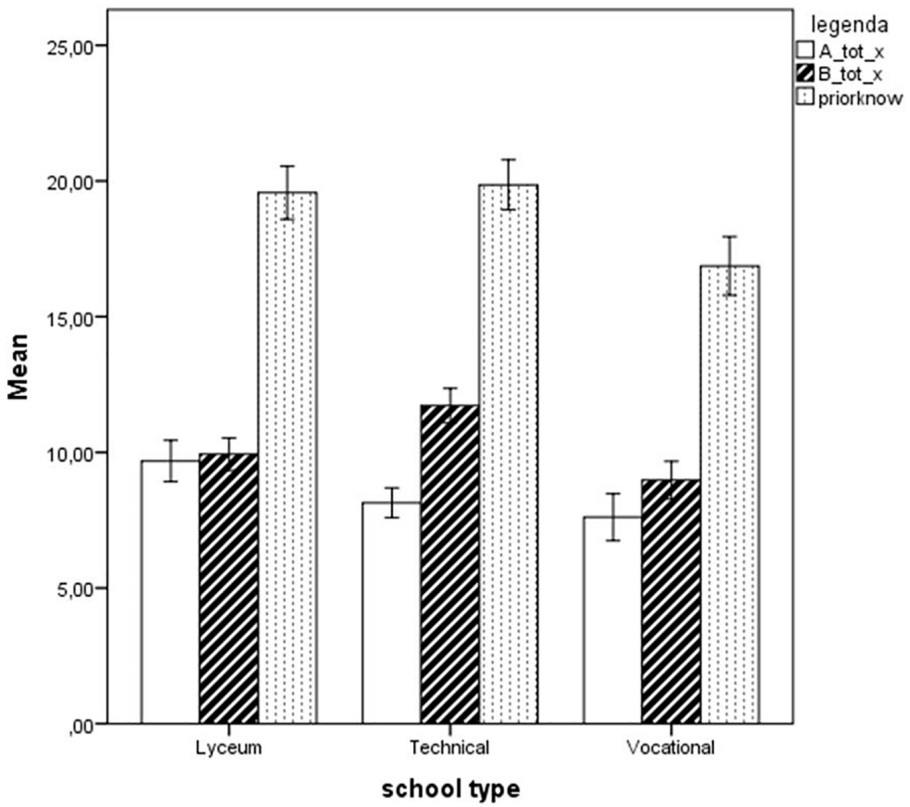


**FIGURE 3 |** Performance of students from the three different school types on the two topics A (music and brain) and B (electromagnetic waves and health), and total prior knowledge (the sum of the two). Error bars illustrate 95% Confidence Intervals for the Mean.

36.51 versus 17) and in the Text information percent scale (Mann–Whitney's $U = 119.5$, $Z = -1.905$, *one-tailed* $p = 0.029$, mean rank 36.04 versus 21.07). These results are illustrated in **Figure 5**.

## Comparison Between Males and Females

No significant differences emerged from either version between the two sexes, with the exception of the Graph scale, where males ($M = 60.53$, $SD = 35.18$) scored higher than females ($M = 43.55$, $SD = 33.52$), $F(1,68) = 4.147$, $p = 0.046$, partial $\eta^2 = 0.058$ (see **Figure 6**).

## Correlations Between the Aggregate Scores and Performances on Cognitive-Neuropsychological Tests

First of all, the scores obtained on all cognitive and neuropsychological tests were compared in the two groups of students who had been administered Version 1 and Version 2. No significant differences emerged in any of the variables (all $p_s > 0.05$).

Then, Pearson's bivariate correlation indices were calculated to analyze associations between the variables. *Z*-scores were considered whenever available for the standardized tests, so as to partial out the effects of (small) age differences among the students. Significant correlations are reported in bold in **Table 1**.

## Regression Analysis

A stepwise regression analysis was performed on total scores Vers1TOT and Vers2TOT to identify the predictors of performance on each of the two versions, based on the results of correlation analysis and *a priori* expectations. Criteria for entering the regression equation related to probability of F: $p < 0.05$ for entering, $p \geq 0.1$ for removal.

Predictors were entered in three subsequent steps: IQ was entered first, followed by all subscales related to web use (frequency, competence, technical abilities, and surfing habits) and prior knowledge; then all other potential predictors were entered, more precisely language skills (vocabulary and syntactic abilities), attentional variables (expressed by both objective and self-reported measures concerning organization capacity, attentional shifting and focusing, sustained attention, concentration, self-management and emotional control), reading (words and non-words, speed, and accuracy), and memory measures (forward and backward digit span). Due to missing scores for the neuropsychological tests, 43 participants out of 71 were included for Version 1 and 51 out of 77 were included for Version 2.

For Version 1 (reduced to 24 item), seven variables were included in the equation in subsequent steps: IQ (explaining 13.6% of total variance), prior knowledge (20.9%), Offline reading comprehension (9.6%), self-reported attentional skills

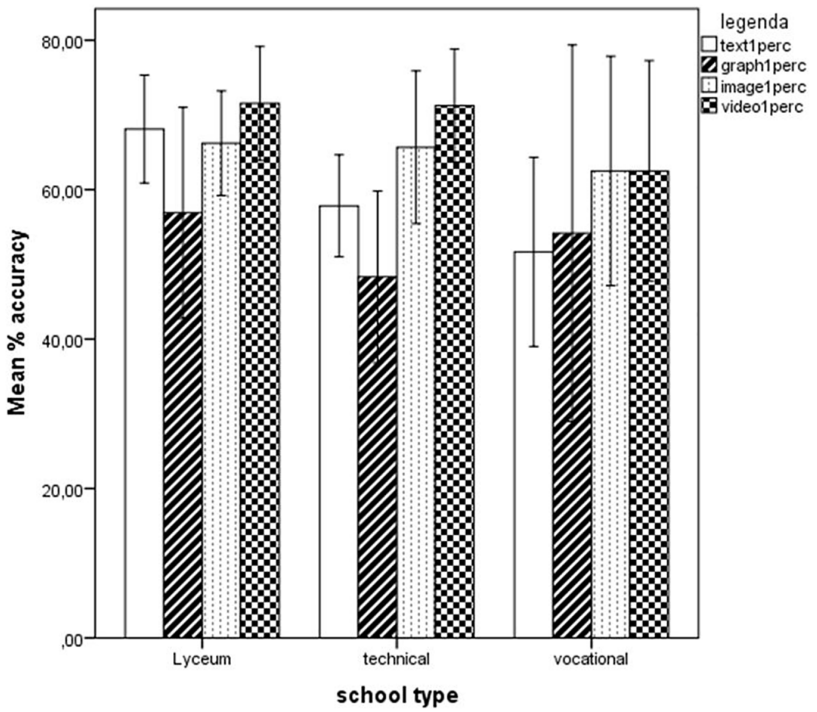


**FIGURE 4 |** Performance of students from the three different school types on the four subscales Text, Graph, Image, and Video, expressed as percent scores. Error bars illustrate 95% Confidence Intervals for the Mean.

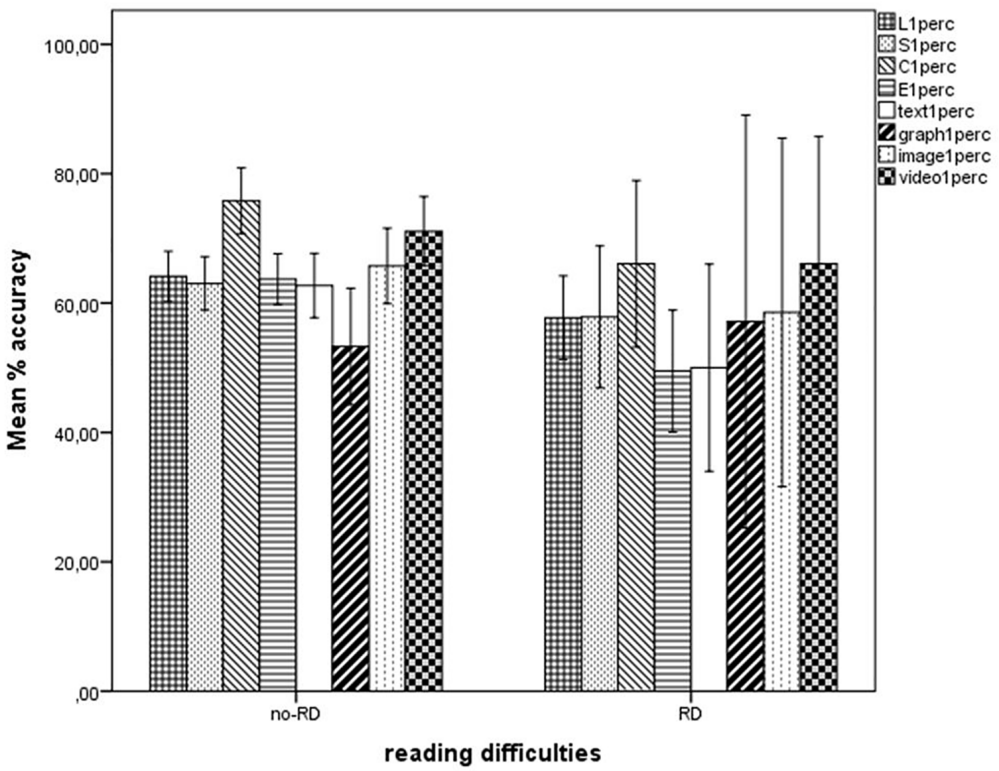


**FIGURE 5 |** Results obtained by students with and without reading difficulties on the four subscales Locate, Evaluate, Synthesize, and Communicate and on the four types of materials Text, Graph, Image, and Video, expressed as percent scores. RD, reading difficulties. Error bars illustrate 95% Confidence Intervals for the Mean.

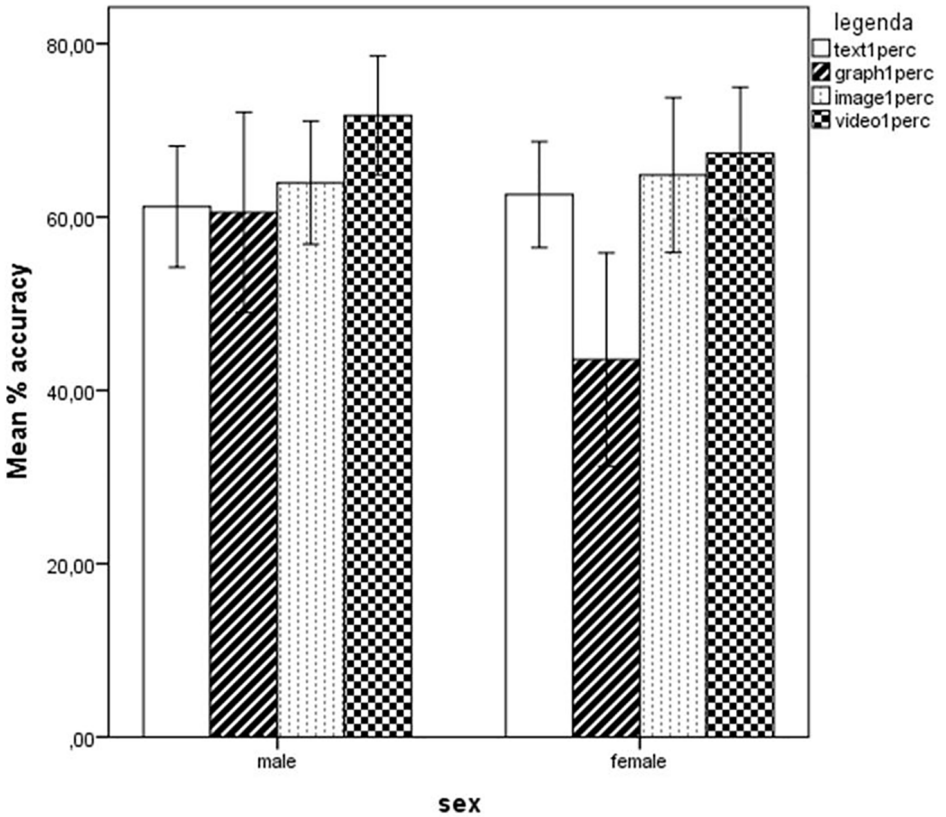


**FIGURE 6 |** Results obtained by males and females on the four subscales Text, Graph, Image, and Video, expressed as percent scores. Error bars illustrate 95% Confidence Intervals for the Mean.

related to Working Memory (5.9%), Forward Digit span (7.2%), attentional skills related to Concentration (6.4%), self-reported attentional skills related to concentration (4.4%); overall, these predictors explained 68% of variance, $F(7,40) = 9.999$, $p < 0.001$. No multicollinearity issues were present (VIF ranged between 1.131 and 1.924).

For Version 2 (reduced to 17 items), four variables entered the regression equation: IQ (explaining 9.4% of total variance), attention/concentration scores (14.9%), non-word reading accuracy (7.5%) syntax (6%). Overall, 38% of total variance was explained, $F(4,50) = 6.993$, $p < 0.001$. No multicollinearity issues were present (VIF ranged between 1.019 and 1.493).

The coefficients for the two regression equations are reported in **Tables 2**, **3**.

## DISCUSSION

The study described a new online test of online research and comprehension developed for the Italian population, that was named ORCA.IT.

Several questions, mostly regarding a number of *a priori* predictions, were addressed, which are reported and discussed in the following, under separate subheadings.

### What Are the General Psychometric Characteristics of the New Test?

The new test turned out to have good psychometric properties after accurate item construction and final selection. In particular, Version 1 showed better characteristics, with Skewness and Kurtosis for total score distribution being 0.073 and 0.723, respectively. The total scale was subdivided into subscales reflecting the structure that had inspired item construction, i.e., the so-called LESC structure (Locate, Evaluate, Synthesize, and Communicate online information). Also these subscales have good psychometric properties with close-to-normal distributions. As to Version 2, it was found to have poorer internal consistency and it was reduced to a 17-item scale in order to improve it, thus reaching sufficient reliability in terms of inter-item correlations, item-to-total scale correlations and Cronbach's Alpha. For this reason, no subscales were computed and the scale cannot be considered as an equivalent, parallel version of Version 1, but rather as a different, shorter version with different characteristics.

### Are There Any Differences Between the Different Types of School?

The different school types involved in the study are representative of the school population in the Italian system, but can also be

**TABLE 1 |** Pearson correlations between ORCA.IT variables and prior knowledge (digital competence and familiarity with topics), IQ, language skills (vocabulary and syntactic abilities), attentional variables (expressed by both objective and self-reported measures concerning organization capacity, attentional shifting and focusing, sustained attention, concentration, self-management, and emotional control), reading (words and non-words, speed, and accuracy), and memory measures (forward and backward digit span).

| | Version 1 total | L1 | S1 | C1 | E1 | Text 1 | Graph 1 | Image 1 | Video 1 |
|---|---|---|---|---|---|---|---|---|---|
| PRIOR KNOWL | **0.367**** | **0.355**** | **0.424**** | **0.088** | **0.405**** | **0.315**** | 0.131 | 0.114 | 0.222 |
| | *0.002* | *0.002* | *0.000* | *0.466* | *0.000* | *0.007* | *0.275* | *0.344* | *0.063* |
| | (71) | (71) | (71) | (71) | (71) | (71) | (71) | (71) | (71) |
| IQ | **0.285*** | 0.227 | 0.200 | 0.212 | **0.244*** | **0.309**** | 0.005 | 0.137 | 0.186 |
| | *0.018* | *0.060* | *0.100* | *0.080* | *0.044* | *0.010* | *0.964* | *0.260* | *0.127* |
| | (69) | (69) | (69) | (69) | (69) | (69) | (69) | (69) | (69) |
| FW D_SPAN | **0.272*** | 0.138 | 0.109 | 0.236 | 0.155 | **0.259*** | **0.255*** | −0.115 | 0.160 |
| | *0.025* | *0.262* | *0.375* | *0.053* | *0.208* | *0.033* | *0.036* | *0.352* | *0.192* |
| | (68) | (68) | (68) | (68) | (68) | (68) | (68) | (68) | (68) |
| BW D_SPAN | 0.235 | 0.107 | 0.182 | **0.281*** | 0.116 | **0.286*** | −0.067 | 0.014 | 0.049 |
| | *0.054* | *0.387* | *0.137* | *0.020* | *0.348* | *0.018* | *0.587* | *0.911* | *0.692* |
| | (68) | (68) | (68) | (68) | (68) | (68) | (68) | (68) | (68) |
| OFFL TEXT COMPR | **0.414**** | **0.337**** | **0.347**** | 0.090 | **0.430**** | 0.228 | −0.054 | **0.391**** | **0.251*** |
| | *<0.001* | *0.005* | *0.003* | *0.463* | *<0.001* | *0.060* | *0.660* | *0.001* | *0.037* |
| | (69) | (69) | (69) | (69) | (69) | (69) | (69) | (69) | (69) |
| PEABODY | **0.379*** | **0.343*** | **0.382**** | **0.295*** | **0.416**** | **0.315*** | −0.076 | **0.328*** | **0.350*** |
| | *0.010* | *0.021* | *0.010* | *0.049* | *0.005* | *0.035* | *0.622* | *0.028* | *0.018* |
| | (45) | (45) | (45) | (45) | (45) | (45) | (45) | (45) | (45) |
| SYNTAX | 0.263 | 0.281 | 0.208 | 0.031 | 0.318* | 0.232 | **−0.326*** | **0.385**** | 0.237 |
| | *0.077* | *0.058* | *0.165* | *0.838* | *0.031* | *0.121* | *0.027* | *0.008* | *0.112* |
| | (46) | (46) | (46) | (46) | (46) | (46) | (46) | (46) | (46) |
| NW_READ_ACC | **0.243*** | 0.041 | 0.037 | 0.135 | **0.268*** | **0.308*** | −0.174 | 0.125 | 0.020 |
| | *0.044* | *0.739* | *0.761* | *0.269* | *0.026* | *0.010* | *0.154* | *0.304* | *0.870* |
| | (69) | (69) | (69) | (69) | (69) | (69) | (69) | (69) | (69) |
| ATTENTION_CORR_RESP | **0.252** | 0.165 | 0.170 | 0.156 | **0.256** | **0.251** | −0.131 | 0.216 | 0.149 |
| | *0.081* | *0.257* | *0.242* | *0.284* | *0.075* | *0.082* | *0.369* | *0.137* | *0.307* |
| | (49) | (49) | (49) | (49) | (49) | (49) | (49) | (49) | (49) |
| ATTENTION_OMISS | −0.053 | 0.155 | 0.034 | 0.022 | −0.141 | **−0.260** | −0.010 | 0.055 | 0.191 |
| | *0.719* | *0.287* | *0.819* | *0.878* | *0.332* | *0.071* | *0.946* | *0.710* | *0.190* |
| | (49) | (49) | (49) | (49) | (49) | (49) | (49) | (49) | (49) |
| ATTENTION_CONC | 0.184 | 0.012 | 0.115 | 0.073 | 0.250 | **0.311*** | −0.096 | 0.103 | 0.011 |
| | *0.206* | *0.933* | *0.433* | *0.620* | *0.084* | *0.030* | *0.511* | *0.479* | *0.938* |
| | (49) | (49) | (49) | (49) | (49) | (49) | (49) | (49) | (49) |
| SR_ATT_CONC | **−0.302*** | −0.178 | **−0.296*** | −0.164 | **−0.283*** | **−0.310*** | −0.050 | −0.171 | −0.063 |
| | *0.035* | *0.221* | *0.039* | *0.261* | *0.049* | *0.030* | *0.731* | *0.240* | *0.669* |
| | (49) | (49) | (49) | (49) | (49) | (49) | (49) | (49) | (49) |

*KNOWL, knowledge; FW D_SPAN, forward digit span; BW D_SPAN, backward digit span; OFFL TEXT COMPR, offline text comprehension; NW_READ_ACC, non-word reading accuracy; ATTENTION_CORR_RESP, attention test – correct responses; ATTENTION_OMISS, attention test – omissions; ATTENTION_CONC, attention test – concentration; SR_ATT_CONC, self-reported attention test – concentration. Significant correlations are reported in bold, p-values in italics, N in parentheses.*

easily compared with the educational systems of other countries, with more theory-focused schools (Lyceum, here represented by a scientific and by a sport + science curriculum), technical schools (here represented by graphics and chemistry curricula), and more practice-oriented, vocational schools (here represented by commercial, education, and mechanics curricula). The three types of schools did not reveal clear differences with respect to students' general performance on the ORCA.IT test. This is encouraging since it suggests that the different educational pathways have provided similarly effective (though certainly different in content) opportunities to develop online search and reading abilities. Furthermore, students from the three types of schools with different educational and professional orientation were found not to differ in digital skills (on digital competence and web-surfing habits and skills). This preliminary data is comforting if we consider the choice of the school as an expression of a socio-economic condition. In fact, research has pointed to correlations between the users' socio-economic and cultural background and their use of ICTs (Gui and Argentin, 2011; Buffardi and Taddeo, 2017). Indeed, the failure to acquire digital skills can reproduce or even increase existing social inequalities (Hargittai, 2008; Bracciale and Mingo, 2015; van Deursen and Helsper, 2015). It is probable that the greater familiarity and greater exposure to

**TABLE 2 |** Coefficients of the regression equation for Version 1 total score (24 items).

| | Non-standardized coefficients | | Standardized coefficients | *t* | *p*-value |
|---|---|---|---|---|---|
| | B | Standard Error | Beta | | |
| (Constant) | −21.335 | 10.883 | | −1.96 | 0.058 |
| IQ | 0.095 | 0.085 | 0.133 | 1.128 | 0.268 |
| PRIOR KNOWL | 0.524 | 0.183 | 0.299 | 2.857 | 0.007 |
| OFFL TEXT COMPR | 1.534 | 0.507 | 0.35 | 3.023 | 0.005 |
| SR_ATT_WM | 1.027 | 0.253 | 0.531 | 4.064 | <0.001 |
| FWD_SPAN | 1.473 | 0.416 | 0.378 | 3.544 | 0.001 |
| ATTENTION_CONC | 0.047 | 0.016 | 0.352 | 2.951 | 0.006 |
| SR_ATT_CONC | −0.395 | 0.186 | −0.29 | −2.12 | 0.042 |

*KNOWL, knowledge; OFFL TEXT COMPR, offline text comprehension; SR_ATT_WM, self-reported attention test – working memory; FW D_SPAN, forward digit span; ATTENTION_CONC, attention test – concentration; SR_ATT_CONC, self-reported attention test – concentration.*

**TABLE 3 |** Coefficients of the regression equation for Version 2 total score (17 items).

| | Non-standardized coefficients | | Standardized coefficients | *t* | *p*-value |
|---|---|---|---|---|---|
| | B | Error Standard Deviation | Beta | | |
| (Constant) | −13.987 | 8.891 | | −1.573 | 0.123 |
| IQ | 0.113 | 0.094 | 0.169 | 1.195 | 0.238 |
| ATTENTION_CONC | 0.046 | 0.013 | 0.411 | 3.501 | 0.001 |
| NW_READ_ACC | −1.305 | 0.499 | −0.309 | −2.616 | 0.012 |
| SYNTAX | 1.095 | 0.518 | 0.300 | 2.114 | 0.040 |

*ATTENTION_CONC, attention test, concentration; NW_READ_ACC, non-word reading accuracy.*

written information (text) or the opportunities provided by the socio-cultural background can account for much of this difference. Differences between school types in Italy have also been found with regard to reading speed. In other words, the ability to read and comprehend offline has sometimes been found to significantly differ between school types, with poorer performance emerging in vocational institutes compared to lyceums (Stella and Tintoni, 2007). What is suggested by the present results, by contrast, is that online reading and web-based activities may have the potential to minimize such differences and reduce the gap in career opportunities and professional and personal satisfaction.

Nonetheless, differences across school types emerged in the qualitative profiles. Specifically, the comprehension and processing of textual information was found to be more difficult for students of the Vocational Schools. This may reflect a preference by students with higher verbal abilities to choose more theory-oriented schools, and vice versa, for students with lower verbal but higher practical skills to choose Vocational schools. Indeed, this is also supported by the much higher percentage of students with a diagnosis of Specific Reading Disorders in Vocational (and Technical) schools as compared to Lyceum where they are rare. It is also interesting to note that students from Technical schools were found to have even higher levels of prior knowledge on technical – oriented topics, whereas students from Lyceum were more familiar with more scientific topics, reflecting the relevance of the specific topic rather than of general abilities and general education on performance on the ORCA.IT test. In other terms, the two topics, which are both represented in both versions of the test, seem to have been effective in balancing and reducing the effects of possible specific prior educational differences. Clearly, as shown also by regression analysis, prior knowledge does have an influence on performance: this is in line with most previous studies (though not all those examining performance on ORCA instruments) and with our hypotheses, and a complete elimination of such effects would probably be both non-realistic and inappropriate.

## Are There Any Differences Between Males and Females?

Generally, males and females showed similar performances. Males were found to score better than females on a specific subscale only, i.e., the Graph subscale of Version 1. This might be related more to the usually reported higher ability of males on visual-spatial tasks than to specific digital skills or web-based search and comprehension skills. Similarly, in a study on Singapore secondary school students (Wu, 2004), male students performed better in graph reading, female students in graph construction, whereas no gender differences were found in graph interpretation and evaluation. While a general advantage for males in online reading performance cannot be documented due to the very little difference found in our sample (with marginal statistical significance), it can rather certainly be excluded that a general advantage for females (described for instance by Kanniainen et al., 2019 or Forzani, 2016) can be found in the Italian school

system when accessing and using digital information. This, following Jang and Ryoo's (2018) suggestions, might reflect a less positive attitude toward digital media in Italian girls, even for academic purposes (different, in this case, from both North-European and Asian female students). Moreover, a similar advantage for boys in digital reading had been described also by Rasmusson and Åberg-Bengtsson (2015), who put this in relation with the habit for males (more than for females) to spend many hours in playing with digital games, a difference that likely characterizes Italian adolescents too.

## Are There Any Differences Between Students With and Without Reading Difficulties?

Significant differences emerged between students with and without reading difficulties on total performance for Version 1, more specifically in the Evaluate subscale and for textual information. Such differences confirm our hypotheses about the disadvantage that students with reading difficulties would face when confronted with text (even if voice-to-speech technology was available to support reading). This is likely to have its origin in the differences that were found between the groups in offline reading comprehension, syntactic comprehension, verbal memory (in addition to obvious impairments in reading speed and accuracy). As to the Evaluate component, the difference emerging from the results also points to a lower level of critical appraisal and capacity to assess the reliability of online sources and choose the best source of information with respect to a specific question. This is in line with what reported by previous studies describing dyslexic students' difficulties in understanding page layout, navigating, understand color function and complex texts overall (e.g., Kurniawan and Conroy, 2007). The differences in performance found for dyslexic students also point out that voice-to-speech support is probably not sufficient to help them overcome their comprehension difficulties. This should be taken as a suggestion to intensify research on further technological aids (see Nickerson, 2005) that could help not only decoding (which seems not to be so crucial) but especially text parsing and information collection. Use of colors is an example, but further and more sophisticated strategies could exploit the much greater flexibility of online texts and sources with respect to offline ones. Some online services, for instance, already offer the possibility to change text features in real-time according to the individual needs of dyslexic readers (e.g., changing spatial placement of text boxes and figures, types of characters used, inter-character, and inter-lines spacing etc.), and international groups are working on the development of systems able to simplify the syntactic structure of sentences.

## Are Online Reading Abilities Related to Offline Reading Skills?

The answer to this question needs some distinctions to be made between reading skills (decoding) and comprehension abilities. Offline reading comprehension is very clearly related to online reading comprehension, as emerging from both correlation and regression analyses. This confirms what already reported in the literature (Leu et al., 2011; Kiili et al., 2018). As to decoding abilities, they show very little correlations with ORCA.IT variables and they enter the regression equation for Version 2 only, together with other language skills, suggesting that some linguistic difficulty in item formulation and in the texts to be read has significantly contributed to test performance, possibly due to the presence of students with very varying levels of verbal and reading abilities (often correlated in turn). No contribution of decoding skills has been shown for the most reliable version of the test, Version 1. Indeed, it has been previously shown that the relationship between decoding and comprehension is stronger when comprehension is assessed with a cloze test, but weaker when using multiple-choice questions (Francis et al., 2005); moreover, decoding skills have more impact on comprehension for younger/less skilled readers than for older/more able ones; further, for short rather than long passages (Keenan et al., 2008). Thus, the materials used in the present study, as well as the characteristics of participants, were likely to make the test more independent of decoding.

## What Are the Predictors and Components of Online Reading Abilities (Digital Experience, Prior Knowledge, Offline Reading Comprehension, STM, WM, and Executive Skills)?

The regression analysis showed a very interesting set of underlying skills explaining a large portion (68%) of the variance for Version 1, and a different set of variables explaining a smaller proportion (36%) of variance for Version 2. Specifically, performance on Version 1 is explained by general non-verbal reasoning, prior knowledge on the topics, offline text reading comprehension, attention (especially concentration) and memory skills. This is totally consistent with what reported in published studies about the predictors and components of online reading skills (Cromley et al., 2010; Tarchi, 2010; Snowling, 2013; Kanniainen et al., 2019). A different explanation is probably necessary for the results of the regression analysis for Version 2, where linguistic difficulty seems to have been central to performance, and may also have biased the reliability and internal consistency of the results. In this case, the contribution of decoding skills, as suggested above, could be linked to general language ability (reading and language abilities being moderately associated) rather than representing an independent underlying factor. It should also be considered that the students could use text-to-speech to support reading, thus further minimizing the effect of decoding ability on text comprehension. Rather, higher-level language skills involved in text analysis and syntactic comprehension, together with attention and concentration, appear to have strongly influenced performance levels. A special mention goes to the fact that general intelligence measured by non-verbal IQ scales reveals little

(and almost non-significant) contribution to online reading and comprehension skills.

It is also interesting to note that prior knowledge had a large effect in our results, differently from what was found by Kanniainen et al. (2019). Indeed, following Coiro (2011), it could be hypothesized that prior knowledge helps compensate lacking comprehension skills, but it is rather uninfluential when comprehension and online search skills are adequate; since the group of students in our study included several students with decoding difficulties, this may have strengthened the role played by familiarity with the topics.

As to the lack of effects by digital competence, we argue that general digital competence (which could be described as a more "technical" and procedural skill) as investigated in the pre-test module has no or little impact on comprehension of online information, which represents a complex, exquisitely cognitive integration and evaluation process. Indeed, such a clear-cut dissociation between procedural and processing skills may also reflect the fact that Internet searching in ORCA.IT was not real but simulated (so as to be less unpredictable, as a desired characteristic, but also less complex, an undesired side-effect) and the greater focus on information processing and comprehension than on more specific digital skills (since these are the object of existing assessment tools). The results show that the present test is in fact capable to capture such processing skills, which are, in our opinion, crucial requirements for the construction of real competence and knowledge, and relevant for integration, adaptation and success in educational and professional contexts.

## CONCLUSION

In conclusion, the newly developed test, especially in the most complete version (Version 1) appears to capture students' web searching abilities and online reading comprehension in an objective, accurate and reliable way. The tool is able to highlight differences in ability between students with and without reading difficulties, not penalizing overall performance but allowing very specific weaknesses to be pointed out. Similarly, it seems to be able to capture differences due to both educational pathways (different school types) and social attitudes (differences between males and females). Even more interestingly, it shows to be clearly resting on specific cognitive and neuropsychological abilities which explain much of the total variance. Prior knowledge also influences the results, as expected. No contribution from digital competence is found, which is rather unexpected. This can be seen as a potentially positive feature of the instrument, which turns out to be rather independent of previous Internet experience and to measure more cognitively grounded processes related to information gathering, processing and communicating. Also, no influence of decoding ability emerged, possibly thanks (at least in part) to the text-to-speech facility implemented in the software.

The tool has thus the potential to be used as a screening tool to identify students who need a special training to improve their specific skills necessary for effective use of online information. It could also be useful in a clinical setting where adolescents with language and reading disorders undergo special trainings for cognitive empowerment and remediation.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Scientific and Ethics Committee of the Department of Psychology of the Catholic University of Milan. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

MC contributed to the conception of the study, definition of the experimental materials, performed data collection, and contributed to the writing of the manuscript. MG assisted in the definition of the experimental materials, helped in data collection, and wrote a section of the manuscript. AT performed part of the statistical analyses and participated in the interpretation and description of the results. MM assisted in the recruitment of participants. DS contributed to the definition and correction of neuropsychological tests. MV participated in the conception of the study, and organized and supervised data collection. ML took care of the conception and definition of the experimental design, assisted in the definition of experimental materials, participated in the interpretation of results, performed most statistical analyses, contributed to the writing of the manuscript, and supervised the whole study. All authors contributed to manuscript revision, read and approved the submitted version.

## FUNDING

This publication was supported by the CARIPLO Foundation (Grant 2017-NAZ-0131 "New technologies for education and their impact on students' well-being and inclusion") and by the Italian Ministry of Health (Grant code RC 2018-2019).

## ACKNOWLEDGMENTS

We acknowledge the precious contribution of Eng. Giorgio Sidari for the development of the ORCA.IT software.

## REFERENCES


Afflerbach, P. A., and Cho, B.-Y. (2010). "Determining and describing reading strategies: internet and traditional forms of reading," in *Metacognition, Strategy Use, and Instruction*, eds H. S. Waters, and W. Schneider, (New York, NY: Guilford), 201–225.

Amadieu, F., Tricot, A., and Mariné, C. (2009). Prior knowledge in learning from a non-linear electronic document: disorientation and coherence of the reading sequences. *Comput. Hum. Behav.* 25, 381–388. doi: 10.1016/j.chb.2008.12.017

Arina, S., Iervolino, I., and Stella, G. (2013). Prima raccolta di dati normativi per la valutazione della dislessia evolutiva negli adolescenti su un campione di scuola secondaria di secondo grado. [First norms for dyslexia assessment in adolescents from a secondary school sample]. *Dislessia* 10, 9–38.

Baumgartner, S. E., van der Schuur, W. A., Lemmens, J. S., and te Poel, F. (2017). The relationship between media multitasking and attention problems in adolescents: results of two longitudinal studies. *Hum. Commun. Res.* 44, 3–30. doi: 10.1093/hcre.12111

Becta. (2008). *Harnessing Technology: Next Generation Learning 2008-2014*. Coventry: Becta.

Bilal, D. (2000). Children's use of the Yahooligans! Web search engine: I. Cognitive, physical, and affective behaviors on fact-based search tasks. *J. Am. Soc. Inform. Sci.* 51, 646–665. doi: 10.1002/(sici)1097-4571(2000)51:7<646::aid-asi7>3.0.co;2-a

Bishop, D. V., and Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: same or different? *Psychol. Bull.* 130, 858–886. doi: 10.1037/0033-2909.130.6.858

Bracciale, R., and Mingo, I. (2015). "Digital divide in time of crisis in europe: do the rich get richer, the poor get poorer?," in *Exploring the Crisis: Theoretical Perspectives and Empirical Investigations*, eds A. Borghini, and E. Campo, (Pisa: Pisa University Press).

Bråten, I., Strømsø, H. I., and Britt, M. A. (2009). Trust matters: examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Read. Res. Q.* 44, 6–28. doi: 10.1598/rrq.44.1.1

Brickenkamp, R., Schmidt-Atzert, L., and Liepmann, D. (2010). *Test d2-R-Revision. Aufmerksamkeits-and Konzentrationstest [Test d2-R-Revision: attention and concentration test]*. Göttingen: Hogrefe.

Britt, M. A., and Gabrys, G. L. (2001). "Teaching advanced literacy skills for the World Wide Web," in *Learning and Teaching on the World Wide Web*, ed. C. R. Wolfe, (San Diego, CA: Academic Press), 73-90.

Brown, T. E. (1996). *Brown Attention-Deficit Disorder Scales: Manual*. San Antonio, TX: The Psychological Corporation.

Buckingham, D. (2007). Digital media literacies: rethinking media education in the age of the Internet. *Res. Comp. Int. Educ.* 2, 43–55. doi: 10.2304/rcie.2007.2.1.43

Buckingham, D. (2010). "Defining digital literacy," in *Medienbildung in Neuen Kulturräumen*, ed. B. Bachmair, (Berlin: VS Verlag für Sozialwissenschaften), 59–71. doi: 10.1007/978-3-531-92133-4_4

Buffardi, A., and Taddeo, G. (2017). The Web 2.0 skills of italian students: an empirical study in southern Italy. *Ital. J. Sociol. Educ.* 9, 45–76. doi: 10.14658/pupj-ijse-2017-1-4

Cattell, R. B. E., and Cattell, A. K. S. (1981). *Misurare l'intelligenza con i test "Culture Fair" [Measuring intelligence with culture-fair tests]*. Firenze: Organizzazioni Speciali.

Coiro, J. (2003). Reading comprehension on the internet: expanding our understanding of reading comprehension to encompass new literacies. *Read. Teach.* 56, 458–464.

Coiro, J. (2011). Predicting reading comprehension on the internet: contributions of offline reading skills, online reading skills, and prior knowledge. *J. Lit. Res.* 43, 352–392. doi: 10.1177/1086296x11421979

Coiro, J., and Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth grade skilled readers to search for and locate information on the Internet. *Read. Res. Q.* 42, 214–257. doi: 10.1598/rrq.42.2.2

Coiro, J., and Kennedy, C. (2011). The Online Reading Comprehension Assessment (ORCA) Project: Preparing Students for Common Core Standards and 21st Century Literacies. White Paper Based on Work Supported by the United States Department of Education under Award No. R305G050154 and R305A090608. Available at: http://tinyurl.com/7y5dt65 (accessed March 30, 2017).

Cope, B., and Kalantzis, M. (eds) (2000). *Multiliteracies: Literacy learning and the design of social futures*. Abingdon: Routledge.

Cornoldi, C., Baldi, A. P., and Giofrè, D. (2017). *Prove MT Avanzate-3-Clinica. [MT Tests for the Clinics - Advanced]*. Firenze: Giunti O.S.

Cornoldi, C., Pra Baldi, A., and Friso, G. (2010). *Prove MT avanzate di lettura e matematica 2 per il biennio della scuola secondaria di secondo grado [MT reading and mathematics tests for secondary school - advanced]*. Firenze: Giunti. O.S.

Cromley, J. G., Snyder-Hogan, L. E., and Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: a domain-specific test of the direct and inferential mediation model of reading comprehension. *J. Educ. Psychol.* 102, 687–700. doi: 10.1037/a0019452

Davies, R. S., and West, R. E. (2014). "Technology integration in schools," in *Handbook of Research on Educational Communications and Technology*, eds J. M. Spector, M. D. Merrill, J. Elen, and M. J. Bishop, (New York, NY: Springer), 841–853.

Dong, G., and Potenza, M. N. (2015). Behavioural and brain responses related to Internet search and memory. *Eur. J. Neurosci.* 42, 2546–2554. doi: 10.1111/ejn.13039

Dunn, L. M., and Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-PPVT*, 3rd Edn, Minneapolis, MN: American Guidance Service.

Eynon, R., and Malmberg, L. E. (2011). A typology of young people's Internet use: implications for education. *Comput. Educ.* 56, 585–595. doi: 10.1016/j.compedu.2010.09.020

Fabos, B. (2004). *Wrong Turn on the Information Superhighway: education and the commercialization of the Internet*. New York, NY: Teachers College Press.

Fabos, B. (2008). "The price of information: critical literacy, education, and today's Internet," in *Handbook of Research on New Literacies*, eds J. Coiro, M. Knobel, C. Lankshear, and D. J. Leu, (Mahwah, NJ: Lawrence Erlbaum Associates), 839–870.

Facer, K., and Furlong, R. (2001). Beyond the myth of the 'cyberkid': young people at the margins of the information revolution. *J. Youth Stud.* 4, 451–469. doi: 10.1080/13676260120101905

Firth, J., Torous, J., Stubbs, B., Firth, J. A., Steiner, G. Z., Smith, L., et al. (2019). The "online brain": how the Internet may be changing our cognition. *World Psychiatry* 18, 119–129. doi: 10.1002/wps.20617

Forzani, E. (2016). *Individual Differences in Evaluating the Credibility of Online Information in Science: Contributions of Prior Knowledge, Gender, Socioeconomic Status, and Offline Reading Ability. Doctoral Dissertation*. Storrs, CT: University of Connecticut.

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., and Gebhardt, E. (2014). *Preparing for Life in a Digital Age: The IEA International Computer and Information Literacy Study International report*. Berlin: Springer Open.

Francis, D. J., Fletcher, J. M., Catts, H. W., and Tomblin, J. B. (2005). "Dimensions affecting the assessment of reading comprehension," in *Children's Reading Comprehension and Assessment*, eds S. G. Paris, and S. Stahl, (Mahwah, NJ: Erlbaum), 369–394.

García, J. R., and Cain, K. (2014). Decoding and reading comprehension: a meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Rev. Educ. Res.* 84, 74–111. doi: 10.3102/0034654313499616

Giovagnoli, G., Vicari, S., Tomassetti, S., and Menghini, D. (2016). The role of visual-spatial abilities in dyslexia: age differences in children's reading? *Front.Psychol.* 7:1997. doi: 10.3389/fpsyg.2016.01997

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10. doi: 10.1177/074193258600700104

Graesser, A. C., Wiley, J., Goldman, S. R., O'Reilly, T., Jeon, M., and McDaniel, B. (2007). SEEK Web tutor: fostering a critical stance while exploring the causes of volcanic eruption. *Metacogn. Learn.* 2, 89–105. doi: 10.1007/s11409-007-9013-x

Gremigni, E. (2019). Competenze digitali e media education: potenzialità e limiti del piano nazionale scuola digitale [Digital competences and media education: potentiality and limits of the national plan for the digital school]. *Rivista Trimestrale di Scienza dell'Amministrazione* 1–21.

Gui, M., and Argentin, G. (2011). The digital skills of the internet natives. different forms of digital literacy in a random sample of northern italian high school students. *New Media Soc.* 13, 963–980. doi: 10.1177/1461444810389751

Hahn, N., Foxe, J. J., and Molholm, S. (2014). Impairments of multisensory integration and cross-sensory learning as pathways to dyslexia. *Neurosci. Biobehav. Rev.* 47, 384–392. doi: 10.1016/j.neubiorev.2014.09.007

Hahnel, C., Goldhammer, F., Naumann, J., and Kröhne, U. (2016). Effects of linear reading, basic computer skills, evaluating online information, and navigation on reading digital text. *Comput. Hum. Behav.* 55, 486–500. doi: 10.1016/j.chb.2015.09.042

Hargittai, E. (2008). "The Digital Reproduction of Inequality," in *Social Stratification*, ed. D. Grusky, (Boulder, CO: Westview Press).

Hargittai, E., and Hinnant, A. (2008). Digital inequality: differences in young adults' use of the Internet. *Commun. Res.* 35, 602–621. doi: 10.1177/0093650208321782

Henry, L. (2006). SEARCHing for an answer: the critical role of new literacies while reading on the Internet. *Read. Teach.* 59, 614–627. doi: 10.1598/rt.59.7.1

Jang, B. G., and Ryoo, J. H. (2018). Multiple dimensions of adolescents' reading attitudes and their relationship with reading comprehension. *Read. Writ.* 32, 1769–1793. doi: 10.1007/s11145-018-9926-6

Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. New York, NY: New York University Press.

Kanniainen, L., Kiili, C., Tolvanen, A., Aro, M., and Leppänen, P. H. (2019). Literacy skills and online research and comprehension: struggling readers face difficulties online. *Read. Writ.* 32, 2201–2222. doi: 10.1007/s11145-019-09944-9

Keenan, J. M., Betjemann, R. S., and Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: differential dependence on decoding and oral comprehension. *Sci. Stud. Read.* 12, 281–300. doi: 10.1080/10888430802132279

Kiili, C., Laurinen, L., Marttunen, M., and Leu, D. J. (2012). Working on understanding during collaborative online reading. *J. Lit. Res.* 44, 448–483. doi: 10.1177/1086296X12457166

Kiili, C., Leu, D. J., Marttunen, M., Hautala, J., and Leppänen, P. H. (2018). Exploring early adolescents' evaluation of academic and commercial online resources related to health. *Read. Writ.* 31, 533–557. doi: 10.1007/s11145-017-9797-2

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. New York, NY: Cambridge University Press.

Korte, W. B., and Hüsing, T. (2006). *Benchmarking Access and Use of ICT in European Schools 2006: Results from Head Teacher and A Classroom Teacher Surveys in 27 European Countries*. Brussels: European Commission.

Kuiper, E., Volman, M., and Terwel, J. (2008). Students' use of Web literacy skills and strategies: searching, reading and evaluating Web information. *Inform. Res.* 13, paper 351. Available at: http://informationr.net/ir/13-3/paper351.html (accessed February 15, 2018).

Kurniawan, S., and Conroy, G. V. (2007). "Comparing comprehension speeds and accuracy of online information in students with and without dyslexia," in *Advances in Universal Web Design and Evaluation*, eds S. Kuriawan, and P. Zaphiris, (Hersey: Idea Group), 271–292.

Law, N., Pelgrum, W. J., and Plomp, T. (2008). *Pedagogical Practices and ICT Use Around the World: Findings from an International Comparative Study. CERC Studies in Comparative Education)*. Hong Kong/Dordrecht: Comparative Education Research Centre, University of Hong Kong/Springer.

Leu, D. J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., and Timbrell, N. (2015). The new literacies of online research and comprehension: rethinking the reading achievement gap. *Read. Res. Q.* 50, 37–59. doi: 10.1002/rrq.85

Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., and Henry, L. A. (2013). "New literacies and : a dual-level theory of the changing nature of literacy, instruction, and assessment," in *Theoretical Models and Process of Reading*, eds D. Alvermann, N. Unruh, and R. B. Ruddell, (Newark: International Reading Association), 1150–1181. doi: 10.1598/0710.42

Leu, D. J., Kulikowich, J. M., Sedransk, N., Coiro, J., Liu, C., Cui, W., et al. (2014). "The ORCA Project: designing technology-based assessments for online research comprehension, and communication," In *Processdings of the Conference on the AERA*, Philadelphia, PA.

Leu, D. J., McVerry, J. G., O'Byrne, W. I., Kiili, C., Zawilinski, L., Everett Cacopardo, H., et al. (2011). The new literacies of online reading comprehension: expanding the literacy and learning curriculum. *J. Adolesc. Adult Lit.* 55, 5–14.

Livingstone, S., and Helsper, E. (2007). Gradations in digital inclusion: children, young people and the digital divide. *New Media Soc.* 9, 671–696. doi: 10.1177/1461444807080335

Logan, S., and Johnston, R. (2009). Gender differences in reading ability and attitudes: examining where these differences lie. *J. Res. Read.* 32, 199–214. doi: 10.1111/j.1467-9817.2008.01389.x

Logan, S., and Johnston, R. (2010). Investigating gender differences in reading. *Educ. Rev.* 62, 175–187. doi: 10.1080/00131911003637006

Lupo, S., Jang, B. G., and McKenna, M. (2017). The relationship between reading achievement and attitudes toward print and digital texts in adolescent readers. *Lit. Res.* 66, 264–278. doi: 10.1177/2381336917719254

McCarthy, J. E., and Swierenga, S. J. (2010). What we know about dyslexia and web accessibility: a research review. *Univers. Access Inf.* 9, 147–152. doi: 10.1007/s10209-009-0160-5

McDonald, S., and Stevenson, R. J. (1996). Disorientation in hypertext: the effects of three text structures on navigation performance. *Appl. Ergon.* 27, 61–68. doi: 10.1016/0003-6870(95)00073-9

McKenna, M. C., Conradi, K., Lawrence, C., Jang, B. G., and Meyer, J. P. (2012). Reading attitudes of middle school students: results of a U.S. survey. *Read. Res. Q.* 47, 283–306. doi: 10.1371/journal.pone.0133544

Meelissen, M. R. M., and Drent, M. (2008). Gender differences in computer attitudes: does the school matter? *Comp. Hum. Behav.* 24, 969–985. doi: 10.1089/g4h.2014.0056

Miur. (2015). *Piano Nazionale Scuola Digitale [National Plan for Digital Education]*. Rome: MIUR.

Nicholas, D., Rowlands, I., Clark, D., and Williams, P. (2011). Google generation II: web behaviour experiments with the BBC. *Aslib Proc.* 63, 28–45. doi: 10.1108/00012531111103768

Nickerson, R. S. (2005). "Technology and cognitive amplification," in *Intelligence and Technology: The Impact of Tools on the Nature and Development of Human Abilities*, eds R. J. Sternberg, and D. D. Preiss, (Mahwah, NJ: Lawrence Erlbaum), 3–27.

Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* 2, 175–220. doi: 10.1037//1089-2680.2.2.175

Office of Communications [Ofcom], (2004). *Strategies and Priorities for the Promotion of Media Literacy: A Statement*. London: Ofcom.

Ophir, E., Nass, C., and Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15583–15587. doi: 10.1073/pnas.0903620106

Pollak, D. (2001). *Dyslexia, the Self and Higher Education*. London: Trentham Books.

Rasmusson, M., and Åberg-Bengtsson, L. (2015). Does performance in digital reading relate to computer game playing? A study of factor structure and gender patterns in 15-year-olds' reading literacy performance. *Scand. J. Educ. Res.* 59, 691–709. doi: 10.1080/00313831.2014.965795

Raven, J. (2003). "Raven progressive matrices," in *Handbook of Nonverbal Assessment*, (Boston, MA: Springer), 223–237. doi: 10.1007/978-1-4615-0153-4_11

Rouet, J. F. (2006). *The Skills Of Document Use: From Text Comprehension to Web-Based Learning*. Abingdon: Routledge.

Salmerón, L., García, A., and Vidal-Abarca, E. (2018). The development of adolescents' comprehension-based Internet reading activities. *Learn. Individ. Differ.* 61, 31–39. doi: 10.1016/j.lindif.2017.11.006

Sanchez, C. A., Wiley, J., and Goldman, S. R. (2006). "Teaching students to evaluate source reliability during Internet research tasks," in *Proceedings of the Seventh International Conference On the Learning Sciences*, eds S. A. Barab, K. E. Hay, and D. T. Hickey, (Mahwah, NJ: Lawrence Erlbaum Associates), 662–666.

Sartori, G., Job, R., and Tressoldi, P. E. (2007). *DDE-2: Batteria per la Valutazione della Dislessia e della Disortografia Evolutiva – 2, Seconda Edizione [Battery for the Assessment Of Developmental Dyslexia and Dysorthographia, Second Edition]*. Firenze: Giunti O.S (Organizzazioni Speciali).

Selwyn, N. (2009). The digital native–myth and reality. *Aslib Proc.* 61, 364–379. doi: 10.1108/00012530910973776

Shewbridge, S., Ikeda, M., and Schleicher, A. (2006). *Are Students Ready for a Technology-rich World?: What PISA Studies Tell Us*. Paris: OECD.

Snowling, M. J. (2013). Early identification and interventions for dyslexia: a contemporary view. *J. Res. Spec. Educ. Needs* 13, 7–14. doi: 10.1111/j.1471-3802.2012.01262.x

Snowling, M. J., and Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: a meta-analysis and review. *Psychol. Bull.* 142, 498–545. doi: 10.1037/bul0000037

Sparrow, B., Liu, J., and Wegner, D. M. (2011). Google effects on memory: cognitive consequences of having information at our fingertips. *Science* 333, 776–778. doi: 10.1126/science.1207745

Stella, G., and Tintoni, C. (2007). Indagine e rilevazione sulle abilità di lettura nelle scuole secondarie di secondo grado [Investigation and report on reading abilities in secondary schools]. *Dislessia* 4, 271–285.

Swart, N. M., Muijselaar, M. M., Steenbeek-Planting, E. G., Droop, M., de Jong, P. F., and Verhoeven, L. (2017). Cognitive precursors of the developmental relation between lexical quality and reading comprehension in the intermediate elementary grades. *Learn. Individ. Differ.* 59, 43–54. doi: 10.1016/j.lindif.2017.08.009

Taboada, A., and Guthrie, J. T. (2006). Contributions of student questioning and prior knowledge to construction of knowledge from reading information text. *J. Lit. Res.* 38, 1–35. doi: 10.1207/s15548430jlr3801_1

Tarchi, C. (2010). Reading comprehension of informative texts in secondary school: a focus on direct and indirect effects of reader's prior knowledge. *Learn. Individ. Differ.* 20, 415–420. doi: 10.1016/j.lindif.2010.04.002

Torppa, M., Eklund, K., Sulkunen, S., Niemi, P., and Ahonen, T. (2018). Why do boys and girls perform differently on PISA Reading in Finland? The effects of reading fluency, achievement behaviour, leisure reading and homework activity. *J. Res. Read.* 41, 122–139. doi: 10.1111/1467-9817.12103

Tsai, C. C., Chuang, S. C., Liang, J. C., and Tsai, M. J. (2011). Self-efficacy in Internet-based learning environments: a literature review. *J. Educ. Technol. Soc.* 14, 222–240. doi: 10.1016/j.soard.2011.02.009

Uncapher, M. R., and Wagner, A. D. (2018). Minds and brains of media multitaskers: current findings and future directions. *Proc. Natl. Acad. Sci. U.S.A.* 115, 9889–9896. doi: 10.1073/pnas.1611612115

van Deursen, A. J., and Helsper, E. J. (2015). The third-level digital divide: who benefits most from being online? *Commun. Inf. Technol. Annu.* 10, 29–52. doi: 10.1108/S2050-206020150000010002

van Strien, J. L., Kammerer, Y., Brand-Gruwel, S., and Boshuizen, H. P. (2016). How attitude strength biases information processing and evaluation on the web. *Comput. Hum. Behav.* 60, 245–252. doi: 10.1016/j.chb.2016.02.057

Vernice, M., Matta, M., Tironi, M., Caccia, M., Lombardi, E., Guasti, M. T., et al. (2019). An online tool to assess sentence comprehension in teenagers at risk for school exclusion: *evidence from L2 Italian students. Front. Psychol.* (in press). doi: 10.3389/fpsyg.2019.02417

Vuorikari, R., Punie, Y., Gomez, S. C., and Van Den Brande, G. (2016). *DigComp 2.0: The Digital Competence Framework for Citizens, Update Phase 1: The Conceptual Reference Model (No. JRC101254)*. Seville: Institute for Prospective Technological Studies.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children - Fourth Edition*. San Antonio, TX: Pearson Education, Inc.

Wu, Y. (2004). "Singapore secondary school students understanding of statistical graphs," in *Proceedings of the 10th International Congress on Mathematics Education*, Copenhagen.

Zhao, L., Lu, Y., Huang, W., and Wang, Q. (2010). Internet inequality: the relationship between high school students' internet use in different locations and their Internet self-efficacy. *Comput. Educ.* 55, 1405–1423. doi: 10.1016/j.compedu.2010.05.010

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 05 November 2019
doi: 10.3389/fpsyg.2019.02417

# An Online Tool to Assess Sentence Comprehension in Teenagers at Risk for School Exclusion: Evidence From L2 Italian Students

*Mirta Vernice[1]*, Michael Matta[2], Marta Tironi[1], Martina Caccia[3], Elisabetta Lombardi[4], Maria Teresa Guasti[1], Daniela Sarti[5] and Margherita Lang[6]*

*[1] Department of Psychology, University of Milano-Bicocca, Milan, Italy, [2] Department of Psychological, Health and Learning Science, University of Houston, Houston, TX, United States, [3] Scientific Institute IRCCS Eugenio Medea, Bosisio Parini, Italy, [4] Department of Psychology, Catholic University of the Sacred Heart, Milan, Italy, [5] Foundation IRCCS Carlo Besta Neurological Institute, Milan, Italy, [6] A.R.P. Associazione per la Ricerca in Psicologia Clinica, Milan, Italy*

This study presents a web-based sentence comprehension test aimed at identifying high school students who are at risk for a language delay. By assessing linguistic skills on a sample of high school students with Italian as an L2 and their monolingual peers, attending a vocational school, we were able to identify a subgroup of L2 students with consistent difficulties in sentence comprehension, though their reading skills were within the average range. The same subgroup revealed to experience a lack of support within the school context, suggesting that poor L2 skills might be a critical variable to consider in order to identify students at risk for school exclusion. Regarding the test, accuracy to the on-line sentence comprehension task was significantly predicted by reading abilities and vocabulary skills, thus indicating that this test might represent a rapid but efficient way to assess linguistic abilities at school. We recommend that establishing a valid and practical procedure for the evaluation of linguistic skills in bilingual students who struggle with their L2 is the first step toward promoting social inclusion in the multilingual classroom, in order to increase their ability to actively participate in school and social activities.

**Keywords: sentence comprehension, L2 speakers, adolescents, school exclusion, online tool**

OPEN ACCESS

***Edited by:***
*Meryem Yilmaz Soylu, University of Nebraska–Lincoln, United States*

***Reviewed by:***
*Paolo Canal, University Institute of Higher Studies in Pavia, Italy*
*Xiaoxun Sun, Australian Council for Educational Research, Australia*
*Judit García-Martín, University of Salamanca, Spain*

****Correspondence:***
*Mirta Vernice*
*mirta.vernice@unimib.it*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *30 June 2019*
***Accepted:*** *10 October 2019*
***Published:*** *05 November 2019*

***Citation:***
*Vernice M, Matta M, Tironi M, Caccia M, Lombardi E, Guasti MT, Sarti D and Lang M (2019) An Online Tool to Assess Sentence Comprehension in Teenagers at Risk for School Exclusion: Evidence From L2 Italian Students. Front. Psychol. 10:2417. doi: 10.3389/fpsyg.2019.02417*

## INTRODUCTION

The Italian educational system is currently undergoing a significant change toward promoting inclusion of students coming from the most diverse backgrounds. To this aim, in the last few years, a tremendous transformation took place in terms of digital technology enhancement in the classroom. According to the national educational policy known as "The Italian National Plan for Digital Education" [Piano Nazionale Scuola Digitale] (Italian Ministry of Education [MIUR], 2015b), digital technology deployment in the classroom is aimed at fostering student engagement by creating a learner-centered environment that promotes inclusion of all students, in particular those with special needs (Italian Ministry of Education [MIUR], 2015a,b). However, in the very same years, the rapid growth in the number of students with learning difficulties of the most heterogeneous nature (e.g., specific learning disorders; atypical language acquisition; language or cultural deprivation) had a critical impact on the schooling system making it difficult to foresee educational programs able to integrate and include the greatest number of pupils and in particular those at risk of school exclusion (Contini, 2013).

Moving from the assumption that, in the digital classroom, teachers must be equipped with online instruments able to identify those students who might show a learning or linguistic problem that might prevent their full integration within the classroom, we present a web-based tool that allows discriminating those students that belong to a population at potential risk for a language problem, and thus, for school exclusion: teenagers speaking Italian as an L2 (henceforth L2 students).

## L2 Students: A Population at Risk for School Exclusion

In the last years, consistent growth in the number of immigrants has caused profound modifications in the Italian educational system. According to data from the Italian Ministry of Education [MIUR] (2014, 2017), the number of L2 students increased from 60,000 in 1997 to over 800,000 in 2016 (excluding universities). Crucially, in 10 years, the rate of L2 students attending a high school increased in percentage significantly more in comparison with other school levels. In fact, from 2007 to 2017, there were 82% more L2 students in high schools, as compared to lower secondary schools (45% more), primary schools and kindergartens (56% and 76%, respectively) (Italian Ministry of Education [MIUR], 2017). Considering only high schools, national reports indicate a total of 23% of immigrant students (including both first- and second-generation immigrants), with consistent differences across regions and type of high school (Italian Ministry of Education [MIUR], 2017). For instance, out of the total number of L2 students attending a high school, 92% of them, in 2016/2017, opted for a vocational-technical school. Additionally, Lombardy, the region where the current research took place, appears to be characterized by the highest presence of L2 students (25%) in high school classrooms with respect to the national average.

A recent report by the National Institute of Statistics (ISTAT, 2016) further indicates that L2 Italian students appear to be significantly more vulnerable than monolingual students in terms of academic failures. They tend to achieve lower academic outcomes, with higher proportions of dropouts and lower-levels of school attainment (ISTAT, 2016). Some studies based on Italian data (Murineddu et al., 2006) went further by indicating that low academic achievements in this population are associated with learning problems. Research has shown that the presence of L2 students in the classroom might exert an adverse effect on the well-being and inclusion in the classroom. For instance, Brunello and Rocco (2013) analyzed to what extent the presence of immigrant students exerts an impact on the school performance of L1 students at 15 years of age, finding evidence of a significant negative effect, increasing with the level of segregation of immigrants, as evidenced by larger dropout rates.

A number of previous studies investigated the difficulties of the educational pathway of immigrant students in Italy, by addressing the multiple factors that might concur in determining such condition. For instance, Mussino and Strozza (2012) presented a national survey (ITAGEN2) involving more than 20,000 lower-secondary school students, half of whom were L2 speakers. Starting from the observation that L2 students showed higher drop-out rates and fewer academic achievements, the survey allowed to identify a number of possible factors causing a lack of integration of immigrant students in the Italian school system, among which the authors included educational delay and socioeconomic deprivation. However, lack of school integration was significant even when immigrant students showed comparable socioeconomic and housing conditions in comparison with their Italian peers. The authors concluded that it was mainly socialization with peers that appeared to decrease the risk of isolation, facilitating a faster inclusion in the school context.

As for the role of academic and learning skills, a recent Italian study observed that students who are struggling in reading and writing, regardless of whether they were L1 or L2 speakers, developed low motivation with respect to their abilities, helpless behavior and anxiety in being involved in school activities (Andolfi et al., 2015). Again, this study confirmed that only support within the peer group reduced the risk of school failure and, most importantly, the feeling of exclusion.

The above-mentioned findings suggest that when studying the possible causes of school exclusion, it is crucial to consider not only aspects related to the socioeconomic level and learning profile, but also factors such as students well-being, including relationships with classmates, engagement, and support from peers (Deci and Ryan, 2008).

Interestingly, to our knowledge, relatively little research has directly addressed the role of L2 proficiency in promoting school inclusion of L2 students. Indeed, while it is uncontroversial that being proficient in the language of instruction represents, in general, a protective factor for L2 students (cf., in primary school, Whiteside et al., 2017), it is not yet consolidated that poor L2 skills might be directly related to risk of school exclusion in high school students. In the current study, we hypothesize that L2 proficiency of immigrant teens, tested at school by means of an *ad hoc* created test of sentence comprehension, could be regarded as a critical variable in order to identify those students that might experience difficulties in their educational pathway. That is, we propose that L2 students' proficiency in the language of instruction might offer valuable information to signal those students who could be at risk for school exclusion.

Before we move on to describe the test used to investigate language proficiency in L2 teens, we need to discuss the profile of bilingualism at stake in the current study, taking into consideration some factors, such as the cultural and socioeconomic background (i.e., socioeconomic status, SES) as well as the age of exposure to the L2 of bilingual students. In fact, in the current study, we are concerned with L2 students who come from a context of relative socioeconomic disadvantage and that were exposed to Italian since birth (except for two, who came to Italy when they were infants).

With regards to the SES, in Italy, a growing number of L2 students comes from immigrant families that live below the poverty level and have, on average, lower-levels of education compared to non-immigrant families (ISTAT, 2016). Research suggests that in monolingual development, students who come from a deprived context tend to be exposed to a less varied

and less rich linguistic input than that of children from high SES families (Myers and Botting, 2008). The same applies to immigrant children living in a context of relative socioeconomic disadvantage: L2 children often begin schooling without being able to speak the majority language and then start acquiring it during the preschool years (Spencer et al., 2012). Usually, a child becomes dominant in his/her L2 only when formal education begins, while dominance in other languages decreases (Montrul, 2008, 2015; Benmamoun et al., 2013; Puig-Mayenco et al., 2018). However, research indicates that many of them acquire a limited L2 proficiency to catch up to their monolingual peers. Crucially, in such a case, language problems cannot be attributed to bilingual development, rather to the deprivation in terms of language input (Spencer et al., 2017). Note indeed that studies conducted in countries where bilingualism is actively promoted (such as Canada), show that bilingual children are not at risk for a language problem, rather, that (early) exposure to two languages is related to several cognitive benefits (Bialystok, 1999; Bialystok and Martin, 2004; Bialystok and Feng, 2009; Luk et al., 2011; Paradis et al., 2011; Poarch and Bialystok, 2015; also with immigrant children see Robinson and Sorace, 2019).

A second factor in determining bilingual language outcome is the age at which children are first exposed to each language. According to the literature, children who are exposed to two languages from early infancy (the so-called "early bilinguals") generally achieve greater proficiency than speakers exposed to a language after 3 years of age (Flege et al., 1999; Kovelman et al., 2008). However, as Unsworth (2014, 2016) notes, what is crucial is not only the length of exposure to the L2 input but also the quality and the real amount of language use. For instance, L2 children are often exposed to input from both native and non-native L2 speakers. However, predominant exposure to non-native speakers is often associated with poorer levels of language proficiency. Therefore, it is important to consider the fact that even pupils that are exposed to Italian since birth might show a language delay in L2.

In the current study, we involved a sample of L2 participants as homogeneous as possible in terms of age of first exposure to the L2. We are aware of the fact that, by controlling this variable, the number of L2 participants was limited. As a consequence, the current study should be considered as an initial contribution aiming at characterizing the linguistic competence of early bilingual teens, coming from a deprived socioeconomic context.

## Sentence Comprehension: A Pivotal Ability Underlying Higher-Level Comprehension Processes

The ability to comprehend a syntactically complex sentence is a language skill purporting a variety of activities that characterize high school programs, such as reading complex passages, summarizing them, creating well-structured texts that effectively communicate the content. In adolescence, once the decoding process in reading is automatized, lower-level comprehension skills (i.e., receptive vocabulary and grammar) appear to support the ability to make inferences about the internal structure of complex texts (Gough and Tunmer, 1986) as well as about its lexical content (Kieffer and Box, 2013). That is, lower-level comprehension processes (related to vocabulary and grammar) facilitate the extraction of semantic and syntactic information that supports reading comprehension of connected text (van Gelderen et al., 2007). Thus, language comprehension appears as a critical prerequisite of text reading comprehension not only at sentence-level but also at text-level, promoting lexical and grammatical inference, and thus supporting academic achievement in later grades of school (Kamhi and Catts, 2012).

In L2 students, research suggests that even though typically developing L2 students show appropriate word reading skills since the very first years of primary school, their reading comprehension scores fall below average with respect to their peers (Menken, 2008). Along the course of development this difference in text reading comprehension increases in the upper-grade levels (Galindo and Reardon, 2006). For this reason, being able to identify those students who struggle with language comprehension is critical to the task of creating high-quality academic education and effective inclusion strategies.

Up-to-date research has focused mostly on children and there has been little investigation into reading comprehension skills of L2 adolescents. In general, assessment of language skills in adolescence is difficult due to the lack of standardized tests for this age group in Italian. Additionally, according to Italian data regarding language and communication difficulties, adolescents are rarely assessed for previously undetected language problems (Brizzolara et al., 2011). This study, therefore, represents one first attempt to identify language difficulties in Italian L2 teens, by using an online sentence comprehension text. To this aim, we created a task that was based on 20 multiple-choice trials. Each trial involved a target sentence of varying syntactic complexity that had to be read silently. Then four additional sentences followed. One was a paraphrase of the target sentence, equivalent in meaning to the target sentence; the second sentence contradicted the meaning of the target sentence; the third one involved a meaning compatible with the target sentence, but not equivalent; the fourth one was a distractor involving a different content. Correct responses were based on the number of equivalent sentences detected. The sentence comprehension test was presented using Google forms. The choice of a web-based tool, instead of a canonical paper and pencil test, moved from the consideration that L2 adolescents tend to show low levels of participation and involvement to research (Myers and Botting, 2008; Clegg et al., 2009), suggesting that this population may be hard to reach and to be involved in projects. Digital technology might come in hand, being a particularly useful tool for motivating learning (Zhang, 2008). In particular, many observations about motivating L2 students in classroom situations support the idea that digital technology provides "motivational affordances" (ibidem). Students show a feeling of intrinsic motivation in performing an online task when they experience competence, autonomy and social connection (Ryan and Deci, 2000). Therefore, in such a case, the opportunity to use a digital device to run the assessment would result in a more committed participation in the study.

To sum up, the current study aims: (1) to compare the performance of L1 and L2 students on a series of standardized

tests as well as on a web-based assessment of sentence comprehension; (2) to identify the profile of L2 students whose performance on the sentence comprehension assessment proved to be insufficient to allow the student to actively participate in school activities. In general, this study aimed at promoting the use of a web-based tool in order to facilitate the identification of subjects at risk for a language problem. In fact, by gaining further information about the type of L2 linguistic development of the student and the eventual occurrence of language difficulties, teachers might be able to effectively supply to the needs of students.

# MATERIALS AND METHODS

## Participants

A total of 44 teenagers who attended a vocational high school in the region of Lombardy (Italy) participated in this study. Participants ranged in age from 14;9 (years; months) to 17;4 (Mean age in months = 189.66, SD = 7.09). Participants were divided into two groups: L1 Italian students (henceforth L1 group; n = 22, 3 boys; Age range: 14;9 – 17;1; Mean age in months = 189.59, SD = 6.90) and L2 Italian (L2; $n$ = 22, 3 boys; Age range: 14;10 – 17;40; Mean age in months = 189.73, SD = 7.43), matched for gender and chronological age ($t = -0.063$, $p = 0.950$).

First languages of L2 students were Chinese (3), Arabic (7), Spanish (1), Albanian (3), Roumanian (4), Punjabi (4). Twenty of them were born in Italy, while the remaining two came to Italy during early infancy (i.e., before than 2 years of age; Mean total exposure to Italian in months = 182.63, SD = 23.47). L2 students could thus be considered early bilinguals (Kovelman et al., 2008). All participants received their formal education only in Italy.

To be eligible to participate, students had to meet a number of criteria. First, they reported no history of learning, cognitive, neurological, or sensory disorders. Second, their cognitive level fell within the normal range (above the 25th percentile). Third, they did not receive any special educational support (according to school reports) because of language problems.

Written informed consent was obtained from the parents of all participating students. The protocol was approved by the Ethical Committees of the University of Milano-Bicocca and of the Catholic University of the Sacred Heart.

## Materials

To address our research questions, participants took part in an online assessment of sentence comprehension. Additionally, they were administered a battery of standardized tests of cognitive level, reading fluency and comprehension, vocabulary, and SES.

## Standardized Tests

Participants were administered a series of tasks from standardized tests, assessing, respectively: (a) non-verbal cognitive level; (b) receptive vocabulary (The Peabody Picture Vocabulary Test Form L; PPVT); (c) reading fluency and passage reading comprehension; (d) SES; (e) level of bilingual exposure (only for the bilingual group); (f) psychological well-being (CIT).

**Table 1** offers a summary of the different factors that concur in determining a condition of vulnerability at school, namely, L2 proficiency and dominance, learning profile, SES, psychological well-being. On each row, we report constructs investigated and tests used. Tests and questionnaires are described below.

To investigate the non-verbal cognitive level, we used the Raven's Progressive Matrices Test (SPM version; Raven and Court, 1998). SPM measures fluid intelligence (Horn and Cattell, 1966) which correlates strongly with IQ scores (Kvist and Gustafsson, 2008) and involves low degrees of cultural loading and linguistic demand (Flanagan et al., 2013). We used this measure to the extent of identifying participants whose cognitive non-verbal level was within the normal range (above the 25th percentile).

Vocabulary skills were assessed by means of The Peabody Picture Vocabulary Test Form L (PPVT, Dunn and Dunn, 1981). The primary aim of this test is to measure the receptive vocabulary for standard Italian. Note that as norms for the Italian population are not yet available, we conducted analysis only on raw scores.

In order to evaluate bilingual exposure and dominance, we opted for a sensitive instrument that provides a comprehensive description of bilingualism that applies to a broad range of contexts: the Language and Social Background Questionnaire (LSBQ; Anderson et al., 2018). A copy of the Italian version of the questionnaire is available for view at the following link: https://drive.google.com/file/d/1VVubXIjaLq9-Ep_KN-CsmsekzPWf04kt/view?usp=sharing. The questionnaire includes a detailed description of bilingual usage patterns across different situations. The questionnaire provides a composite factor score that represents the overall level of bilingualism. Critically, this composite score can be used as both a continuous variable or as a cut-off to discriminate groups categorically. According to the authors, students with a composite score of less than −3.13 could be classified as monolinguals (in that one language was significantly more dominant than the other), whereas participants with a composite score above 1.23 are considered

**TABLE 1 |** Summary of the different factors that were considered in order to evaluate the risk of school exclusion in our sample, and related tests that were used to measure each factor.

| Factors | Skills and constructs evaluated | Tests |
|---|---|---|
| L2 Proficiency and use | L2 Receptive vocabulary, dominance and use of L2, L2 sentence comprehension | PPVT, LSBQ, Sentence comprehension assessment |
| Learning profile | Reading fluency and comprehension | MT 3 Advanced (passage reading comprehension), DDE-2 (word and non-word reading) |
| Socioeconomic status | Family wellness | Family Affluence Scale (FAS) |
| Psychological well-being | General well-being including engagement, quality of relationships and support from peers | Comprehensive Inventory of Thriving (CIT) |

bilinguals. It is important to note that information about language dominance in such sample needs to be considered with caution. Indeed it could be possible that an L2 student could be classified as monolingual because s/he mostly uses Italian in everyday life, but it is also possible to be regarded as monolingual being more dominant in the L1. Note that the information about language dominance could be used both categorically and continuously. In the current paper we considered the score as a cut-off in order to discriminate between those L2 students that were dominant in their second language versus those that were dominant in their L1.

Regarding the learning profile, word and non-word reading fluency, and passage reading comprehension scores were obtained through the administration of the following Italian standardized tests: (1) "Test of word and non-word reading" drawn from "Batteria per la valutazione della dislessia evolutiva DDE-2" [Battery for the Assessment of Dyslexia and Developmental Dysorthography 2] (Sartori et al., 2007). Norms for high school students were based on the study of Arina et al., 2013; (2) "Passage reading comprehension test" drawn from "Prove MT Avanzate-3-clinica" ([MT 3 Advanced], Cornoldi et al., 2017), which provides accuracy scores for passage reading comprehension.

As for the third factor, SES, we administered the Family Affluence Scale (FAS; Currie et al., 2008) to identify students' SES. The questionnaire collects information about family wellness.

Psychological well-being was assessed through the Comprehensive Inventory of Thriving (CIT; Su et al., 2014), a validated self-report questionnaire. It is composed by 54 items assessing 18 facets of positive functioning that represent seven dimensions of psychological well-being. The responses to each item vary on a Likert-scale from 1 ("Strongly Disagree") to 5 ("Strongly Agree"). In addition to total scores, we also considered eighteen sub-scales: Support, Community, Trust, Respect, Loneliness, Belongingness, Flow, Skill, Learning, Lack of Control, Accomplishment, Self-efficacy, Self-worth, Meaning, Optimism, Life satisfaction, Positive emotions, and Negative emotions. Note that, even though the CIT is not directly meant to offer an indication of school exclusion, it provides a reliable and valid multi-component measure of well-being in a social context. In particular, we put the attention to the Support scale, which investigates the perception of how much others, as well as the external context, act as a support for the subject.

## Sentence Comprehension Assessment

In order to evaluate the proficiency in L2 sentence comprehension, we created a simple test aimed at evaluating the ability of teen students to comprehend complex sentences. The sentence comprehension assessment involved 20 multiple-choice items made up of a target sentence and four possible options. In order to create a set of sentences for the sentence comprehension assessment, we proceeded as follows. First, we selected several items drawn from the well known psychodiagnostic tests MMPI-RF (Ben-Porath and Tellegen, 2008; 46 items), PAI (Morey and Boggs, 1991; 21 items) e DAPP-BQ (Livesley and Jackson, 2009; 16 items). Sentences were chosen with respect to their syntactic complexity: they might involve a subject or an object relative clause (Traxler et al., 2002; Staub, 2010; Yang, 2013), a double negation within the same NP (i.e., a negative reversal structure, Chierchia, 2013), a hypothetical conditional or multiple subordinate clauses. We provide below an example of a sentence drawn from the MMPI-RF that was selected, followed by the English equivalent:

(1) Non sarei preoccupato se qualcuno dei miei famigliari si trovasse nei guai con la legge.
'It wouldn't make me nervous if any members of my family got into trouble with the law.'

In general, we selected structures that were regarded in the literature to cause slower processing (Grodner and Gibson, 2005) or difficulty in anticipating the structure (Levy, 2008).

We then asked 44 native Italian adults (age range: 19–43, 7 M), to rate the grammatical acceptability of the 83 sentences so identified on a 1 to 5 Likert scale. By doing so, we selected 20 sentences that were rated as less acceptable: 11 of them were drawn from the MMPI-RF, eight from the DAPP-BQ, one from the PAI.

To create the target sentences of the sentence comprehension test, for each sentence, we changed the lexical content, maintaining unaltered the syntactic structure. Thus, (1) was transformed into (2):

(2) Elena non sarebbe tranquilla se qualcuno dei suoi amici si trovasse in difficoltà con un esame.
'Elena could not relax if any of her friends were in trouble with an exam.'

We further developed four types of response options starting from the original target sentence. The first option was a paraphrase of the target sentence with equivalent meaning (henceforth, "equivalent", see 2a below). The second type of sentence involved a contradictory meaning with respect to the target sentence ("contradictory"; 2b); the third one presented a meaning compatible with the target sentence, but not equivalent ("compatible"; 2c); the fourth option involved an unrelated content ("distractor"; 2d). Participants were asked to identify the sentence with the same meaning of the target sentence. Correct responses therefore refer to the number of equivalent sentences. We are aware of the fact that several procedures might be developed for scoring the sentence comprehension test. In the current study, we opted for the most straightforward procedure, namely, to compute proportions of correct responses, i.e., the raw number of equivalent sentences identified. The full list of target sentences and the four response options are reported in **Supplementary Appendix A**.

(2a) Se un amico di Elena facesse fatica con un esame, lei sarebbe in pensiero per lui.
'If a friend of Elena were struggling with an exam, she would be worried about him.'

(2b) Se tra gli amici di Elena ce ne fosse qualcuno in difficoltà con un esame, a lei non importerebbe molto.
'If any of Elena's friends were in trouble with an exam, she would not care much.'

(2c) A Elena importa sapere che i suoi amici stanno bene.
'Elena cares about her friends being fine.'
(2d) Se uno si trovasse in difficoltà con un esame non potrebbe essere tranquillo.
'If one were struggling with an exam, he could not relax.'

The sentence comprehension test was presented using Google forms. The test began with a statement outlining the overall purpose of the research. After a short demographic section, participants were asked their reading habits through several short multiple-choice questions. After a couple of practice trials, the test began. Participants had no time limit, and they could go back to the previous sentence whenever they wanted. The test ended by providing the subject with the number of correct responses obtained.

## Procedure

The purpose of the study was explained to all potential participants in an assembly. Participants were provided with an information sheet and consent form to give to their parents. The signed consent form was returned to school, where it was collected. Participants also completed a consent form.

The assessment took place at different times of the school year, in individual and collective administering sessions. Some tests were presented individually in a quiet room (i.e., reading tests); others were proposed through a collective administration (i.e., syntactic assessment and text comprehension, FAS and LSBQ). All tests were administered by qualified psychologists. At the beginning of each session, participants were given the opportunity to ask any questions and to withdraw from the study if they wished. Students were tested in three collective sessions lasting approximately 30–40 min each and in one individual session. As for collective assessments, tests were proposed in the following order: SPM Raven and MT-3 during the first session; Peabody and Sentence Comprehension Test during the second; CIT, FAS, and LSBQ (only for bilingual students) in the third session. Reading fluency, tested through the DDE-2 word/non-word reading task, was individually assessed. Administration of the Sentence Comprehension test lasted about 15 min.

# RESULTS

## Comparison of L1 vs. L2 on Standardized Tests

We provide a short summary of the statistical results comparing L1 vs. L2 students on **Table 2**. Descriptive statistics (Means, M and Standard Deviations, SD) for all the standardized tests included in the study are reported on **Table 2**. We compared the performance of L1 vs. L2 students on standardized tests by means of a series of independent samples comparisons (*t*-tests). For simplicity sake, as regards to the CIT test, we report the only Scale that resulted significant, namely Support.

As shown in **Table 2**, there were some differences between the two groups regarding word reading fluency, passage comprehension, SES and psychological well-being. As for reading scores, it is important to note that, though appearing within the range of normality (recall that none of the participants were reported to school either for a learning disorder, or for a special educational need), word reading was significantly slower and less accurate in L2 students than in the L1 group. It is important to note that we did not find the same pattern in non-word reading. In such a case, L2 did not differ with respect to L1 readers.

Also passage reading comprehension scores indicated a poorer performance of the L2 students. However, again, L1 outperformed their L2 peers, though showing scores within the range of normality. Differential occurrences in the category of performance to the passage reading comprehension test (i.e., Need for immediate intervention; Attention is needed; Sufficient performance; Complete performance; see **Table 3**) attested significant across groups differences [$\chi^2(3) = 13.23$, $p < 0.004$]. Interestingly, 77% of L2 students showed an average performance, while 59% of L1 students involved an optimal performance. As regards to vocabulary skills, we did not find any difference in the raw scores of Peabody. Descriptive data indicate that L1 students were slightly better as compared to their L2 peers, however, such difference was not significant.

Scores at the FAS test differed across groups: although students were all drawn from the same school, living in an area of relative socioeconomic disadvantage, we did find a difference in terms of SES between the two groups. Finally, a significant difference was also found with respect to the Support scale in the CIT: L2 were less likely to feel supported by the school context as compared to their L1 peers. Overall, these findings provide evidence for the fact that the L2 group might be considered a sample at risk for

**TABLE 2 |** Vocabulary, reading and socioeconomic assessment (mean and standard deviations, independent sample *t*-tests, *p*-values and Cohen's *d*) for L1 and L2 students on the standardized tests.

| Assessment | | Mean | SD | *t* (42) | *p* | Cohen's *d* |
|---|---|---|---|---|---|---|
| Peabody (raw scores) | L1 | 149.59 | 20.34 | 1.476 | 0.147 | 0.445 |
| | L2 | 138.68 | 28.08 | | | |
| Passage comprehension (*z* scores) | L1 | 0.50 | 0.78 | 2.439 | 0.019* | 0.735 |
| | L2 | 0.00 | 0.56 | | | |
| Word reading accuracy (*z* scores) | L1 | 0.02 | 0.86 | 2.041 | 0.048* | 0.615 |
| | L2 | −0.52 | 0.88 | | | |
| Word reading speed (*z* scores) | L1 | 0.18 | 0.76 | 2.233 | 0.031* | 0.673 |
| | L2 | −0.41 | 0.98 | | | |
| Non-word reading accuracy (*z* scores) | L1 | 0.04 | 1.01 | 0.704 | 0.486 | 0.212 |
| | L2 | −0.17 | 0.99 | | | |
| Non-word reading speed (*z* scores) | L1 | 0.08 | 0.84 | 1.499 | 0.141 | 0.452 |
| | L2 | −0.38 | 1.16 | | | |
| FAS (log-transformed)[1] | L1 | 7.04 | 1.17 | 4.116 | 0.049* | |
| | L2 | 5.71 | 1.49 | | | |
| CIT (Support)[1] | L1 | 4.00 | 0.47 | 6.567 | 0.014* | |
| | L2 | 3.51 | 1.04 | | | |

*Note that, as regards to the FAS and the CIT (Support scale), the assumption of homogeneity of variance was not supported, we conducted a Levene's test. [1]Test of Equality of Variances (Levene's) used because of the violation of the equal variance assumption. *Significant at the 0.05 level (two-tailed).*

**TABLE 3 |** Contingency table for the performance categories to the passage reading comprehension test by group.

| | NI | AN | SP | CP | Total |
|---|---|---|---|---|---|
| L1 | 0 (0%) | 2 (9%) | 7 (32%) | 13 (59%) | 22 (100%) |
| L2 | 1 (5%) | 2 (9%) | 17 (77%) | 2 (9%) | 22 (100%) |
| Total | 1 (2, 5%) | 4 (9%) | 24 (54%) | 15 (34%) | 44 (100%) |

*Performance categories refer to NI, need for immediate intervention; AN, attention is needed; SP, sufficient performance; CP, complete performance.*

school exclusion. In fact, both socioeconomic as well as personal well-being factors seem to concur to draw an unsettling scenario for these pupils.

With respect to the LSBQ, by examining the composite score, only 10 of the L2 participants could be considered "bilingual" (composite score above 1.23). The remaining 12 students were classified as "monolingual" (showing a composite score of less than −3.13; $N = 6$) or "undifferentiated" (composite score falling between −3.13 and 1.23; $N = 6$). Note that the LSBQ provides information about language dominance in terms of patterns of preferred use of and exposure to L1 and L2 in everyday life. Therefore one should not infer anything with regard to proficiency. In general, the inspection of additional factors such as "L1 Home Use and Proficiency" and "L2 Social Use" revealed that most of the participants showed conditions of an unbalanced bilingualism that allowed us to categorize L2 "monolinguals" and "undifferentiated" as dominant in L2 (henceforth "L2-dominant"). L2 students classified as "bilinguals", although proved to use both languages in their everyday life to a greater extent that L2 "monolingual" peers, revealed a prevalent tendency to use their L1 in most social and personal activities. For this reason we defined such group as "L2-non-dominant".

## Comparison of L1 vs. L2 on the Sentence Comprehension Test

Descriptive statistics (Means and Standard Deviations, SD) of proportions of correct responses to the sentence comprehension task by group are reported in **Table 4** (first section). There were slightly more correct responses (7%) in the L1 group as compared to the L2 one.

Data analyses were conducted using R (R Core Team, 2015). Results obtained in the sentence comprehension test were analyzed with mixed-effects models, using the "lmerTest" package (Kuznetsova et al., 2017). The dependent variable was fitted to a series of mixed effects models. As accuracy to the sentence comprehension test was a binary variable we used the "binomial family". In each model, we first tested whether the fixed effects of group (L1 vs. L2) and group based on the LSBQ questionnaire categories (L1 vs. L2-non-dominant vs. L2-dominant) contributed to the model's fit. We examined whether fixed and random effects added significant information to the model by means of a series of likelihood ratio tests based on a stepwise removal procedure (e.g., Jaeger, 2008). Additionally, participants and items were included as random effects to take into account their variability in each mixed-effects model. On **Table 5** we report the coefficients in the final models thus identified, with $p$-values approximated by the normal distribution (Barr et al., 2013). Additionally, we run a series of mixed effects models on the L1 vs. L2 datasets separately, in order to test whether reading fluency, vocabulary and passage reading comprehension predicted accuracy to the sentence comprehension test to a different extent in the L1 vs. L2 groups.

First, L1 students were more accurate than L2 in the sentence comprehension test, but this difference only approached significance ($p = 0.09$). However, interestingly, when we split the L2 group according to their relative dominance in L2, the picture changed drastically. As reported on the second section of **Table 4**, the group less dominant in L2 (L2-non-dominant) were 12% less accurate than L1 students and, on average, they produced 9% less correct responses than their peers that were L2-dominant. When we added group [$\chi^2(3) = 6.45$, $p < 0.039$] to the model, we observed that the L2-non-dominant group significantly differed both from L1 and L2-dominant. Importantly, the L2-dominant group showed a performance at the comprehension test comparable to that of L1 students.

We conducted a series of linear mixed effects models, based on the L1 and L2 data, respectively, to test whether accuracy to the sentence comprehension test was predicted by reading skills (word and non-word reading fluency, vocabulary skills, text comprehension, SES, chronological age, and, only for the bilingual dataset, length of exposure to the Italian language). For simplicity sake, we will report only significant results, omitting the $\chi^2$ values and the corresponding $p$-values when the predictors

**TABLE 4 |** Mean proportions and standard deviations (SD) of correct responses to the sentence comprehension test by group (L1 vs. L2) and by group as defined by the LSBQ (L2-dominant vs. L2-non-dominant).

| | *N* | Mean | SD |
|---|---|---|---|
| L1 | 22 | 0.90 | 0.30 |
| L2 | 22 | 0.83 | 0.37 |
| L2-dominant | 12 | 0.87 | 0.33 |
| L2-non-dominant | 10 | 0.78 | 0.41 |

**TABLE 5 |** Summary of mixed effects models based on the accuracy to the sentence comprehension test.

| | β | SE | *Wald Z* | *p* |
|---|---|---|---|---|
| **Group (L1 vs. L2)** | | | | |
| Intercept | 2.91 | 0.39 | 7.39 | 0.001 |
| Group (L1 vs. L2) | −0.69 | 0.42 | −1.65 | 0.09 |
| **LSBQ Group (L1 vs. L2-dominant vs. L2-non-dominant)** | | | | |
| Intercept | 2.71 | 0.45 | 6.004 | 0.001 |
| Group (L1 vs. L2-dominant) | 0.15 | 0.48 | 0.32 | 0.75 |
| Group (L1 vs. L2-non-dominant) | −1.25 | 0.51 | −2.44 | 0.01** |
| Group (L2-dominant vs. L2-non-dominant) | −1.07 | 0.54 | −1.95 | 0.05* |

*Contrast were coded using dummy coding with L1 and L2-dominant used as reference levels. **Significant at the 0.01 level (two-tailed). *Significant at the 0.05 level (two-tailed).*

did not contribute to the fit of the model (and thus had to be excluded).

Regarding L1 data, accuracy to the sentence comprehension test was significantly predicted by word reading accuracy (Estimate = 0.084, SE = 0.370, $t$ = 2.23, $p < 0.02$) and by vocabulary (Estimate = 0.030, SE = 0.014, $t$ = 2.09, $p < 0.03$). When we considered the L2 data, the only significant predictors were non-word reading accuracy (Estimate = 0.211, SE = 0.96, $t$ = 2.18, $p < 0.03$) and accuracy to the passage comprehension test (Estimate = 0.59, SE = 0.31, $t$ = 1.89, $p < 0.05$). While it is perfectly reasonable that sentence comprehension (considered the fact that it involved reading) was significantly predicted by word reading and vocabulary, as appeared in the L1 group, it is remarkable that in the L2 dataset, sentence comprehension was predicted by non-word reading as well as by text comprehension. One might not exclude that, in this population, a task that requires reading underlies a sub-lexical decoding process, that is also involved in non-word reading. We will go back to these findings in the General Discussion.

## Comparison of L2-Non-dominant vs. L2-Dominant vs. L1 on Standardized Tests

Given the differences observed at the sentence comprehension test, we compared the performance of L2-non-dominant vs. L2-dominant vs. L1 on standardized tests by means of a series of Analysis of Variance using the Bonferroni correction. Inferential Statistics for each variable of the standardized tests are reported on **Table 6** (F, *p*-values and *post hoc* tests) and graphically presented on **Figure 1**.

Regarding reading tests, there was a significant difference in word reading accuracy and speed, with L2-non-dominant being significantly slower than L2-dominant and L1, while no difference was found between L2-dominant and L1 groups. With respect to non-word reading, L2-dominant outperformed both L1 and L2-non-dominant, being significantly faster than both groups. L1 did not differ from L2-non-dominant. No differences were found with respect to non-word reading accuracy. Regarding passage reading comprehension, L2-non-dominant were significantly less accurate, as compared to L1; L2-dominant lay between the other groups and did not differ from any of them.

Importantly, L2-non-dominant were significantly below their peers also when considering the Support scale of the CIT, while no difference was found between L2-dominant and L1. Again, this represents an evidence of the fact that, within the L2 population, there might be students showing distress in the school context, and that only by inspecting their language profile it is possible to identify vulnerable students.

Interestingly, the only test where L2-dominant patterned similarly to L2-non-dominant was the FAS. That is, the two L2 groups showed a comparable SES, that resulted significantly below their L1 peers. This finding is extremely important because it suggests that, although coming from a context of relative socioeconomic disadvantage, this subgroup of L2 students

**TABLE 6 |** Summary of the statistical analyses (Anova and *Post hoc* tests) of the assessment results by experimental group L1, L2-dominant and L2-non-dominant.

| **Test** | ***F*(1, 42)** | ***p*** | $\eta^2$ |
|---|---|---|---|
| Peabody | 1.085 | 0.347 | 0.050 |
| | ***Post hoc* Test – *t*** | ***p* Bonf** | |
| L1 vs. L2-non-dominant | 1.28 | 0.62 | |
| L1 vs. L2-dominant | 0.20 | 1.00 | |
| L2-dominant vs. L2-non-dominant | −1.11 | 0.81 | |
| | ***F*(1, 42)** | ***p*** | $\eta^2$ |
| Passage reading comprehension (*z* scores) | 3.255 | 0.049* | 0.137 |
| | ***Post hoc* Test – *t*** | ***p* Bonf** | |
| L1 vs. L2-non-dominant | 2.400 | 0.063 | |
| L1 vs. L2-dominant | −1.614 | 0.343 | |
| L2- dominant vs. L2-non-dominant | 0.785 | 1.000 | |
| | ***F*(1, 42)** | ***p*** | $\eta^2$ |
| Word reading accuracy (*z* scores) | 8.540 | < 0.001** | 0.294 |
| | ***Post hoc* Test – *t*** | ***p* Bonf** | |
| L1 vs. L2-non-dominant | 3.917 | 0.001** | |
| L1 vs. L2-dominant | −0.057 | 1.000 | |
| L2- dominant vs. L2-non-dominant | 3.441 | 0.004** | |
| | ***F*(1, 42)** | ***p*** | $\eta^2$ |
| Word reading speed (*z* scores) | 8.044 | 0.001** | 0.282 |
| | ***Post hoc* Test – *t*** | ***p* Bonf** | |
| L1 vs. L2-non-dominant | 3.885 | 0.001** | |
| L1 vs. L2-dominant | −0.351 | 1.000 | |
| L2- dominant vs. L2-non-dominant | 3.167 | 0.009** | |
| | ***F*(1, 42)** | ***p*** | $\eta^2$ |
| Non-word reading accuracy (*z* scores) | 0.495 | 0.486 | 0.012 |
| | ***Post hoc* Test – *t*** | ***p* Bonf** | |
| L1 vs. L2-non-dominant | 3.885 | 0.001** | |
| L1 vs. L2-dominant | 0.005 | 1.000 | |
| L2- dominant vs. L2-non-dominant | 1.102 | 0.830 | |
| | ***F*(1, 42)** | ***p*** | $\eta^2$ |
| Non-word reading speed (*z* scores) | 4.969 | 0.012* | 0.195 |
| | ***Post hoc* Test – *t*** | ***p* Bonf** | |
| L1 vs. L2-non-dominant | 2.932 | 0.016* | |
| L1 vs. L2-dominant | 0.119 | 1.000 | |
| L2- dominant vs. L2-non-dominant | 2.711 | 0.029* | |
| | ***F*(1, 42)** | ***p*** | $\eta^2$ |
| FAS | 5.469 | 0.008** | 0.215 |

*(Continued)*

**TABLE 6 |** Continued

| | *Post hoc* Test – *t* | *p* Bonf | |
|---|---|---|---|
| L1 vs. L2-non-dominant | 2.632 | 0.036* | |
| L1 vs. L2-dominant | −2.745 | 0.027* | |
| L2- dominant vs. L2-non-dominant | −0.022 | 1.000 | |
| | ***F*(1, 42)** | ***p*** | $\eta^2$ |
| CIT (Support scale) | 3.600 | 0.036* | 0.153 |
| | ***Post hoc* Test – *t*** | ***p* Bonf** | |
| L1 vs. L2-non-dominant | 2.680 | 0.032 | |
| L1 vs. L2-dominant | 0.735 | 1.000 | |
| L2- dominant vs. L2-non-dominant | 1.718 | 0.281 | |

***Significant at the 0.01 level (two-tailed). *Significant at the 0.05 level (two-tailed).*

showed a linguistic and learning performance, as well as a level of psychological well-being, basically overlapping with that of their L1 peers. It further indicates that the critical variable that allows to discriminate immigrant teens that might show a vulnerability in the school context is their inherent proficiency in L2.

## DISCUSSION

Up to date, the lack of empirical studies on the language and learning profile of Italian L2 teenagers has made it difficult to clarify whether this population might be effectively considered at risk for school exclusion, as national reports seem to indicate (ISTAT, 2016). Therefore, to shed light on the nature of the problems that might characterize L2 adolescents, their learning profile as well as their language ability, the present study compared a group of L2 participants with a group of L1 peers that attended the same vocational school. L2 and L1 groups were tested on a series of standardized cognitive, linguistic and reading tests, as well as on an online sentence comprehension test. The test was created to identify subjects struggling with their L2.

Moving from the assumption that being proficient in the language of instruction might represent a protective factor for L2 students, we tested whether language comprehension difficulties in L2 might allow identifying students at risk for school exclusion, in accordance with other variables (i.e., SES, learning skills, and psychological well-being) that are known to be associated with this construct (e.g., Lam et al., 2014). We will first discuss results considering the L1 and L2 groups. Then, we will focus on the specific contribution of L2 proficiency in discriminating those students who might be at risk for school exclusion.

When comparing the experimental groups L1 and L2, results of the sentence comprehension test revealed that L2 were less accurate than the other group, but the difference did not reach significance. However, the analysis provided interesting results about the linguistic underpinnings of sentence comprehension process in monolingual and bilingual students. While in the L1 group, accuracy to the sentence comprehension task was

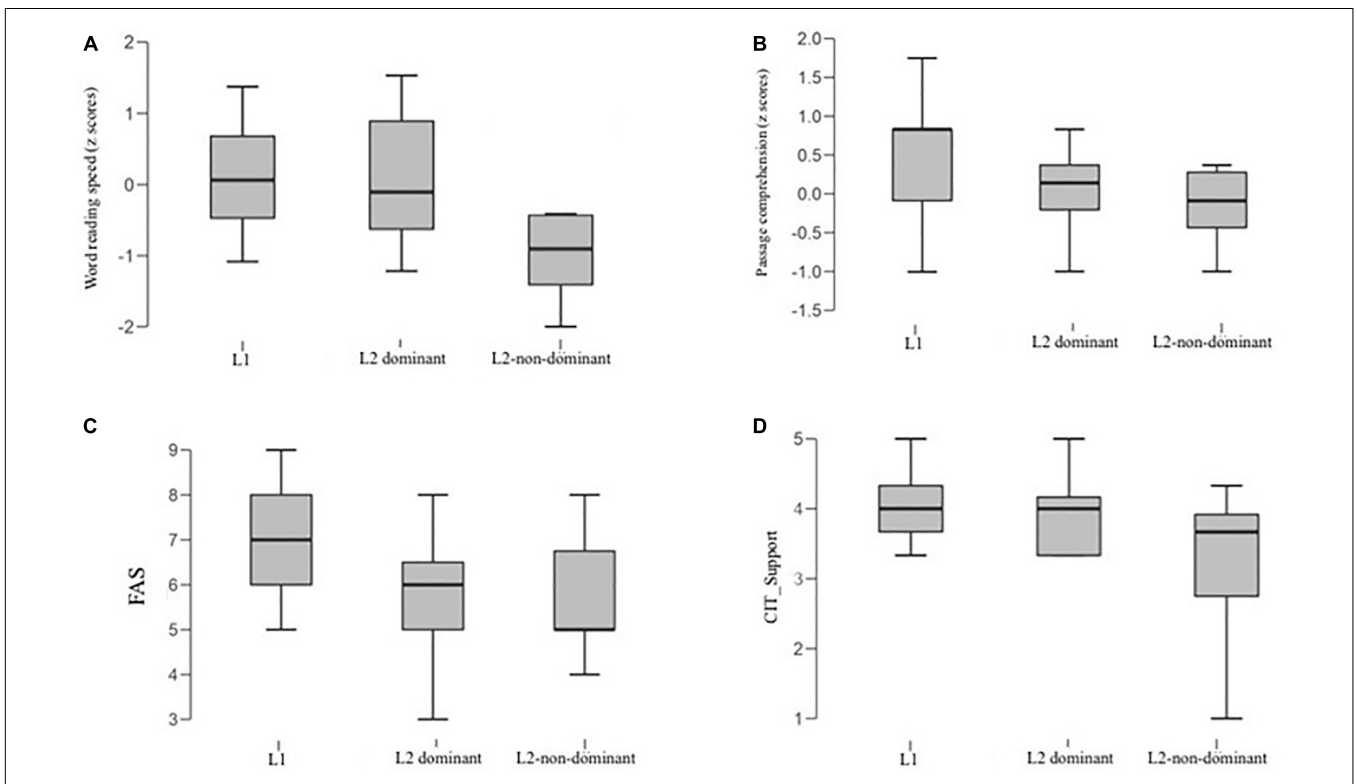


**FIGURE 1 |** The four box plots represent the distribution of **(A)** word reading speed; **(B)** Passage reading comprehension; **(C)** scores to the FAS questionnaire on socioeconomic level; **(D)** CIT, Support scale as a function of group defined with respect to language dominance (i.e., L2 dominant vs. L2-non-dominant vs. L1).

predicted by vocabulary and word reading fluency, in the L2 sample, text reading comprehension as well as non-word reading speed were the only significant predictors.

These findings suggest that L2 students, when reading a sentence, tend to activate sub-lexical decoding, as well as higher-level mechanisms underlying text comprehension processes to comprehend it. Such findings offer an important clue about sentence comprehension mechanisms at stake in bilingual development. Namely, one possibility is that bilingual readers are more likely to rely on decoding processes, instead of lexical ones, due to a reduced L2 vocabulary size (though not significant in our study). Therefore higher-level processes underlying text comprehension are exploited only later, to make inferences about the internal structure of complex sentences.

Regarding standardized tests, the two groups behaved similarly only on non-word reading fluency and vocabulary skills. Conversely, there were significant differences in word reading fluency and passage reading comprehension scores. The fact that L2 students' accuracy on word reading was significantly lower as compared to their monolingual peers, while non-word reading did not differ, confirmed that poorer reading performance in L2 students was not caused by a learning disability. Note additionally that in both groups, scores on reading tests, though significantly different, remained within the range of normality. Instead, this pattern of results indicated that bilinguals over-relied on sub-lexical mechanisms even at 15 years of age. However, sub-lexical decoding processes, while facilitating non-word reading, hampered word reading fluency.

This finding is in line with previous research based on early L2 Italian primary school children that outperformed their monolingual peers in non-word reading. In contrast, their ability in other reading and linguistic tasks appeared significantly lower with respect to monolinguals (Vernice and Pagliarini, 2018; see also Vender et al., 2016 for a comparable effect in pre-schoolers on non-word repetition). Note that it was not the aim of the current paper to disentangle whether word reading in L2 teens was based on lexical or sub-lexical mechanisms. Our results simply indicate that up to 15 years of age, bilingual students still rely on sub-lexical decoding in reading.

As for the family well-being, even though we tried to keep the SES as comparable as possible within the two groups, the L2 lived in a condition of socioeconomic disadvantage as compared to their monolingual peers. Additionally, when tested on a measure of psychological well-being in the classroom, L2 students were less likely to feel supported from the external environment.

We now discuss the contribution of L2 proficiency and use in order to better characterize the profile of students who could be at potential risk for school exclusion. Using a questionnaire aimed at testing language use in daily life contexts, we were able to identify two subgroups of L2 participants. The first one showed a higher use and exposure to Italian as compared to their L1 (L2-dominant group); the second one, in contrast, was characterized by use and exposure to both languages in different contexts (L2-non-dominant group). The L2-dominant group appeared to perform comparably to the L1 group in all standardized tests as well as on the sentence comprehension test. Conversely, the L2-non-dominant students proved to be significantly below their monolingual and bilingual (L2-dominant) peers in most of the standardized tests as well as on the sentence comprehension test.

As regards to psychological well-being, the CIT- Scale Support indicated that the L2-non-dominant group was significantly less likely to feel supported by the external context in comparison with L1, while L2-dominant did not differ from the other two groups. The L2-dominant group showed comparable scores with respect to L2-non-dominant group only regarding the SES. To sum up, the L2-non-dominant group appeared to be not only more impaired on most of learning and linguistic tasks in comparison with L1 and L2-dominant groups, but was also more vulnerable in terms of psychological well-being. Crucially, the co-occurrence of language and learning difficulties, together with the perception of scarce support within the external context, suggests that L2-non-dominant students are potential candidates for being at risk for school exclusion.

The striking difference we found on L2 participants based on proficiency, dominance and use of the L2, provides significant clues about the role of language dominance in determining the (linguistic and learning) profile of the bilingual speaker. Namely, we observed that language dominance exerts a critical effect in defining the linguistic and academic skills of the L2 student, and this finding is in line with previous studies indicating that dominance in a language, but not necessarily proficiency, determine the linguistic outcome in bilingual development (Perpiñán, 2017). Additionally, such finding reveals that, in line with our initial hypothesis, proficiency and use of the language of instruction represent a crucial asset in promoting effective inclusion in the classroom, exerting important implications for educational programs in high schools. Our data suggest that teachers and educators should be challenged to promote activities that enhance L2 proficiency targeting not only late bilinguals or newly arrived immigrants, but also L2 students who were born in Italy, showing a limited use of the majority language.

We are conscious of the fact that our L2-dominant speakers might be regarded as Heritage Language Speakers (henceforth, HLS). HLS are L2 speakers who have acquired their L1 simultaneously with the majority language (L2) or exclusively the L1 if the immigration has occurred during infancy (Benmamoun et al., 2013). They might use a minority language (their L1) in the home environment, though showing a predominant use and exposure in the daily life context to their L2 (Valdés, 2005). Research on bilingual speakers of minority languages has further shown that HLS might fail to acquire full linguistic competence in the heritage language, i.e., their L1 (Polinsky and Kagan, 2007). As a consequence, in adulthood, linguistic outcomes highlight atypical acquisition patterns (Kupisch and Rothman, 2016). Importantly, in our study, we observed that L2-dominant participants showed a distinctive profile as compared to their bilingual and monolingual peers. These students tended to use their L2, Italian, to a greater extent in daily life, as compared to their L2-non-dominant peers. However, given the fact that, in order to properly define HLS, one needs to collect information about the real language proficiency in L1, we might not be sure that L2-dominant can be regarded as HLS. For this reason, we safely categorized the groups based on use and exposure to L2 language.

With respect to the current study, it is important to discuss a number of potential limitations. First, finding a homogenous sample of participants in terms of first age and length of exposure to L2, implied trimming a consistent number of L2 students that attended the vocational high school involved in the study. As a consequence, the current research, given the small sample of participants, offers only an initial contribution to the study of the relation between weakness in the L2 and risk for school exclusion. We are conscious that our results do not allow to draw conclusions on the population of teens of immigrant families in Italy. Thus, in future work, we aim to extend our sample size, ideally involving a consistent number of participants with diverse L1 backgrounds. By doing so, it would be possible to investigate, for instance, gender and cultural differences between subgroups.

A second aspect that we were not able to take into account in the current research, but that is deeply intertwined with the first point discussed above, refers to the inclusion of a measure of academic achievement. We believe that this variable could provide additional information about cross-cultural differences in academic attainment and motivation to learn (Bempechat and Drago-Severson, 1999). Additionally, an objective measure of school achievement would clarify whether the sentence comprehension test positively correlates with school outcomes, at least in disciplines where reading comprehension is crucial. Given the importance of this point, it would be worth further investigation.

Third, regarding the sentence comprehension test, we deliberately opted for a simple scoring procedure. We are aware of the fact that there exist alternative ways to score the current data. For instance, one could assign intermediate scores to compatible sentences or penalize a contradictory response with a negative score. Alternatively, it would be even possible to transform raw scores according to the Item Response Theory approach (IRT; Hambleton and Jones, 1993). By doing so, the test would allow to assess the ability of each student in the sentence comprehension test with respect to the difficulty of the individual test items. However, since the current study represents only an initial phase of our research, we preferred to follow the simplest scoring procedure. Recall that the current test is not meant to offer a diagnostically sophisticated measure of linguistic competence, rather to provide a quick screening task to be used by teachers and educators in schools. It is possible that, in future steps of our research, we might develop a more complex scoring procedure for the sentence comprehension test.

## CONCLUSION

In conclusion, we observed that, in a sample of early bilinguals, the relative disadvantage of some L2 students at the linguistic and academic level was large. The linguistic problems were importantly captured by a rapid but efficient test of sentence comprehension, that we created on purpose. The sentence comprehension test exerted several advantages: first, it could be carried out by teachers or educators following simple instructions, in contrast to tests of reading fluency and comprehension that must be administered by health professionals. Second, the possibility to use it in schools as an on-line tool would allow to address the urgent need of teachers to adequately but quickly identifying those students that are struggling with sentence comprehension.

To sum up, our study indicates that an intervention aimed at targeting the sentence comprehension ability may help teachers and educators to assess the language skills of students. In fact, language proficiency is of greatest importance for personal and academic realization, and, as we observed in the current study, it represents a critical variable in signaling students who might manifest a vulnerability within the school context. By using a simple online tool, such as the sentence comprehension test presented here, readily available to educators, teachers might offer adequate educational opportunities to all the students, promoting cohesion in the classroom.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Catholic University of the Sacred Heart. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

MV: creation of the "Sentence Comprehension Test" with MM conception and design of the work, data collection, data analysis, interpretation, and drafting the manuscript. MM: conception and creation of the "Sentence Comprehension Test". MT: data collection and coding, and drafting of the section Materials and Methods. MC and EL: data collection and coding. DS: conception of the design and critical revision of the draft. MG and ML: conception of the "Sentence Comprehension Test". All the authors contributed to the critical revision of the manuscript.

## FUNDING

MV, MT, MC, and EL were supported by a grant from the CARIPLO Foundation [Grant 2017-NAZ-0131, Tecnologie Innovative per il Benessere e l'Integrazione Scolastica (IBIS) (New technologies for education and their impact on students' well-being and inclusion)].

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02417/full#supplementary-material

## REFERENCES


Anderson, J. A., Mak, L., Keyvani Chahi, A. K., and Bialystok, E. (2018). The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behav. Res. Methods* 50, 250–263. doi: 10.3758/s13428-017-0867-9

Andolfi, V. R., Confalonieri, E., Nardò, F., and Traficante, D. (2015). Star bene a scuola: il ruolo dell'abilità di letto-scrittura. *Psicol. Dell'educazione* 2, 133–152.

Arina, S., Iervolino, I., and Stella, G. (2013). Prima raccolta di dati normativi per la valutazione della dislessia evolutiva negli adolescenti su un campione di scuola secondaria di secondo grado. *Dislessia* 10, 9–38.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001

Bempechat, J., and Drago-Severson, E. (1999). Cross-national differences in academic achievement: beyond etic conceptions of children's understandings. *Rev. Educ. Res.* 69, 287–314. doi: 10.2307/1170541

Benmamoun, E., Montrul, S., and Polinsky, M. (2013). Heritage languages and their speakers: opportunities and challenges for linguistics. *Theor. Linguist.* 39, 129–181. doi: 10.1515/tl-2013-0009

Ben-Porath, Y. S., and Tellegen, A. (2008). *MMPI-2-RF, Minnesota Multiphasic Personality Inventory-2 Restructured Form: Manual for Administration, Scoring and Interpretation*. Minnesota: University of Minnesota Press.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* 70, 636–644. doi: 10.1111/1467-8624.00046

Bialystok, E., and Feng, X. (2009). Language proficiency and executive control in proactive interference: evidence from monolingual and bilingual children and adults. *Brain Lang.* 109, 93–100. doi: 10.1016/j.bandl.2008.09.001

Bialystok, E., and Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev. Sci.* 7, 325–339. doi: 10.1111/j.1467-7687.2004.00351.x

Brizzolara, D., Gasperini, F., Pfanner, L., Cristofani, P., Casalini, C., and Chilosi, A. M. (2011). Long-term reading and spelling outcome in Italian adolescents with a history of specific language impairment. *Cortex* 47, 955–973. doi: 10.1016/j.cortex.2011.02.009

Brunello, G., and Rocco, L. (2013). The effect of immigration on the school performance of natives: cross country evidence using PISA test scores. *Econ. Educ. Rev.* 32, 234–246. doi: 10.1016/j.econedurev.2012.10.006

Chierchia, G. (2013). *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford, UK: Oxford University Press.

Clegg, J., Stackhouse, J., Finch, K., Murphy, C., and Nicholls, S. (2009). Language abilities of secondary age pupils at risk of school exclusion: a preliminary report. *Child Lang. Teaching Ther.* 25, 123–139. doi: 10.1177/0265659008098664

Contini, D. (2013). Immigrant background peer effects in Italian schools. *Soc. Sci. Res.* 42, 1122–1142. doi: 10.1016/j.ssresearch.2013.02.003

Cornoldi, C., Baldi, A. P., and Giofrè, D. (2017). *Prove MT Avanzate-3-Clinica*. Firenze: Giunti OS.

Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., and Richter, M. (2008). Researching health inequalities in adolescents: the development of the Health Behaviour in School-Aged Children (HBSC) family affluence scale. *Soc. Sci. Med.* 66, 1429–1436. doi: 10.1016/j.socscimed.2007.11.024

Deci, E. L., and Ryan, R. M. (2008). Self-determination theory: a macrotheory of human motivation, development, and health. *Can. Psychol.* 49, 182–185. doi: 10.1037/a0012801

Dunn, L., and Dunn, L. (1981). *Peabody Picture Vocabulary Test–Revised: Manual for forms L and M*. Circle Pines, MN: American Guidance Service.

Flanagan, D. P., Ortiz, S. O., and Alfonso, V. C. (2013). *Essentials of Cross-Battery Assessment*, 3rd Edn, New York, NY: Wiley.

Flege, J. E., Yeni-Komshian, G., and Liu, H. (1999). Age constraints on second-language acquisition. *J. Mem. Lang.* 41, 78–104. doi: 10.1006/jmla.1999.2638

Galindo, C., and Reardon, S. F. (2006). *Hispanic Students' Educational Experiences and Opportunities During Kindergarten Report to the National Task Force on Early Childhood Education for Hispanics*. Tempe, AZ: Arizona State University.

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10. doi: 10.1177/074193258600700104

Grodner, D., and Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cogn. Sci.* 29, 261–290. doi: 10.1207/s15516709cog0000_7

Hambleton, R. K., and Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educ. Measur.* 12, 38–47. doi: 10.1111/j.1745-3992.1993.tb00543.x

Horn, J. L., and Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *J. Educ. Psychol.* 57:253. doi: 10.1037/h0023816

ISTAT (2016). *L'integrazione Scolastica e Sociale Delle Seconde Generazioni. Anno 2015*. Available at: https://www.istat.it/it/files//2016/03/Integrazione-scolastica-stranieri.pdf (accessed May 22, 2019).

Italian Ministry of Education [MIUR] (2014). *Alunni Con Cittadinanza Non Italiana. La Scuola multiculturale nei contesti locali. Rapporto Nazionale A.s. 2014/2015*. Available at: http://www.istruzione.it/allegati/2016/Rapporto-Miur-Ismu-2014_15.pdf (accessed May 22, 2019).

Italian Ministry of Education [MIUR] (2015a). *Diversi Da Chi? Online*. Available at: http://www.istruzione.it/allegati/2015/cs090915_all2.docx (accessed May 22, 2019).

Italian Ministry of Education [MIUR] (2015b). *Piano Nazionale Scuola Digitale*. Available at: https://www.miur.gov.it/documents/20182/50615/Piano+nazionale+scuola+digitale.pdf/5b1a7e34-b678-40c5-8d26-e7b646708d70?version=1.1&t=1496170125686 (accessed May 22, 2019).

Italian Ministry of Education [MIUR] (2017). *Gli Alunni Con Cittadinanza Non Italiana*. Available at: https://www.miur.gov.it/documents/20182/0/FOCUS+16-17_Studenti+non+italiani/be4e2dc4-d81d-4621-9e5a-848f1f8609b3?version=1.0 (accessed May 22, 2019).

Jaeger, T. F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *J. Mem. Lang.* 59, 434–446. doi: 10.1016/j.jml.2007.11.007

Kamhi, A. G., and Catts, H. W. (2012). "Language and reading: convergences and divergences," in *Language and reading disabilities*, eds A. G. Kamhi, and H. W. Catts, (New York, NY: Pearson).

Kieffer, M. J., and Box, C. D. (2013). Derivational morphological awareness, academic vocabulary, and reading comprehension in linguistically diverse sixth graders. *Learn. Individ. Differ.* 24, 168–175. doi: 10.1016/j.lindif.2012.12.017

Kovelman, I., Baker, S. A., and Petitto, L. A. (2008). Bilingual and monolingual brains compared: a functional magnetic resonance imaging investigation of syntactic processing and a possible "neural signature" of bilingualism. *J. Cogn. Neurosci.* 20, 153–169. doi: 10.1162/jocn.2008.20.1.153

Kupisch, T., and Rothman, J. (2016). "Interfaces in the acquisition of syntax," in *Handbook on Interfaces*, eds S. Fischer, and C. Gabriel, (Berlin: De Gruyter).

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Statist. Softw.* 82, 1–26.

Kvist, A. V., and Gustafsson, J.-E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: a test of Cattell's investment theory. *Intelligence* 36, 422–436. doi: 10.1016/j.intell.2007.08.004

Lam, S., Jimerson, S., Wong, B. P. H., Kikas, E., Shin, H., Veiga, F. H., et al. (2014). Understanding and measuring student engagement in school: the results of an international study from 12 countries. *Sch. Psychol. Q.* 29, 213–232. doi: 10.1037/spq0000057

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition* 106, 1126–1177. doi: 10.1016/j.cognition.2007.05.006

Livesley, W. J., and Jackson, D. (2009). *Manual for the Dimensional Assessment of Personality Pathology–Basic Questionnaire*. Port Huron, MI: Sigma.

Luk, G., Bialystok, E., Craik, F. I., and Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *J. Neurosci.* 31, 16808–16813. doi: 10.1523/JNEUROSCI.4563-11.2011

Menken, K. (2008). *English Learners Left Behind: Standardized Testing as Language Policy*. Bristol: Multilingual Matters.

Montrul, S. (2015). *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press.

Montrul, S. A. (2008). *Incomplete Acquisition in Bilingualism: Re-Examining the Age Factor*. Amsterdam: John Benjamins Publishing.

Morey, L. C., and Boggs, C. (1991). *Personality Assessment Inventory (PAI)*. Lutz, FL: PAR.

Murineddu, M., Duca, V., and Cornoldi, C. (2006). Difficoltà di apprendimento scolastico degli studenti stranieri. *Difficoltà di Apprendimento* 12, 1–15.

Mussino, E., and Strozza, S. (2012). The delayed school progress of the children of immigrants in lower-secondary education in Italy. *J. Ethn. Migr. Stud.* 38, 41–57. doi: 10.1080/1369183X.2012.640014

Myers, L., and Botting, N. (2008). Literacy in the mainstream inner-city school: its relationship to spoken language. *Child Lang. Teaching Ther.* 24, 95–114. doi: 10.1177/0265659007084570

Paradis, J., Genesee, F., and Crago, M. (2011). *Dual language development and disorders: A handbook on bilingualism and second language learning*, 2nd Edn, Baltimore, MD: Brookes.

Perpiñán, S. (2017). Catalan-Spanish bilingualism continuum. *Linguist. Approaches Biling.* 7, 477–513. doi: 10.1075/lab.15004.per

Poarch, G. J., and Bialystok, E. (2015). Bilingualism as a model for multitasking. *Dev. Rev.* 35, 113–124. doi: 10.1016/j.dr.2014.12.003

Polinsky, M., and Kagan, O. (2007). Heritage languages: in the 'wild' and in the classroom. *Lang. Linguist. Compass* 1, 368–395. doi: 10.1111/j.1749-818x.2007.00022.x

Puig-Mayenco, E., Cunnings, I., Bayram, F., Miller, D., Tubau, S., and Rothman, J. (2018). Language dominance affects bilingual performance and processing outcomes in adulthood. *Front. Psychol.* 9:1199. doi: 10.3389/fpsyg.2018.01199

R Core Team, (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Raven, J. C., and Court, J. H. (1998). *Raven's Progressive Matrices and Vocabulary Scales*. Oxford: Oxford psychologists Press.

Robinson, M. G., and Sorace, A. (2019). The influence of collaborative language learning on cognitive control in unbalanced multilingual migrant children. *Eur. J. Psychol. Educ.* 34, 255–272. doi: 10.1007/s10212-018-0377-x

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037/0003-066X.55.1.68

Sartori, G., Job, R., and Tressoldi, P. (2007). *DDE-2: Giunti OS Organizzazioni Speciali: Batteria Per la Valutazione Della Dislessia e Della Disortografia Evolutiva-2: Manuale*. Giunti OS: Organizzazioni Speciali.

Spencer, S., Clegg, J., and Stackhouse, J. (2012). Language and disadvantage: a comparison of the language abilities of adolescents from two different socioeconomic areas. *Int. J. Lang. Commun. Disord.* 47, 274–284. doi: 10.1111/j.1460-6984.2011.00104.x

Spencer, S., Clegg, J., Stackhouse, J., and Rush, R. (2017). Contribution of spoken language and socio-economic background to adolescents' educational achievement at age 16 years. *Int. J. Lang. Commun. Disord.* 52, 184–196. doi: 10.1111/1460-6984.12264

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition* 116, 71–86. doi: 10.1016/j.cognition.2010.04.002

Su, R., Tay, L., and Diener, E. (2014). The development and validation of the Comprehensive Inventory of Thriving (CIT) and the Brief Inventory of Thriving (BIT). *Appl. Psychol.* 6, 251–279. doi: 10.1111/aphw.12027

Traxler, M. J., Morris, R. K., and Seely, R. E. (2002). Processing subject and object relative clauses: evidence from eye movements. *J. Mem. Lang.* 47, 69–90. doi: 10.1006/jmla.2001.2836

Unsworth, S. (2014). "Comparing the role of input in bilingual acquisition across domains," in *Input and Experience in Bilingual Development*, eds T. Grüter, and J. Paradis (Amsterdam: John Benjamins), 181–201. doi: 10.1075/tilar.13.10uns

Unsworth, S. (2016). "Quantity and quality of language input in bilingual language development," in *Lifespan Perspectives on Bilingualism*, eds E. Nicoladis, and S. Montanari (Berlin: de Gruyter), 136–196. doi: 10.1037/14939-007

Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: opportunities lost or seized? *Mod. Lang. J.* 89, 410–426. doi: 10.1111/j.1540-4781.2005.00314.x

van Gelderen, A., Schoonen, R., Stoel, R. D., de Glopper, K., and Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: a longitudinal analysis of constituent components. *J. Educ. Psychol.* 99, 477–491. doi: 10.1037/0022-0663.99.3.477

Vender, M., Garraffa, M., Sorace, A., and Guasti, M. T. (2016). How early L2 children perform on Italian clinical markers of SLI: a study of clitic production and nonword repetition. *Clin. Linguist. Phon.* 30, 150–169. doi: 10.3109/02699206.2015.1120346

Vernice, M., and Pagliarini, E. (2018). Is morphological awareness a relevant predictor of reading fluency and comprehension? new evidence from italian monolingual and arabic-italian bilingual children. *Front. Commun.* 3:11. doi: 10.3389/fcomm.2018.00011

Whiteside, K. E., Gooch, D., and Norbury, C. F. (2017). English language proficiency and early school attainment among children learning english as an additional language. *Child Dev.* 88, 812–827. doi: 10.1111/cdev.12615

Yang, J. (2013). Mobile assisted language learning: a review of the recent applications of emerging mobile technologies. *Engl. Lang. Teaching* 6, 19–25. doi: 10.5539/elt.v6n7p19

Zhang, P. (2008). Technical opinion motivational affordances: reasons for ICT design and use. *Commun. ACM* 51, 145–147.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 November 2019
doi: 10.3389/fpsyg.2019.02369

# Tell Me a Story: Socio-Emotional Functioning, Well-Being and Problematic Smartphone Use in Adolescents With Specific Learning Disabilities

*Daniela Sarti[1]*, Roberta Bettoni[1,2], Ilaria Offredi[1], Marta Tironi[2], Elisabetta Lombardi[3], Daniela Traficante[3] and Maria Luisa Lorusso[4]*

*[1] Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy, [2] Department of Psychology, University of Milano-Bicocca, Milan, Italy, [3] Department of Psychology, Catholic University of the Sacred Heart, Milan, Italy, [4] Istituto IRCCS Eugenio Medea, Lecco, Italy*

OPEN ACCESS

**Edited by:**
*Claudio Longobardi, University of Turin, Italy*

**Reviewed by:**
*Roger C. Ho, National University of Singapore, Singapore*
*Samuel P. L. Veissière, McGill University, Canada*

***Correspondence:**
*Daniela Sarti*
*daniela.sarti@istituto-besta.it;*
*danielamaria.sarti@gmail.com*

**Specialty section:**
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

**Received:** *30 June 2019*
**Accepted:** *04 October 2019*
**Published:** *06 November 2019*

**Citation:**
*Sarti D, Bettoni R, Offredi I, Tironi M, Lombardi E, Traficante D and Lorusso ML (2019) Tell Me a Story: Socio-Emotional Functioning, Well-Being and Problematic Smartphone Use in Adolescents With Specific Learning Disabilities. Front. Psychol. 10:2369. doi: 10.3389/fpsyg.2019.02369*

Although Specific Learning Disabilities (SLD) are described as specific difficulties in one or more academic areas, often socio-emotional problems are also reported to be related to well-being and school engagement. Moreover, recent evidence shows that emotional problems and reduced social support predict problematic use of new technologies, such as a smartphone, that can, in turn, increase these problems. In this study, we aimed to investigate socio-emotional functioning and its relation to well-being, school engagement, and problematic smartphone use. Social and emotional skills of 19 adolescents with a diagnosis of SLD and 19 control adolescents were assessed through a narrative test; adolescents were requested to narrate complete stories elicited by pictures representing social situations. Information on well-being and problematic smartphone use were collected through questionnaires. The comparison between groups showed differences in cognitive-social skills, although no significant differences in emotional functioning were found. Additionally, the perception of the social environment as supportive and trustworthy was related to general well-being for both groups, whereas the perception of limits and rules set by the adult world appeared to be related to a decreased investment in learning processes only for the SLD students. Finally, correlation analysis showed that smartphone use was associated with reduced perception of social support and to a decreased ability to understand and solve social situations. These results assert the critical role played by social information processing and social support in terms of well-being in adolescence, and contribute to enhancing knowledge of the mechanisms underlying problematic smartphone use in a clinical sample.

**Keywords: socio-emotional functioning, well-being, smartphone addiction, SLD, adolescence, narratives**

## INTRODUCTION

In the last 20 years, the use of technology is constantly increasing and has become very common, especially among adolescents (De Pasquale et al., 2017) who are progressively encouraged to use technological devices for learning purposes. Nevertheless, the use of technology needs to be adapted to the increasing heterogeneity of students, who have different cultural and linguistic backgrounds and different neuropsychological profiles (Lawless and Pellegrino, 2007; Perelmutter et al., 2017). In Italy, 2.5–3.5% of students have been diagnosed with Specific Learning Disabilities (SLD), such as dyslexia, dyscalculia, dysgraphia, and dysorthography (MIUR, 2015). SLD are neurodevelopmental disorders characterized by difficulties in specific academic areas despite average intelligence and adequate educational and socio-cultural opportunities (American Psychiatric Association [APA], 2013). In addition, it has been reported that students with SLD often exhibit social-emotional problems (i.e., Mugnaini et al., 2009; Francis et al., 2018), such as an increased risk of developing externalizing and internalizing problems, loneliness, and poor self−esteem (Bryan, 2005; Wilson et al., 2009; Aro et al., 2019). Social competence also seems to be a challenge for many SLD students, especially during adolescence when social functioning becomes increasingly complex and multifaceted (Cosden et al., 2002; Mugnaini et al., 2009). For example, it has been reported that SLD students experience low-quality relationships with their peers, higher rejection, and lower acceptance (Wiener and Schneider, 2002; Al-Yagon and Margalit, 2013), they are less likely to report a secure attachment with their parents, and they have difficulties in perceiving teachers as a secure base (Al-Yagon, 2012). Some researchers suggest that these difficulties in social competence may be a result of the social stigma frequently attached to SLD (Chan et al., 2017) and, in general, to other populations affected by neurodevelopmental disorders (i.e., intellectual disabilities; Ali et al., 2012). Specifically, due to their poor academic performance, students with SLD are often perceived as being less attractive, less successful, and lazier than their peers and, therefore, they may tend to believe and internalize this biased perspective about themselves. This can, in turn, affect the way they behave in social contexts (e.g., Livingston et al., 2018). Some researchers have also examined the mental processes that underlie the understanding of social interactions, investigating how social cues are encoded, interpreted, and how potential responses are generated (Crick and Dodge, 1994). Recent studies have focused on these aspects of children with and without SLD. Using social vignettes described by the examiner, participants are asked to answer a series of questions about different aspects of social information processes; that is, to recognize the problem in the vignette, to clarify goals, to make a response decision, and to evaluate given alternative solutions. SLD students have shown poorer ability to encode social cues, to produce alternative solutions, and to select the most feasible competent one (Bauminger and Kimhi-Kind, 2008; Bauminger-Zviely et al., 2019).

Social competence and perceived support are particularly important because they are related to academic success and engagement (Welsh et al., 2001; Estell and Perdue, 2013). Moreover, acceptance by peers is an essential part of adolescent self-identity and has a strong influence on psychological well-being (La Greca and Harrison, 2005). Difficulties in this area of functioning compounded by a struggle to achieve adequate school results could negatively influence school commitment and general well-being. Unfortunately, these aspects are still scarcely investigated with adolescents.

When discussing adolescents and peer relationships, it is necessary to consider the use of smartphones since they have become an essential part of student lives. Smartphone applications have been proven to be useful for youth in various aspects, including health-related smartphone applications (Do et al., 2018) and beauty and nutrition counseling (Tran et al., 2018). However, the majority of smartphone use is spent on social activities (e.g., using Instagram, Facebook, and Snapchat), especially among Italian adolescents (ISTAT, 2018). This seems to be grounded in a deep evolutionary need to be connected with others and suggests a prosocial nature of smartphone usage (Veissière and Stendel, 2018) which can, in turn, have a positive effect on subjective well-being (Teppers et al., 2014). However, this depends on the amount of time spent on mobile phones and social networks (Twenge, 2019). Increased use of this device has been connected with the perception of diminished quality relationships with classmates (Wang et al., 2017), lower satisfaction with life (Lepp et al., 2014) and poor sleep quality (Zhang et al., 2017). Recent studies have shown that emotional stress and low academic performance seem to mediate the relationship between excessive smartphone use and well-being (Samaha and Hawi, 2016; Hawi and Samaha, 2017). Lack of social support also seems to play an important role in predicting problematic smartphone behaviors, and this relationship has negative consequences on well-being (Herrero et al., 2019). Students that perceive a low level of social support seem to engage more often in smartphone activities that can satisfy their necessity of social connections. In addition, the abuse of smartphones negatively affects the perception of social support over time. A low level of social support together with smartphone addiction can lead to higher levels of psychological distress (Van Deursen et al., 2015; Kim, 2017; Wang et al., 2017; Herrero et al., 2019). In this context, it is possible to hypothesize that students with SLD represent a population vulnerable to the problematic use of a smartphone, as they have reported more internalizing and externalizing problems and more difficulties in reaching satisfying, supportive relationships (Francis et al., 2018; Bauminger-Zviely et al., 2019).

Research on this topic among the SLD Italian population is still scarce, especially about how social support, emotional distress, well-being, and addiction are related to each other. Using a performance-based test, we aimed to investigate whether SLD Italian students are less skillful in understanding and solving social situations, manifest more emotional problems, and perceive less social support than the control group. Moreover, considering the brief review of scientific literature aforementioned, we hypothesized that general well-being and school engagement of both SLD students and the control group could be related to emotional and socio-cognitive functioning. Finally, since extensive use of smartphones could be related to

emotional and socio-cognitive functioning for both SLD and the control group (e.g., Herrero et al., 2019), we explored if SLD represents an at-risk population for smartphone addiction.

To empirically verify these hypotheses, high school students with and without a diagnosis of SLD were tested. Narrative test Roberts-2 (Roberts and Gruber, 2005; for Italian validation: Parolin et al., 2019) was used to provide an index of social cognitive understanding and emotional functioning, while general well-being and school engagement were assessed using self-report measures ("Comprehensive Inventory of Thriving-CIT," Su et al., 2014; "Student Engagement Scale," Mameli and Passini, 2017). Usually the perception and understanding of everyday interpersonal situations and their intrapersonal implications are assessed through self-report instruments or interviews, responses to which can be affected by misrepresentation and poor self-awareness (Hopwood and Bornstein, 2014). In addition to self-report measures, we used a performance-based test to better control for response bias and to maximize the imprint of individuality, since the students were asked to produce their own stories. Given the growing awareness of smartphone impact on social functioning and the heightened vulnerability to addiction during adolescence, we also included a measure of smartphone use ("Smartphone Addiction Scale," Kwon et al., 2013; for Italian standardization: De Pasquale et al., 2017) in order to examine its link with socio-emotional functioning. Although the literature on this topic is still scarce, a greater understanding of adolescents' socio-emotional functioning and its relation to well-being and smartphone use can be of paramount importance for adolescents with SLD because these aspects may operate as risk factors for them.

# MATERIALS AND METHODS

## Participants

The sample is composed of 19 adolescents (male = 5) with a diagnosis of SLD group and with a mean age of 15.16 years (*SD* = 0.36) and 19 typically developing adolescents (CNT group; male = 9) with a mean age of 15.42 years (*SD* = 0.77). All participants were recruited from three high schools located near Milan that were part of the project "New technologies for education and their impact on students' well-being and inclusion." We included students from Scientific Lyceum (21%), Technical School (39.5%), and Vocational School (39.5%). Gender distribution was different for these curricula ($\chi^2 = 19.16$; $p < 0.001$), with girls preferring the Technical Schools (54%) and boys preferring Scientific Lyceum (57%).

All participants voluntarily took part in the study after parents gave their informed consent. The study was approved by the Ethics Committee of Catholic University of the Sacred Heart, according to standards of the Helsinki Declaration (World Medical Association, 2013). Adolescents included in the study met the following inclusion criteria: they were fluent Italian speakers and they obtained an intelligence score within the normal range on a standardized intelligence test (IQ > 25 centiles; Raven's Colored Matrices test, Raven, 2000). All participants with SLD had received a clinical diagnosis based on standard inclusion and exclusion criteria (DSM-V; American Psychiatric Association [APA], 2013). Five students had received a diagnosis of dyslexia, five of dyscalculia, one of dysgraphia, and seven of mixed disabled scholastic skills, and for all of them, Individualized Education Programs were applied in the school context. Comorbidity with other psychopathological conditions were excluded. In order to be included in the CNT group, adolescents had to obtain a score within the normal range (>−1.5 SD) in reading speed, accuracy, passage reading comprehension, and dictation scores, which were obtained through the administration of the following Italian standardized tests: (1) Test of Word and non-word reading tests, drawn from "Batteria per la valutazione della Dislessia e della Disortografia Evolutiva-2, DDE-2" (Battery for the Evaluation of Dyslexia and Evolutionary Disorthography-2; Sartori and Job, 2007). Norms for high school students were based on a study by Arina et al. (2013). (2) Passage reading comprehension test and dictation test, drawn from Prove MT Avanzate-3-Clinica (MT 3 Advanced; Cornoldi et al., 2017). Between SLD and CNT groups, there were no group differences in age: SLD group, M = 15.42, SD = 0.77; CNT group, $M = 15.16$, $SD = 0.36$; $t(36) = 1.34$; $p = 0.188$.

Participant characteristics compared with a series of *t*-tests are shown in **Table 1**. Reflecting the recruitment criteria, the SLD group differed significantly from the CNT group in all reading and orthography scores.

## Procedure

The study took place during school hours. For each participant, the session lasted about 2 hours. The battery comprised the administration in Italian of self-report and performance-based tests: two questionnaires were used to measure well-being both

**TABLE 1 |** Participants characteristics that reflected the recruitment criteria.

| Test | CNT group (*n* = 19) | LD group (*n* = 19) | *t*-test | *P* | Cohen's *d* |
|---|---|---|---|---|---|
| Word reading test accuracy | 0.36 (0.56) | −1.43 (2.26) | 3.34 | 0.002 | 1.11 |
| Word reading test speed | 0.24 (0.56) | −1.03 (1.20) | 4.18 | >0.001 | 1.40 |
| Non-word reading test accuracy | 0.37 (0.76) | −0.78 (1.45) | 3.07 | 0.004 | 1.02 |
| Non-word reading test speed | 0.18 (0.58) | −1.02 (0.92) | 4.82 | >0.001 | 1.61 |
| Reading Comprehension test | 0.56 (0.78) | 0.03 (0.86) | 1.99 | 0.054 | 0.66 |
| Orthography test accuracy | 0.56 (0.56) | −0.26 (1.11) | 2.89 | 0.006 | 0.96 |

*The two groups differed significantly in all reading scores and in the orthography test.*

from a holistic point of view CIT(CIT; Su et al., 2014) and in the school context (Student Agentic Engagement Scale; Mameli and Passini, 2017); one questionnaire was used to account for problematic smartphone use (SAS-SV; De Pasquale et al., 2017); and assessments of general cognitive ability, language, and learning skills along with narrative tests were used to evaluate the ability to understand and resolve social problems (Roberts-2 test; Parolin et al., 2019). Participants engaged in questionnaires collectively, whereas cognitive, language, learning measures, and Roberts-2 tests were administered individually in a quiet and well-lit room provided by the school.

## Measures

### Roberts-2 Test

Roberts-2 is a performance-based narrative test for children and adolescents 6 to 18 years old (Roberts and Gruber, 2005 for Italian version, see: Parolin et al., 2019). The test is composed of 27 black and white stimulus cards differentiated based on gender. The final set for each participant is composed of 16 pictures. The pictures depict social situations of everyday life both with peers and family. For each image, the participant is asked to create a complete story, composed by a beginning, a description of what is happening in the picture, an end, and a description of the characters' feelings. The scoring system allows for the evaluation of the content of each story according to specific areas of functioning. The first area refers to the narrative components necessary to complete the task [scales: Complete Meaning (MEAN), Problem Identification (PID2, PID3, PID4), and Resolution (RES2, RES3, RES4, RES5)]. These variables refer to cognitive psychological functions that enable the adolescent to build a story with a complex and elaborated structure. The second area of functioning is evaluated by scales that measure emotional content and the resources on which the adolescent can rely to solve the problem depicted in the picture [scales: Support Self Feelings (SUPS-F), Support Self Advocacy (SUPS-A), Support Other Feelings (SUPO-F), Support Other Help (SUPO-H), Reliance on Other (REL), Limit Setting (LIM), Anxiety (ANX), Depression (DEP), Aggression (AGG), Rejection (REJ), and Unresolved Outcome (UNRS)]. These scales evaluate the ability to organize stories and include the possibilities to ask for help in a supportive environment, to rely on personal resources, and to take into consideration emotions and feelings.

### Comprehensive Inventory of Thriving

The CIT is a validated self-report questionnaire composed of 54 items assessing seven dimensions of psychological well-being: Relationship, Engagement, Mastery, Autonomy, Meaning, Optimism, and Subjective Well-Being. A total score (Su et al., 2014) is also obtained. Participants were instructed to respond to each item on a scale of 1 ("Strongly Disagree") to 5 ("Strongly Agree").

### Student Engagement Scale

The Student Engagement Scale is composed of 43 items measuring student engagement in school in four dimensions: affective, behavioral, cognitive, and agentic (Mameli and Passini, 2017). The items are rated on a 7-point Likert scale of frequency or agreement (from 1 = never/strongly disagree to 7 = always/strongly agree).

### Smartphone Addiction Scale—Short Version SAS-SV

The Smartphone Addiction Scale (SAS) is a validated screening tool measuring excessive smartphone use among adolescents and young adults (Kwon et al., 2013; for Italian validation, see De Pasquale et al., 2017). The short version of SAS is composed of 10 questions rated on a dimensional scale from 1 "completely disagree" to 6 "strongly agree." The total score ranges from 10 to 60, and higher scores correspond to a higher level of smartphone use.

## Statistics Analysis

Four sets of statistical analysis were conducted. First, to investigate group differences in the ability to produce completed and elaborated solutions on social problems, a set of independent *t*-tests were performed on all subscales of Roberts-2 test. Given that Problem Identification Scales (PID) and Resolution Scales (RES) are mutually exclusive subscales, a composite score for each of them was computed. Following Parolin et al. (2019), the composite score was calculated by multiplying the odds obtained in each scale with their order and adding the resulting scores (i.e., RES1 + RES2 ∗ 2 + RES3 × 3 + RES4 × 4 + RES × 5). Second, correlation analysis was computed to investigate the relationship between the ability to process social information and resolve social problems and the perception of thriving and well-being at school. Third, to investigate differences in problematic smartphone use, an independent *t*-test was run on the scores obtained from the SAS-SV questionnaire for the two groups. Finally, correlation analysis was planned to further explore the relationship between social cognitive functioning (Roberts-2) and problematic smartphone use.

Precisely due to the relatively small sample, Spearman correlations were computed between the scores obtained in the Roberts-2 and the scores obtained by CIT, the Student Engagement Scale, and SAS-SV for each group.

# RESULTS

## Social Cognitive Functioning

Differences between SLD and the control group were assessed by means of an independent *t*-test on the T-scores in each scale of Roberts-2 (Parolin et al., 2019). Specifically, the *t*-test comparison revealed that the control group obtained higher T-scores than the SLD group in POP (CNT group: $M = 47.37$, $SD = 9.51$; SLD Group: $M = 40.89$, $SD = 8.18$; $t(36) = 2.25$, $p = 0.031$, $d' = 0.75$) and in MEAN (CNT group: $M = 45.00$, $SD = 6.52$; SLD group: $M = 40.37$, $SD = 4.34$; $t(36) = 2.58$, $p = 0.014$, $d' = 0.86$). Furthermore, the CNT group showed a greater ability to resolve a social problem positively and with an increased level of elaboration (RES Composite) than SLD group (CNT: $M = 23.21$, $SD = 12.22$; SLD: $M = 15.21$, $SD = 7.52$; $t(36) = 2.43$, $p = 0.020$, $d' = 0.81$). Moreover, the *t*-tests revealed a reliable difference between groups in the scales related to external support as SUPO-H and REL. Specifically, the control group

score was higher than SLD group score in both SUPO-H (CNT: $M$ = 57.42, $SD$ = 9.85; SLD: $M$ = 51.79, $SD$ = 6.75; $t(36)$ = 2.06, $p$ = 0.047, $d$' = 0.69); and REL (CNT: $M$ = 49.47, $SD$ = 7.38; SLD: $M$ = 43.74, $SD$ = 5.08; $t(36)$ = 2.79, $p$ = 0.008, $d$' = 0.93). Finally, the control group obtained a significantly greater score in LIM than SLD (CNT: $M$ = 52.95, $SD$ = 11.05; SLD: $M$ = 45.95, $SD$ = 10.57; $t(36)$ = 2.00, $p$ = 0.054, $d$' = 0.67). No other effect was found ($p_s > 0.132$). Participant scores for each scale are shown in **Figure 1**.

## Social Cognitive Functioning and Well-Being

In order to investigate whether a different functioning in social competence was related to well-being, Spearman correlations analysis was conducted between the Roberts-2 scales (POP, MEAN, RES-Composite, SUPO-H, REL, and LIM), the scores obtained in CIT 8 subscales, and the scores of the 4 subscales of Student Engagement Scale separately for each group. Concerning holistic well-being, a significant positive correlation was found in the CNT group between SUPO-H and Engagement-CIT [$r_s(19) = 0.470$, $p = 0.042$] and also with Total-CIT [$r_s(19) = 0.467$, $p$ = 0.044]. Similarly, a significant positive correlation was found in the SLD group between REL and Meaning-CIT, [$r_s(18)$ = 0.505, $p$ = 0.032]. No other significant effects were found in the CNT group ($p_s > 0.130$) and in the SLD group ($p_s > 0.080$). Concerning school engagement, the analysis for the CNT group revealed that RES-Composite and SUPO-H correlated positively with Affective Engagement [$r_s(17) = 0.534$, $p = 0.027$; $r_s(17) = 0.628$, $p = 0.007$]. On the contrary, the scores obtained in LIM by the SLD group correlated negatively with the Cognitive Engagement Scale [$r_s(18) = -0.561$, $p$ = 0.015]. No other significant effects were found in the CNT group ($p_s > 0.095$) and in the SLD group ($p_s > 0.084$).

### Social Cognitive Functioning and Smartphone Use

Differences between groups in problematic use of smartphones was measured by an independent $t$-test. No significant differences were found between the CNT group ($M$ = 28.28, $SD$ = 10.42) and the SLD group ($M$ = 29.58, $SD$ = 8.66), $t(35) = -0.414$, $p > 0.005$.

To further explore the relationship between different psychosocial functioning profiles raised by Roberts-2 and smartphone use, Spearman correlations between the scores obtained in POP, MEAN, RES-Composite, SUPO-H, REL, and LIM Roberts-2 scales and self-report SAS-SV were run for each group. Concerning the CNT group, the analysis revealed that POP, MEAN, RES-Composite, and REL were negatively linked to smartphone abuse [$r_s(18) = -0.455$, $p$ = 0.058; $r_s(18) = -0.474$, $p = 0.047$; $r_s(18) = -0.541$, $p = 0.020$; $r_s(18) = -0.572$, $p = 0.013$]. No other significant correlations were found either for the CNT group ($p_s > 0.206$) or for the SLD group ($p_s > 0.324$).

# DISCUSSION

Although several studies have shown the presence of socio-cognitive and emotional difficulties in SLD children, these aspects have been less investigated in adolescence. Moreover, relatively little research has analyzed the connections between social functioning, well-being, and smartphone use through both performance-based and self-report measures. Previous studies (Bauminger et al., 2005; Bauminger-Zviely et al., 2019) have demonstrated the presence of difficulties in SLD students in understanding and solving social situations, abilities that support a more general social competence. Our findings partially overlap with previous research and deepen the knowledge about this area of functioning. Specifically, our results show significant differences between adolescents with and without a diagnosis of SLD in terms of the functions that underlie the ability to tell a complex and articulated story, such as recognize, explain, and resolve a problematic social situation. Specifically, the SLD group was less able to identify the problem depicted in the picture, to provide adaptive and complex solutions to it, and to produce complete stories by selecting and organizing the main elements in the pictures. These results indicate that

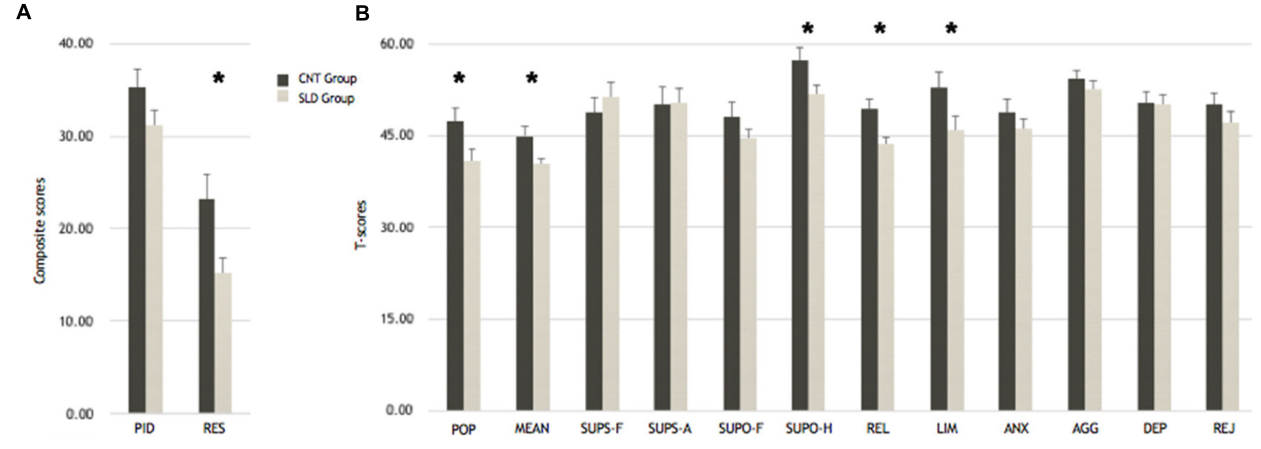


**FIGURE 1 |** Results from $t$-tests comparison on Roberts test scores between groups. **(A)** Composite scores calculated on raw scores obtained in PID and RES scales. **(B)** Comparison on T-scores obtained in cognitive psychological functions and emotional scales. *$p \leq 0.05$.

SLD students have difficulties in perceiving everyday social situations accurately and adaptively and in understanding the process by which social problems can be solved. These abilities are essential to behave appropriately in a social context, to understand other people's behavior, and, therefore, to succeed in creating supportive relationships. As far as social support is concerned, our data suggest that SLD students perceive the social environment as less helpful and trustworthy. As stated before, this can be partially explained by parent, teacher and classmate perceptions of SLD students that is frequently negative and can worsen the self-esteem and self-efficacy of these students, making it even more difficult to ask and receive help (Livingston et al., 2018). However, our results don't allow us to establish the nature of these social difficulties, which should be the topic of further studies. Understanding the processes underlying social functioning can be highly informative, especially in SLD students. Indeed, the perception of support by others and the feeling that it is possible to ask for help when needed are aspects that tap into established relational expectations and are key resources when facing problems in life. This is especially true in school, which is both a learning and a social context, where the possibility to ask and receive help is fundamental to overcome day-to-day challenges. Research has revealed that these abilities affect SLD achievement more than for their peers, especially in higher grades of education (Trainin and Swanson, 2005). In addition to this, the significant difference on the Limit Setting Scale between groups suggests that SLD students perceive the adult world as less able to administer fitting consequences for problem behaviors. A recent study on parental styles suggested that parents of SLD children experience higher levels of distress (Bonifacci et al., 2016) and struggle to maintain a stable parental style, possibly moving from granting independence to over-reactivity. These results are particularly meaningful if compared to the fatigue reported by teachers in dealing with students with special needs, and the students' perceptions of a worse relationship with their teachers (Tobia and Marzocchi, 2015). The difficulties felt by the adult world in handling children with SLD can partially explain our results, as they could be responsible for the perception of a less constructive presence of parents and teachers.

Contrary to the literature (i.e., Mugnaini et al., 2009), adolescents with SLD did not show a higher level of internalizing problems compared to typically developing peers. However, it should be noted that the literature is not consistent in reporting evidence for these problems in the SLD population (e.g., Miller et al., 2005), suggesting that the difference between SLD and control groups might be sensitive to the assessment battery and to environmental factors, such as receiving an intervention and an early diagnosis (e.g., Mugnaini et al., 2009). Moreover, the lack of internalizing problems may, at least in part, be explained by the age of the sample. Adolescence is an emotionally challenging period for every student, so it is less likely that a difference would be found due to increased individual variability in both groups. Unfortunately, our data do not enable us to specify what type of mechanism may underpin the processing of emotion in the SLD population, and further investigations are needed.

Concerning the second hypothesis, the results suggest a similar profile of general well-being functioning for both SLD and control students. An increasing sense of thriving is related to growing trust and perceived support from others. In particular, the possibility to ask for help when needed and to trust others seems more important for SLD adolescents, while a key for the control group is the role played by the concrete help provided by the social environment. On the contrary, school engagement seems to be related to different aspects of socio-cognitive functioning in SLD and control groups. Existing literature has pinpointed a link between parent, peer, and teacher support and school affective engagement (Furrer and Skinner, 2003; Estell and Perdue, 2013), emphasizing the role of support from others as a fundamental school commitment predictor. This is precisely what we found in typically developing adolescents. Our study further enhances these results, as it identifies positive relations between affective school engagement and a deep understanding of the process underlying the resolution of a social situation. This is of fundamental importance, as the class is a social context as well as an academic one and learning how to interact successfully with peers and adults is an essential developmental task that is deeply connected with school engagement. This relation seems absent in SLD students, for whom other aspects are more relevant to school engagement. Notably, their investment in learning processes decreases the more they perceive parents and teachers as ready to impose rules and limits for their shortcomings. This is interesting given that SLD students perceive external regulation less than their peers, however, when they do, it seems complicated to attribute a constructive meaning to it. Unfortunately, our results do not allow us to determine the direction of this result.

Turning to our third hypothesis, our results do not highlight a significative difference between groups in terms of smartphone use. Nevertheless, the correlations between socio-emotional functioning and problematic smartphone use highlight a different functioning in the two groups. In typically developing adolescents, increasing smartphone use seems related to a decrease in perceived support by others, which is confirmed by recent studies on this topic (Wang et al., 2017; Herrero et al., 2019). Furthermore, smartphone use is negatively related to social-cognitive variables, such as the ability to produce an increasingly complex solution to social problems taking into consideration all the main elements. This result links problematic smartphone use not only with social support but also with more in-depth aspects of social understanding and interpretation. Surprisingly, this pattern was not found in SLD students, whose scarce perception of social support seems unrelated to the use of the smartphone. To the best of our knowledge, researchers have not yet addressed this issue, so we can only argue that the diminished social support and trust in others may be a problem for SLD that does not depend on problematic smartphone use, but that is probably due to other variables, such as parenting styles (Bonifacci et al., 2016) and/or self and social-stigma associated with SLD (Chan et al., 2017). To note, the main limitation of this study is the relatively small sample size; therefore, our results are to be interpreted carefully and confirmed by further investigations.

While the present study takes into consideration just some of the variables involved in socio-cognitive and emotional functioning, it contributes to our understanding of psychosocial

variables relevant to adolescents with SLD. It also highlights the relevance of a better understanding of socio-emotional functioning, especially in adolescents with SLD, in order to improve their academic experience and, in general, their quality of life.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Catholic University of the Sacred Heart. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

DS has conceived the study, has predisposed the clinical test to be administered and the inclusion criteria for enrollment of the subjects, as supervisor of the study and discussed the results. RB was responsible for collecting the data, data analysis, drafting of all sections of the manuscripts and critical revision of the final version of the manuscript in collaboration with IO. MT has contributed to the data collection and coding (narrative test). EL has contributed to the data collection and coding (self-report). DT has contributed to the definition of self-report. ML led the discussion on test and results.

## FUNDING

This work was supported by the CARIPLO Foundation (Grant 2017-NAZ-0131 "New technologies for education and their impact on students' well-being and inclusion").

## ACKNOWLEDGMENTS

The authors wish to express their gratitude to all participants and schools who donated their time to participate in the study. A special thanks, also, to the CARIPLO Foundation to support this work. Finally, a sincere thank also to the Reviewers, who helped to improve the quality of the present manuscript.

## REFERENCES

Ali, A., Hassiotis, A., Strydom, A., and King, M. (2012). Self stigma in people with intellectual disabilities and courtesy stigma in family carers: a systematic review. *Res. Dev. Disabil.* 33, 2122–2140. doi: 10.1016/j.ridd.2012.06.013

Al-Yagon, M. (2012). Adolescents with learning disabilities: socioemotional and behavioral functioning and attachment relationships with fathers, mothers, and teachers. *J. Youth Adolesc.* 41, 1294–1311. doi: 10.1007/s10964-012-9767-6

Al-Yagon, M., and Margalit, M. (2013). "Social cognition of children and adolescents with LD: intrapersonal and interpersonal perspectives," in *Handbook of Learning Disabilities*, eds H. L. Swanson, K. Harris, and S. Graham (New York, NY: Guilford Press), 278–292.

American Psychiatric Association [APA] (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. Washington, DC: American Psychiatric Association.

Arina, S., Iervolino, I., and Stella, G. (2013). Prima raccolta di dati normativi per la valutazione della dislessia evolutiva negli adolescenti su un campione di scuola secondaria di secondo grado. *Dislessia* 10, 9–38.

Aro, T., Eklund, K., Eloranta, A. K., Närhi, V., Korhonen, E., and Ahonen, T. (2019). Associations between childhood learning disabilities and adult-age mental health problems, lack of education, and unemployment. *J. Learn. Disabil.* 52, 71–83. doi: 10.1177/0022219418775118

Bauminger, N., and Kimhi-Kind, I. (2008). Social information processing, security of attachment, and emotion regulation in children with learning disabilities. *J. Learn. Disabil.* 41, 315–332. doi: 10.1177/0022219408316095

Bauminger, N., Schorr Edelsztein, H., and Morash, J. (2005). Social information processing and emotional understanding in children with LD. *J. Learn. Disabil.* 38, 45–61. doi: 10.1177/00222194050380010401

Bauminger-Zviely, N., Alon, M., Brill, A., Schorr-Edelsztein, H., David, T., Tubul, G., et al. (2019). Social information processing among children with ASD, SLD, and typical development: the mediational role of language capacities. *J. Spec. Educ.* 53, 1–13. doi: 10.1177/0022466918821400

Bonifacci, P., Storti, M., Tobia, V., and Suardi, A. (2016). Specific learning disorders: a look inside children's and parents' psychological well-being and relationships. *J. Learn. Disabil.* 49, 532–545. doi: 10.1177/0022219414566681

Bryan, T. (2005). Science-based advances in the social domain of learning disabilities. *Learn. Disabil. Q.* 28, 119–121. doi: 10.2307/1593608

Chan, Y., Chan, Y. Y., Cheng, S. L., Chow, M. Y., Tsang, Y. W., Lee, C., et al. (2017). Investigating quality of life and self-stigma in hong kong children with specific learning disabilities. *Res. Dev. Disabil.* 68, 131–139. doi: 10.1016/j.ridd.2017.07.014

Cornoldi, C., Baldi, A. P., and Giofrè, D. (2017). *Prove MT Avanzate-3-Clinica*. Giunti OS: Organizzazioni Speciali.

Cosden, M., Brown, C., and Elliott, K. (2002). "Development of self-understanding and self-esteem in children and adults with learning disabilities," in *Social Dimensions of Learning Disabilities*, eds B. Y. L. Wong and M. L. Donahue (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 33–51.

Crick, N. R., and Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychol. Bull.* 115:74. doi: 10.1093/deafed/enw030

De Pasquale, C., Sciacca, F., and Hichy, Z. (2017). Italian validation of smartphone addiction scale short version for adolescents and young adults (SAS-SV). *Psychology* 8:1513. doi: 10.4236/psych.2017.810100

Do, T. T. T., Le, M. D., Van Nguyen, T., Tran, B. X., Le, H. T., Nguyen, H. D., et al. (2018). Receptiveness and preferences of health-related smartphone applications among Vietnamese youth and young adults. *BMC Public Health* 18:764. doi: 10.1186/s12889-018-5641-0

Estell, D. B., and Perdue, N. H. (2013). Social support and behavioral and affective school engagement: the effects of peers, parents, and teachers. *Psychol. Sch.* 50, 325–339. doi: 10.1002/pits.21681

Francis, D. A., Caruana, N., Hudson, J. L., and McArthur, G. M. (2018). The association between poor reading and internalising problems: a systematic review and meta-analysis. *Clin. Psychol. Rev.* 67, 45–60. doi: 10.1016/j.cpr.2018.09.002

Furrer, C., and Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *J. Educ. Psychol.* 95:148. doi: 10.1037/0022-0663.95.1.148

Hawi, N. S., and Samaha, M. (2017). The relations among social media addiction, self-esteem, and life satisfaction in university students. *Soc. Sci. Comp. Rev.* 35, 576–586. doi: 10.1177/0894439316660340

Herrero, J., Urueña, A., Torres, A., and Hidalgo, A. (2019). Socially connected but still isolated: smartphone addiction decreases social support

over time. *Soc. Sci. Comp. Rev.* 37, 73–88. doi: 10.1177/0894439317742611

Hopwood, C. J., and Bornstein, R. F. (2014). *Multimethod Clinical Assessment*. New York, NY: Guilford Press.

ISTAT (2018). *Internet@italia2018. Domanda e Offerta di Servizi Online e Scenari di Digitalizzazione*. Roma: Fondazione Ugo Bordoni.

Kim, J. H. (2017). Smartphone-mediated communication vs. face-to-face interaction: two routes to social support and problematic use of smartphone. *Comp. Hum. Behav.* 67, 282–291. doi: 10.1016/j.chb.2016.11.004

Kwon, M., Kim, D. J., Cho, H., and Yang, S. (2013). The smartphone addiction scale: development and validation of a short version for adolescents. *PLoS One* 8:e83558. doi: 10.1371/journal.pone.0083558

La Greca, A. M., and Harrison, H. M. (2005). Adolescent peer relations, friendships, and romantic relationships: do they predict social anxiety and depression? *J. Clin. Child Adolesc. Psychol.* 34, 49–61. doi: 10.1207/s15374424jccp3401-5

Lawless, K. A., and Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: knowns, unknowns, and ways to pursue better questions and answers. *Rev. Educ. Res.* 77, 575–614. doi: 10.3102/0034654307309921

Lepp, A., Barkley, J. E., and Karpinski, A. C. (2014). The relationship between cell phone use, academic performance, anxiety, and satisfaction with life in college students. *Comp. Hum. Behav.* 31, 343–350. doi: 10.1016/j.chb.2013.10.049

Livingston, E. M., Siegel, L. S., and Ribary, U. (2018). Developmental dyslexia: emotional impact and consequences. *Aust. J. Learn. Diffic.* 23, 107–135. doi: 10.1080/19404158.2018.1479975

Mameli, C., and Passini, S. (2017). Development and validation of an enlarged version of the student agentic engagement scale. *J. Psychoeduc. Assess.* 37, 450–463. doi: 10.1177/0734282918757849

Miller, C. J., Hynd, G. W., and Miller, S. R. (2005). Children with dyslexia: not necessarily at risk for elevated internalizing symptoms. *Read. Writ.* 18, 425–436. doi: 10.1007/s11145-005-4314-4

MIUR (2015). *L'integrazione Scolastica Degli Alunni con Disabilità a.s.* 2014/2015. Rome: MIUR.

Mugnaini, D., Lassi, S., La Malfa, G., and Albertini, G. (2009). Internalizing correlates of dyslexia. *World J. Pediatr.* 5, 255–264. doi: 10.1007/s12519-009-0049-7

Parolin, L., De Carli, P., and Locati, F. (2019). The roberts–2: italian validation on a sample of children and adolescents. *J. Pers. Assess.* 4, 1–15. doi: 10.1080/00223891.2018.1546713

Perelmutter, B., McGregor, K. K., and Gordon, K. R. (2017). Assistive technology interventions for adolescents and adults with learning disabilities: an evidence-based systematic review and meta-analysis. *Comp. Educ.* 114, 139–163. doi: 10.1016/j.compedu.2017.06.005

Raven, J. (2000). The raven's progressive matrices: change and stability over culture and time. *Cogn. Psychol.* 41, 1–48. doi: 10.1006/cogp.1999.0735

Roberts, G. E., and Gruber, C. (2005). *Roberts-2*. Torrance, CA: Western Psychological Services.

Samaha, M., and Hawi, N. S. (2016). Relationships among smartphone addiction, stress, academic performance, and satisfaction with life. *Comp. Hum. Behav.* 57, 321–325. doi: 10.1016/j.chb.2015.12.045 0747-5632

Sartori, G., and Job, R. (2007). *DDE-2: Giunti OS Organizzazioni Speciali: Batteria per la Valutazione Della Dislessia e Della Disortografia Evolutiva-2: Manuale*. Giunti, OS: Organizzazioni Speciali.

Su, R., Tay, L., and Diener, E. (2014). The development and validation of the comprehensive inventory of thriving (CIT) and the brief inventory of thriving (BIT). *Appl. Psychol.* 6, 251–279. doi: 10.1111/aphw.12027

Teppers, E., Luyckx, K., Klimstra, T. A., and Goossens, L. (2014). Loneliness and Facebook motives in adolescence: a longitudinal inquiry into directionality of effect. *J. Adolesc.* 37, 691–699. doi: 10.1016/j.adolescence.2013.11.003

Tobia, V., and Marzocchi, G. M. (2015). Il benessere scolastico: una ricerca su bambini con sviluppo tipico e con bisogni educativi speciali. *Dif. App. Didattica Inclusiva* 3, 221–232.

Trainin, G., and Swanson, H. L. (2005). Cognition, metacognition, and achievement of college students with learning disabilities. *Learn. Disabil. Q.* 28, 261–272. doi: 10.2307/4126965

Tran, B. X., Zhang, M. W., Le, H. T., Nguyen, H. D., Nguyen, L. H., Nguyen, Q. L., et al. (2018). What drives young vietnamese to use mobile health innovations? implications for health communication and behavioral interventions. *JMIR Mhealth Uhealth.* 6:e194. doi: 10.2196/mhealth.6490

Twenge, J. M. (2019). More time on technology, less happiness? Associations between digital-media use and psychological well-being. *Curr. Direct. Psychol. Sci.* 28, 372–379. doi: 10.1177/0963721419838244

Van Deursen, A. J., Bolle, C. L., Hegner, S. M., and Kommers, P. A. (2015). Modeling habitual and addictive smartphone behavior: the role of smartphone usage types, emotional intelligence, social stress, self-regulation, age, and gender. *Comp. Hum. Behav.* 45, 411–420. doi: 10.1016/j.chb.2014.12.039 0747-5632

Veissière, S. P., and Stendel, M. (2018). Hypernatural monitoring: a social rehearsal account of smartphone addiction. *Front. Psychol.* 9:141. doi: 10.3389/fpsyg.2018.00141

Wang, P., Zhao, M., Wang, X., Xie, X., Wang, Y., and Lei, L. (2017). Peer relationship and adolescent smartphone addiction: the mediating role of self-esteem and the moderating role of the need to belong. *J. Behav. Addict.* 6, 708–717. doi: 10.1556/2006.6.2017.079

Welsh, M., Parke, R. D., Widaman, K., and O'Neil, R. (2001). Linkages between children's social and academic competence: a longitudinal analysis. *J. Sch. Psychol.* 39, 463–482. doi: 10.1016/s0022-4405(01)00084-x

Wiener, J., and Schneider, B. H. (2002). A multisource exploration of the friendship patterns of children with and without learning disabilities. *J. Abnorm. Child Psychol.* 30, 127–141.

Wilson, A. M., Deri armstrong, C., Furrie, A., and Walcot, E. (2009). The mental health of Canadians with self-reported learning disabilities. *J. Learn. Disabil.* 42, 24–40. doi: 10.1177/0022219408326216

World Medical Association (2013). *WMA Declaration of Helsinki-Ethical Principles for Medical Research Involving Human Subjects*. Ferney-Voltaire: World Medical Association.

Zhang, M. W., Tran, B. X., Hinh, N. D., Nguyen, H. L. T., Tho, T. D., Latkin, C., et al. (2017). Internet addiction and sleep quality among vietnamese youths. *Asian J. Psychiatr.* 28, 15–20. doi: 10.1016/j.ajp.2017.03.025

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 06 March 2020
doi: 10.3389/fpsyg.2020.00375

# Emotional Intelligence, Self-Regulation, Smartphone Addiction: Which Relationship With Student Well-Being and Quality of Life?

*Maria Lidia Mascia*, Mirian Agus and Maria Pietronilla Penna*

*Department of Pedagogy, Psychology, Philosophy, Faculty of Humanities, University of Cagliari, Cagliari, Italy*

This study emphasizes the importance of analyzing factors that contribute to student well-being, as a result of the multiplicity of factors that can affect their quality of life (QoL). The literature indicates that, among these factors, emotional intelligence and self-regulation play a central role in influencing adolescents' psychological and scholastic well-being. Technology is a fundamental aspect of adolescent life but addiction to the use of smartphones is increasing, which can affect both emotional intelligence and self-regulation, and in turn individual well-being and QoL. Therefore, this study explores the role of smartphone use with respect to these aspects. Participants were 215 Italian students attending middle school. By applying partial least squares structural equation modeling (PLS-SEM), the results confirm that self-regulation affects the QoL of students, but its role varies according to the degree of smartphone addiction. In conclusion, we confirm the relevance of the relationship between self-regulation and smartphone addiction in teaching students to be aware of their time spent using smartphones. Emotional intelligence and, in general, self-regulation should be encouraged to support the well-being and QoL of students in their adolescence at school.

**Keywords: adolescence, smartphone addiction, self-regulation, emotional intelligence, well-being, quality of life**

OPEN ACCESS

***Edited by:***
*Daniela Traficante, Catholic University of the Sacred Heart, Italy*

***Reviewed by:***
*Roger C. Ho, National University of Singapore, Singapore*
*Chung-Ying Lin, The Hong Kong Polytechnic University, Hong Kong*

****Correspondence:***
*Maria Lidia Mascia*
*marialidia.mascia@unica.it;*
*marialidiamascia@gmail.com*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *31 October 2019*
***Accepted:*** *18 February 2020*
***Published:*** *06 March 2020*

***Citation:***
*Mascia ML, Agus M and Penna MP (2020) Emotional Intelligence, Self-Regulation, Smartphone Addiction: Which Relationship With Student Well-Being and Quality of Life? Front. Psychol. 11:375. doi: 10.3389/fpsyg.2020.00375*

## INTRODUCTION

The well-being of students at school is a primary concern for teachers and educators (Stefansson et al., 2018) as it is strictly related to their quality of life (QoL: Camfield and Skevington, 2008). Several studies (Shoshani et al., 2016; Navarro et al., 2017) examine factors that can positively influence student well-being and QoL in adolescence, finding it to be the result of a combination of affective, behavioral and cognitive dimensions. Some literature shows a link between emotional intelligence and well-being (Zeidner and Olnick-Shemesh, 2010), particularly at class level (Balluerka et al., 2016). Another fundamental dimension connected to these two aspects is self-regulation (Thomas et al., 2019). Self-regulation strategies facilitate students' planning and goal-setting prior to learning by enhancing their attention-focusing and self-monitoring processes (self-reflection) during learning or task performance (Zimmerman, 2002; Cleary and Chen, 2009).

Digital society provides numerous opportunities but despite the implied advantages it also brings risks, especially for younger people (Machimbarrena et al., 2019); indeed, use of the internet can

become problematic, leading to consequences for personal well-being. In particular, young people are continually increasing their smartphone use (Humphreys et al., 2013) and internet addiction has become ubiquitous (Haverlag, 2013; Yam et al., 2019). A body of research states that problematic internet use can become addictive but the issue of smartphone use is more complex; undeniably, smartphones can link to the internet and also execute various types of applications (e.g., gaming, gambling, social media use, etc.), consequently causing psychological impairment (Lin et al., 2019; Yam et al., 2019). Adolescents between 16 and 18 years old were less likely to believe in the negative impact of the internet on health than older people (Do et al., 2020). The prevalence of internet addiction is 1.2–4.9% (Mak et al., 2014) in adolescents and as high as 30% in university students (Zhang and Ho, 2017). Most studies on this issue focus on describing behaviors and consequences, including depression, anxiety, alcohol misuse, musculoskeletal discomfort, and sleep problems (Bianchi and Phillips, 2005; Ho et al., 2014; Yang et al., 2017; Zhang and Ho, 2017; Alimoradi et al., 2019; Chen et al., 2020). Generally, studies emphasize that internet addiction is inversely related to the global Life Satisfaction Index (Cheng et al., 2018) and health-related QoL (Tran et al., 2017), leading to the need to spend increasing time on internet gaming and losing interest in hobbies, relationships, and educational opportunities (Ho et al., 2014). Many studies emphasize that self-regulation constructs are adversely affected by smartphone addiction (van Deursen et al., 2015), but self-regulation may contribute to the suppression of addictive behavior (Baumeister and Vonasch, 2015). Other studies hypothesize that people who are able to express and understand emotions and regulate feelings are better adjusted psychologically and socially and have a high level of well-being (Gascó et al., 2018), therefore it is important to preserve this dimension.

What is the relationship between these variables? These premises underline the need for attention to factors that can positively or negatively affect adolescent well-being. This study considers the effects of self-regulation (hypothesis 1a, H1a) and emotional intelligence (H1b) on scholastic well-being. The innovation in this model relates to the role that smartphone dependence plays in these relationships. We assess if smartphone dependence might mediate the effects of self-regulation (H2a) and emotional intelligence (H2b) on scholastic well-being. Furthermore, it is of interest to evaluate the potential moderating effect of smartphone addiction on the relation between self-regulation and well-being (H3a) and between emotional intelligence and well-being (H3b) (**Figure 1**). These hypotheses are based on previous research findings in the literature (Zimmermann and Iwanski, 2014; Verzeletti et al., 2016; Chung, 2019; Xu et al., 2019).

# METHOD

## Participants

This study involved 215 students (mean age 12.7 years; SD = 0.90) attending their last year of middle school in Sardinia (Italy).

## Measures and Procedure

The survey was conducted in the third-year classes of state middle schools during the school timetable, subject to agreement from parents, the headmaster and teachers. Informed consent was given by the students' parents after the features and aims of the study had been explained to them. We established an atmosphere of participation and trust in all classes, allowing the students to choose to participate in the research and motivating them sufficiently for the purpose of the study. To prevent teachers from interfering during the survey, they were asked to adopt a neutral stance if they were present in the classroom. All teachers were helpful and cooperative, leaving the management of the survey to the provider. Data collection was carried out during school hours from 8:30 to 13:00. Administration of the questionnaire took 25–45 min, preceded by a detailed explanation of the objectives of the survey, the structure of the questionnaire, the method of compilation and the anonymity of the test. The study was approved by the ethics committee at the University of Cagliari, Italy (the Department of Pedagogy, Psychology, Philosophy).

The study protocol comprises five sections:

(1) A questionnaire on socio-demographic characteristics collects specific information such as age and educational level.
(2) The *self-regulation* questionnaire (Moè and De Beni, 2000) aims to identify the components of the self-regulation approach to the study, with particular reference to three meta-cognitive dimensions: processing skills, organization and self-evaluation. Research shows how students can organize their study activities with a time-bound work programe that complies with commitments and deadlines (Ley and Young, 1998; Moè and De Beni, 2000), using schema-driven strategies (based on schematization, building diagrams, and tables, notepads, etc.) and the adoption of specific processing methods. Successful students are aware of their own study method, know how to properly assess their own preparation and are more likely to reflect on the best way to deal with their studies. The scale consists of 30 items (10 items for each dimension) to be answered on a Likert scale from 1 to 5, with higher values denoting better skills. Cronbach's alpha coefficient for reliability is good for all dimensions: processing skills, $\alpha = 0.81$; organization, $\alpha = 0.76$; self-evaluation, $\alpha = 0.72$.
(3) The "My Life as a Student" questionnaire (Soresi and Nota, 2003; Nota et al., 2011) allows students to explore their levels of satisfaction and *well-being*. This instrument consists of 26 items (on a five-point Likert scale, with higher scores indicating greater well-being) and examines seven satisfaction factors in several aspects: the school experience ($\alpha = 0.86$); opportunities to make autonomous decisions ($\alpha = 0.66$); relationships with classmates ($\alpha = 0.70$); current living conditions ($\alpha = 0.76$); family relationships ($\alpha = 0.71$); praise received when due ($\alpha = 0.72$); and availability of assistance ($\alpha = 0.79$).
(4) The Emotional Intelligence Scale (EIS) questionnaire (Schutte et al., 1998), conducted in a validated Italian

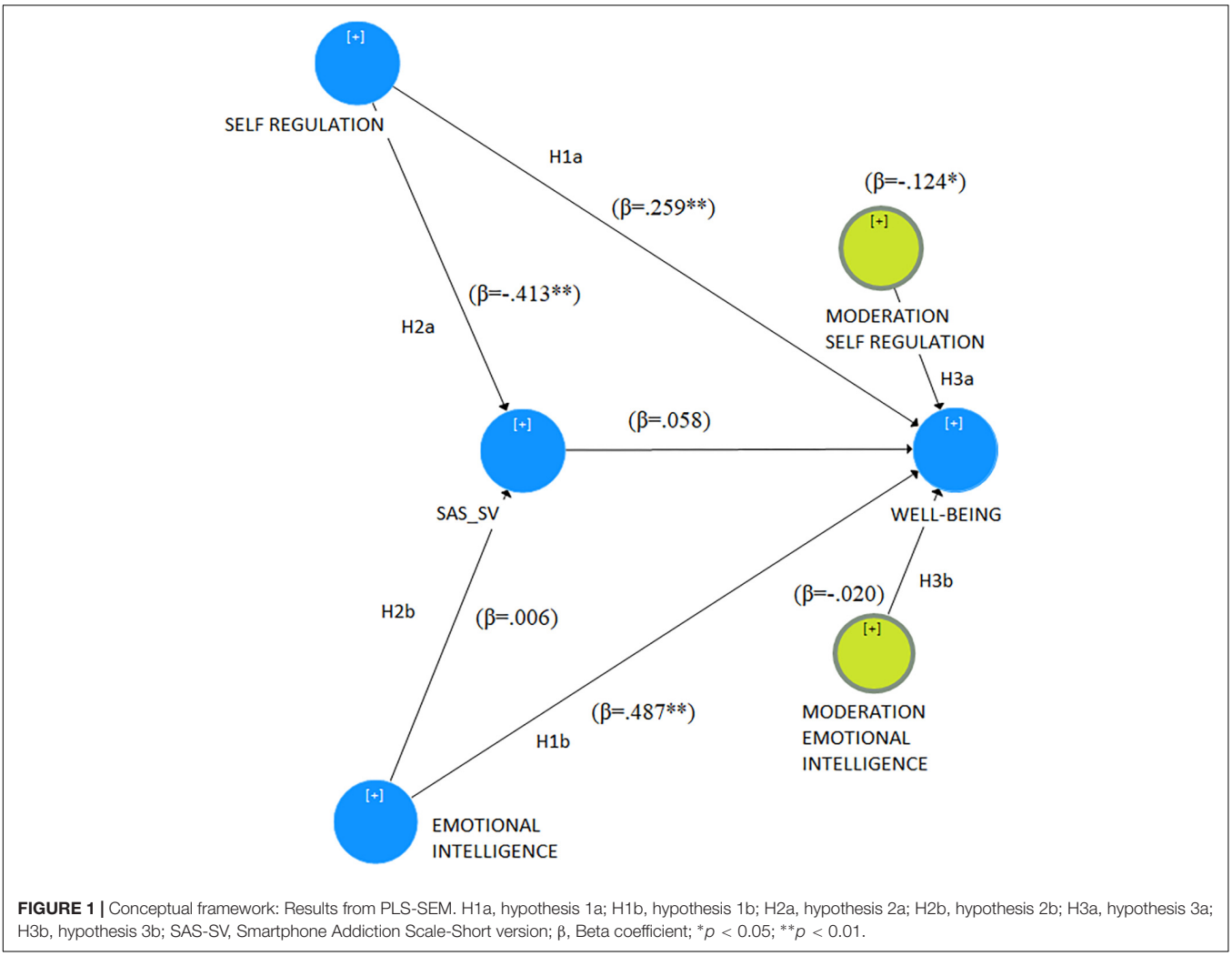


**FIGURE 1 |** Conceptual framework: Results from PLS-SEM. H1a, hypothesis 1a; H1b, hypothesis 1b; H2a, hypothesis 2a; H2b, hypothesis 2b; H3a, hypothesis 3a; H3b, hypothesis 3b; SAS-SV, Smartphone Addiction Scale-Short version; β, Beta coefficient; *$p < 0.05$; **$p < 0.01$.

version (Ciucci et al., 2009), is designed to determine *emotional intelligence* and consists of 33 closed-ended statements (five-point Likert scale, with higher values representing improved emotional intelligence) such as "I am aware of my emotions as I experience them." The three scales identified in the questionnaire measure: emotional appraisal and expression of oneself (13 items, α = 0.64), and others (10 items, α = 0.68); and regulation of emotions in oneself and others (10 items, α = 0.71).

(5) The Smartphone Addiction Scale (SAS) is a validated questionnaire designed to determine the risk level of *smartphone addiction* and identify high-risk groups among adolescents in Korea (Kwon et al., 2013). A short version (SAS-SV) was conducted, validated in Italy by De Pasquale et al. (2017). The questionnaire includes 10 questions (α = 0.79) describing daily disruptions in life, positive expectations, withdrawal, relationships in cyberspace, overuse and tolerance. Participants express their opinions on a six-point scale (1 = strongly disagree; 6 = strongly agree), with higher scores designating greater smartphone addiction (De Pasquale et al., 2017).

In order to verify the research hypotheses, this work proposed a conceptual model including those dimensions that might potentially affect student well-being. The model was devised on the basis of previous relevant works in the literature (Zimmermann and Iwanski, 2014; Gascó et al., 2018; Machimbarrena et al., 2019; Do et al., 2020). The conceptual model was assessed by applying component-based partial least squares structural equation modeling (PLS-SEM), designed to determine the values of the variables in relation to the predicted purpose (Chin, 1998). In this work PLS-SEM was used as the main statistical technique to evaluate our model due to the multiplicity of constructs and relationships to be assessed (Hair et al., 2011).

This statistical approach is particularly suitable for small samples, demonstrates robustness of non-normal data and has fewer restrictive assumptions than factor-based SEM. PLS-SEM analyses both the outer measurement model (referring to the quality, reliability and validity of the construct under study) and the inner model (where paths between latent variables are estimated) (Hair et al., 2012; Sarstedt et al., 2017). Statistical analyses are performed using the software R 3.6.1 (R Core Team,

2019) and Smart-PLS (V.3.2.8) (Ringle et al., 2015). In the model assessed, the subscales related to the constructs of *self-regulation* (processing skills; organization; self-evaluation) (Moè and De Beni, 2000) and *emotional intelligence* (emotions related to others and themselves; regulation and use of emotions) (Ciucci et al., 2009) were used as observed variables. For *smartphone addiction* on the SAS-SV (De Pasquale et al., 2017), all 10 items that explore the construct were used as observed variables. *Well-being* was measured by the seven subscales identified on the "My Life as a Student" questionnaire (Soresi and Nota, 2003).

## RESULTS

Descriptive statistics were performed on each variable to evaluate the distribution (**Table 1**). PLS-SEM was then performed with a reflective measurement model (Hair et al., 2019). **Table 2** illustrates the indicators used for the outer measurement model. The factor loadings obtained vary from to 0 458 to 0.862 for all constructs; the consistent reliability coefficient Rho_A was consistently greater than 0.7, which indicates an acceptable internal reliability for the dimensions (Dijkstra and Henseler, 2015). The constructs indicated an average variance extracted (AVE) value higher than 0.5, indicating convergent validity (Fornell and Larcker, 1981). The adjusted $R^2$ value was 0.161 for *smartphone addiction* and 0.390 for *well-being*, highlighting weak and moderate effects, respectively (Hair et al., 2011).

Concerning the inner model, each path is calculated and assessed by applying the bootstrapping routine (5000 subsamples from the original data), calculating standard errors, *T* values and *p* values. This procedure identifies the significance of each relationship and effect (Hair et al., 2019; **Table 3**). Specifically, the positive effects of *self-regulation* (H1a) ($\beta = 0.259^{**}$) and *emotional intelligence* (H1b) ($\beta = 0.487^{**}$) on *well-being* are confirmed (**Table 3**). Although the negative effect of *self-regulation* on *smartphone addiction* is confirmed ($\beta = -0.413^{**}$), the influence of *smartphone addiction* on *well-being* has not been established as an indirect overall effect, which does not support H2a. H2b has not been confirmed, highlighting that there is no indirect overall effect between *emotional intelligence, smartphone addiction*, and *well-being*. Furthermore, the findings emphasize a significant negative moderation effect of *smartphone addiction* on the relationship between *self-regulation* and *well-being* (H3a) ($\beta = -0.124^{*}$). The moderation effect of *smartphone addiction* on the relationship between *emotional intelligence* and *well-being* (H3b) has not been confirmed (**Table 3**).

## DISCUSSION AND CONCLUSION

The findings of this work highlight the multivariate relationships affecting adolescent well-being, including the role played by their dependence on smartphones. To the best of our knowledge, few works in the literature have referred to the relation between

**TABLE 1 |** Descriptive statistics for the queried variables.

| | **Variables** | **Category** | | | **Fr (%)** | | |
|---|---|---|---|---|---|---|---|
| | Gender | Female | | | 106 (49.3%) | | |
| | | **Minimum** | **Maximum** | **Mean** | **Standard deviation** | **Skewness** | **Kurtosis** |
| | Age | 10 | 15 | 12.7 | 0.907 | −0.172 | 0.564 |
| Self-regulation | Processing skills | 1.60 | 5.00 | 3.29 | 0.736 | 0.011 | −0.466 |
| | Organization skills | 2.20 | 4.80 | 3.53 | 0.520 | 0.001 | −0.254 |
| | Self-evaluation skills | 2.30 | 4.30 | 3.27 | 0.429 | −0.030 | −0.412 |
| Emotional intelligence | Appraisal and expression of emotion in the self | 2.45 | 4.91 | 3.59 | 0.537 | 0.102 | −0.420 |
| | Appraisal and expression of emotion in the others | 1.50 | 5.00 | 3.68 | 0.786 | −0.190 | −0.541 |
| | Regulation and use of emotions | 1.44 | 5.00 | 3.81 | 0.639 | −0.405 | 0.020 |
| Smartphone addiction | Smartphone addiction | 10.00 | 44.00 | 22.90 | 8.80 | 0.381 | −0.757 |
| Scholastic well-being | Satisfaction with the School experience | 7.00 | 35.00 | 26.80 | 6.02 | −0.521 | −0.327 |
| | Satisfaction with opportunities to make decisions autonomously | 5.00 | 25.00 | 17.10 | 3.84 | −0.309 | 0.255 |
| | Satisfaction with relationships with classmates | 3.00 | 15.00 | 11.00 | 2.81 | −0.668 | 0.080 |
| | Satisfaction with Current life conditions | 3.00 | 15.00 | 8.38 | 2.89 | 0.164 | −0.478 |
| | Satisfaction with relationships with family members | 4.00 | 20.00 | 7.80 | 3.37 | 1.220 | 1.540 |
| | Satisfaction with praise received when due | 2.00 | 10.00 | 5.86 | 1.88 | 0.072 | −0.499 |
| | Satisfaction with help availability | 2.00 | 10.00 | 3.96 | 1.97 | 1.010 | 0.631 |

*Fr, frequency.*

**TABLE 2 |** PLS-SEM: Outer model.

| Construct | Observed variables | Latent variable loadings | Rho_A | Average variance extracted | Adjusted $R^2$ |
|---|---|---|---|---|---|
| Self-Regulation | Processing | 0.699 | 0.700 | 0.578 | |
| | Organization | 0.708 | | | |
| | Self-evaluation | 0.862 | | | |
| Emotional Intelligence | Appraisal and expression of emotion in the self | 0.711 | 0.707 | 0.621 | |
| | Appraisal and expression of emotion in the others | 0.803 | | | |
| | Regulation and use of emotions | 0.844 | | | |
| Dependence on smartphone | Sas item1 | 0.652 | 0.881 | 0.457 | 0.161 |
| | Sas item2 | 0.520 | | | |
| | Sas item3 | 0.458 | | | |
| | Sas item4 | 0.664 | | | |
| | Sas item5 | 0.791 | | | |
| | Sas item6 | 0.700 | | | |
| | Sas item7 | 0.757 | | | |
| | Sas item8 | 0.691 | | | |
| | Sas item9 | 0.786 | | | |
| | Sas item10 | 0.666 | | | |
| Well-being | School experience | 0.761 | 0.836 | 0.475 | 0.390 |
| | Opportunities to make decisions autonomously | 0.695 | | | |
| | Relationships with classmates | 0.708 | | | |
| | Current life conditions | 0.502 | | | |
| | Relationships with family members | 792 | | | |
| | Praise received when due | 0.589 | | | |
| | Help availability | 0.732 | | | |

*Rho_A, consistent reliability coefficient; SAS, smartphone addiction scale.*

**TABLE 3 |** PLS-SEM: Inner model.

| Hypothesis | Relationship | Standardized beta | Mean | Standard deviation | *T*-value | *p* | Decision |
|---|---|---|---|---|---|---|---|
| H1a | Self-regulation ->Well-being | 0.259 | 0.257 | 0.086 | 3.014 | 0.003 | Supported |
| H1b | Emotional intelligence ->Well-being | 0.487 | 0.476 | 0.131 | 3.731 | <0.0001 | Supported |
| H2a | Self-regulation ->Dependence on smartphone | −0.413 | −0.423 | 0.056 | 7.405 | <0.0001 | supported |
| | Dependence on smartphone ->Well-being | 0.058 | 0.049 | 0.062 | 0.933 | 0.351 | Not supported |
| | Total indirect effect Self-regulation ->Dependence on smartphone ->Well-being | −0.024 | −0.021 | 0.026 | 0.900 | 0.368 | Not supported |
| H2b | Emotional intelligence ->Dependence on smartphone | 0.006 | 0.005 | 0.083 | 0.070 | 0.944 | Not Supported |
| | Total indirect effect Emotional-intelligence ->Dependence on smartphone ->Well-being | 0.000 | 0.002 | 0.007 | 0.049 | 0.961 | Not supported |
| H3a | Moderation Dependence on smartphone on Self-regulation ->Well-being | −0.124 | −0.117 | 0.062 | 2.018 | 0.044 | Supported |
| H3b | Moderation Dependence on smartphone on Emotional intelligence ->Well-being | −0.020 | −0.015 | 0.060 | 0.335 | 0.738 | Not supported |

*H1a, hypothesis 1a; H1b, hypothesis 1b; H2a, hypothesis 2a; H2b, hypothesis 2b; H3a, hypothesis 3a, H3b: hypothesis 3b; p, probability.*

smartphone addiction, emotional intelligence, self-regulation and well-being. The literature features a series of studies showing that many factors influence well-being and QoL in adolescence (Jovanović, 2016). QoL and satisfaction are defined as cognitive components of subjective well-being (Diener et al., 1999). Many scholars emphasize the multiplicity of factors concerning the emotional and self-regulation processes of young adolescents (e.g., Abe, 2011). Furthermore, recently it has been highlighted that in our daily lives (work, school, leisure) attention is often directed to smartphones. There are many advantages to using technology but the excessive use of smartphones for continuous connectivity can lead to internet addiction (Tonioni and Corvino, 2011) and to the alarming phenomenon of hikikomori (Suwa and Suzuki, 2013).

These relevant facts support the necessity to deepen our knowledge of the relationship between smartphone addiction and well-being, specifically in adolescents. The current model assessed relationships that have seldom been tested empirically

before (e.g., mediation and moderation effects of smartphone addiction in the relationships between self-regulation and well-being and between emotional intelligence and well-being). This study attempted to identify the dimensions affecting adolescent well-being and has highlighted some interesting insights. In a closer look at the relationships between the variables that underlie scholastic QoL, our findings confirm the positive effect of *self-regulation* and *emotional intelligence* on *well-being*. The negative effect of *self-regulation* on *smartphone addiction* was also highlighted. However, the indirect effects of *self-regulation*, *smartphone addiction*, and *well-being* have not been established. Moreover, the indirect overall effects of *emotional intelligence, smartphone addiction*, and *well-being* have not been confirmed. It is of interest that the results emphasize significant negative moderation effects of *smartphone addiction* on the relationship between *self-regulation* and *well-being*, highlighting that the effect of self-regulation on well-being can vary depending on the level of *smartphone addiction*. Specifically, this last significant moderation effect implies that a low level of smartphone addiction enhances the positive relation between self-regulation and well-being; on the other hand, when smartphone addiction is high, the positive relationship between self-regulation and well-being is weakened. These findings shed light on issues that should be taken into consideration to improve adolescent well-being.

Moreover, it should be pointed out that some limitations of these findings might derive from the cross-sectional research design, the non-probabilistic sampling method in the Italian context and the presentation of self-report questionnaires.

New technologies offer endless possibilities for students and schools, but we must find ways to benefit, depending on the level of smartphone addiction. For example, smartphone applications can be used to deliver immersive virtual reality therapy for treating internet addiction in adolescents (Zhang et al., 2017). Education must play an active role in helping digital natives learn about and use these new tools. Emphasis should be placed on education concerning emotional intelligence and self-regulation in order to achieve psychological and social well-being, and in turn global life satisfaction (Huebner et al., 2005).

## DATA AVAILABILITY STATEMENT

The datasets for this study are available from corresponding author on reasonable request.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Department of Pedagogy, Psychology, Philosophy, Faculty of Humanities, University of Cagliari, Cagliari, Italy. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

MM and MP contributed to the design of the study. MA analyzed the data. All authors wrote, read, revised and approved the final manuscript.

## REFERENCES


Abe, J. A. A. (2011). Positive emotions, emotional intelligence, and successful experiential learning. *Pers. Individ. Differ.* 51, 817–822. doi: 10.1016/j.paid.2011.07.004

Alimoradi, Z., Lin, C.-Y., Broström, A., Bülow, P. H., Bajalan, Z., Griffiths, M. D., et al. (2019). Internet addiction and sleep problems: a systematic review and meta-analysis. *Sleep Med. Rev.* 47, 51–61. doi: 10.1016/j.smrv.2019.06.004

Balluerka, N., Gorostiaga, A., Alonso-Arbiol, I., and Aritzeta, A. (2016). Peer attachment and class emotional intelligence as predictors of adolescents' psychological well-being: a multilevel approach. *J. Adolesc.* 53, 1–9. doi: 10.1016/j.adolescence.2016.08.009

Baumeister, R. F., and Vonasch, A. J. (2015). Uses of self-regulation to facilitate and restrain addictive behavior. *Addict. Behav.* 44, 3–8. doi: 10.1016/j.addbeh.2014.09.011

Bianchi, A., and Phillips, J. G. (2005). Psychological predictors of problem mobile phone use. *Cyber. Psychol. Behav.* 8, 39–51. doi: 10.1089/cpb.2005.8.39

Camfield, L., and Skevington, S. M. (2008). On subjective well-being and quality of life. *J. Health Psychol.* 13, 764–775. doi: 10.1177/1359105308093860

Chen, I.-H., Strong, C., Lin, Y.-C., Tsai, M.-C., Leung, H., Lin, C.-Y., et al. (2020). Time invariance of three ultra-brief internet-related instruments: smartphone Application-Based Addiction Scale (SABAS), Bergen Social Media Addiction Scale (BSMAS), and the nine-item Internet Gaming Disorder Scale- Short Form (IGDS-SF9) (Study Part B). *Addict. Behav.* 101:105960. doi: 10.1016/j.addbeh.2019.04.018

Cheng, C., Cheung, M. W. L., and Wang, H. (2018). Multinational comparison of internet gaming disorder and psychosocial problems versus well-being: Meta-analysis of 20 countries. *Comput. Hum. Behav.* 88, 153–167. doi: 10.1016/j.chb.2018.06.033

Chin, W. W. (1998). "The partial least squares approach to structural equation modeling," in *Modern Methods for Business Research*, ed. G. A. Marcoulides, (London: Lawrence Erlbaum), 295–336.

Chung, M. S. (2019). The effect of internet addiction prevention program on adolescents' self-regulation-a systematic review and meta-analysis. *J. Digit. Converg.* 17, 347–355. doi: 10.14400/JDC.2019.17.8.347

Ciucci, E., Menesini, E., Primi, C., Grazzani Gavazzi, I., and Antoniotti, C. (2009). Psychometric properties of the Italian version of the Emotional Intelligence Scale (EIS) with preadolescents counseling. *Ital. J. Res. Interv.* 2, 85–100.

Cleary, T. J., and Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: variations across grade level and math context. *J. Sch. Psychol.* 47, 291–314. doi: 10.1016/j.jsp.2009.04.002

De Pasquale, C., Sciacca, F., and Hichy, Z. (2017). Italian validation of smartphone addiction scale short version for adolescents and young adults (SAS-SV). *Psychology* 08, 1513–1518. doi: 10.4236/psych.2017.810100

Diener, E., Suh, E. M., Lucas, R. E., and Smith, H. L. (1999). Subjective well-being: three decades of progress. *Psychol. Bull.* 125:276. doi: 10.1037/0033-2909.125.2.276

Dijkstra, T. K., and Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Q.* 39, 397–316.

Do, H. N., Onyango, B., Prakash, R., Tran, B. X., Nguyen, Q. N., Nguyen, L. H., et al. (2020). Susceptibility and perceptions of excessive internet use impact on health among Vietnamese youths. *Addict. Behav.* 101:105898. doi: 10.1016/j.addbeh.2019.01.043

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18:39. doi: 10.2307/3151312

Gascó, V. P., Badenes, L. V., and Plumed, A. G. (2018). Trait emotional intelligence and subjective well-being in adolescents: the moderating role of feelings. *Psicothema* 30, 310–315. doi: 10.7334/psicothema2017.232

Hair, J. F., Ringle, C. M., and Sarstedt, M. (2011). PLS-SEM: indeed a silver bullet. *J. Mark. Theory Pract.* 19, 139–152. doi: 10.2753/MTP1069-6679190202

Hair, J. F., Risher, J. J., Sarstedt, M., and Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *Eur. Bus. Rev.* 31, 2–24. doi: 10.1108/EBR-11-2018-0203

Hair, J. F., Sarstedt, M., Pieper, T. M., and Ringle, C. M. (2012). The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long Range Plan.* 45, 320–340. doi: 10.1016/j.lrp.2012.09.008

Haverlag, D. (2013). *Compulsief Gebruik Van Sociale Media op de Smartphone*. Available at: https://dspace.library.uu.nl/handle/1874/280913 (accessed September 27, 2019).

Ho, R. C., Zhang, M. W., Tsang, T. Y., Toh, A. H., Pan, F., Lu, Y., et al. (2014). The association between internet addiction and psychiatric co-morbidity: a meta-analysis. *BMC Psychiatry* 14:183. doi: 10.1186/1471-244X-14-183

Huebner, E. S., Valois, R. F., Paxton, R. J., and Drane, J. W. (2005). Middle school students' perceptions of quality of life. *J. Happiness Stud.* 6, 15–24. doi: 10.1007/s10902-004-1170-x

Humphreys, L., Von Pape, T., and Karnowski, V. (2013). Evolving mobile media: uses and conceptualizations of the mobile internet. *J. Comput. Commun.* 18, 491–507. doi: 10.1111/jcc4.12019

Jovanović, V. (2016). The validity of the Satisfaction with Life Scale in adolescents and a comparison with single-item life satisfaction measures: a preliminary study. *Qual. Life Res.* 25, 3173–3180. doi: 10.1007/s11136-016-1331-5

Kwon, M., Lee, J.-Y., Won, W.-Y., Park, J.-W., Min, J.-A., Hahn, C., et al. (2013). Development and validation of a Smartphone Addiction Scale (SAS). *PLoS One* 8:e56936. doi: 10.1371/journal.pone.0056936

Ley, K., and Young, D. B. (1998). Self-regulation behaviors in underprepared (Developmental) and regular admission college students. *Contemp. Educ. Psychol.* 23, 42–64. doi: 10.1006/ceps.1997.0956

Lin, C.-Y., Imani, V., Broström, A., Nilsen, P., Fung, X. C. C., Griffiths, M. D., et al. (2019). Smartphone application-based addiction among iranian adolescents: a psychometric study. *Int. J. Ment. Health Addict.* 17, 765–780. doi: 10.1007/s11469-018-0026-2

Machimbarrena, J., González-Cabrera, J., Ortega-Barón, J., Beranuy-Fargues, M., Álvarez-Bardón, A., and Tejero, B. (2019). Profiles of problematic internet use and its impact on Adolescents' health-related quality of life. *Int. J. Environ. Res. Public Health* 16, 3877. doi: 10.3390/ijerph16203877

Mak, K.-K., Lai, C.-M., Watanabe, H., Kim, D.-I., Bahar, N., Ramos, M., et al. (2014). Epidemiology of internet behaviors and addiction among adolescents in six asian countries. *Cyberpsychol. Behav. Soc. Netw.* 17, 720–728. doi: 10.1089/cyber.2014.0139

Moè, A., and De Beni, R. (2000). Strategie di autoregolazione e successo scolastico: Uno studio con ragazzi di scuola superiore e universitari. *Psicol. dell'Educazione e della Form.* 2, 31–44.

Navarro, D., Montserrat, C., Malo, S., González, M., Casas, F., and Crous, G. (2017). Subjective well-being: what do adolescents say? *Child Fam. Soc. Work* 22, 175–184. doi: 10.1111/cfs.12215

Nota, L., Soresi, S., Ferrari, L., and Wehmeyer, M. L. (2011). A multivariate analysis of the self-determination of adolescents. *J. Happiness Stud.* 12, 245–266. doi: 10.1007/s10902-010-9191-0

R Core Team, (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Core Team.

Ringle, C. M., Wende, S., and Becker, J.-M. (2015). *SmartPLS 3*. Available at: http://www.smartpls.com (accessed September 09, 2019).

Sarstedt, M., Ringle, C. M., and Hair, J. F. (2017). "Partial least squares structural equation modeling," in *Handbook of Market Research*, eds H. Christian, K. Martin, V. Arnd, (Cham: Springer International Publishing), 1–40. doi: 10.1007/978-3-319-05542-8_15-1

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., et al. (1998). Development and validation of a measure of emotional intelligence. *Pers. Individ. Differ.* 25, 167–177.

Shoshani, A., Steinmetz, S., and Kanat-Maymon, Y. (2016). Effects of the Maytiv positive psychology school program on early adolescents' well-being, engagement, and achievement. *J. Sch. Psychol.* 57, 73–92. doi: 10.1016/j.jsp.2016.05.003

Soresi, S., and Nota, L. (2003). *La mia vita da Studente*. Firenze: ITER–Institute of Training Education and Research.

Stefansson, K. K., Gestsdottir, S., Birgisdottir, F., and Lerner, R. M. (2018). School engagement and intentional self-regulation: a reciprocal relation in adolescence. *J. Adolesc.* 64, 23–33. doi: 10.1016/j.adolescence.2018.01.005

Suwa, M., and Suzuki, K. (2013). The phenomenon of "hikikomori" (social withdrawal) and the socio-cultural situation in Japan today. *J. Psychopathol.* 19, 191–198.

Thomas, V., Muls, J., De Backer, F., and Lombaerts, K. (2019). Exploring self-regulated learning during middle school: views of parents and students on parents' educational support at home. *J. Fam. Stud.* 1–19. doi: 10.1080/13229400.2018.1562359

Tonioni, F., and Corvino, S. (2011). Dipendenza da internet e psicopatologia web-mediata. *Recent. Prog. Med.* 102, 417–420. doi: 10.1701/975.10605

Tran, B. X., Huong, L. T., Hinh, N. D., Nguyen, L. H., Le, B. N., Nong, V. M., et al. (2017). A study on the influence of internet addiction and online interpersonal influences on health-related quality of life in young Vietnamese. *BMC Public Health* 17:138. doi: 10.1186/s12889-016-3983-z

van Deursen, A. J. A. M., Bolle, C. L., Hegner, S. M., and Kommers, P. A. M. (2015). Modeling habitual and addictive smartphone behavior. *Comput. Human Behav.* 45, 411–420. doi: 10.1016/j.chb.2014.12.039

Verzeletti, C., Zammuner, V. L., Galli, C., and Agnoli, S. (2016). Emotion regulation strategies and psychosocial well-being in adolescence. *Cogent. Psychol.* 3:1199294. doi: 10.1080/23311908.2016.1199294

Xu, T.-T., Wang, H.-Z., Fonseca, W., Zimmerman, M. A., Rost, D. H., Gaskin, J., et al. (2019). The relationship between academic stress and adolescents' problematic smartphone usage. *Addict. Res. Theory* 27, 162–169. doi: 10.1080/16066359.2018.1488967

Yam, C.-W., Pakpour, A. H., Griffiths, M. D., Yau, W.-Y., Lo, C.-L. M., Ng, J. M. T., et al. (2019). Psychometric testing of three chinese online-related addictive behavior instruments among hong kong university students. *Psychiatr. Q.* 90, 117–128. doi: 10.1007/s11126-018-9610-7

Yang, S.-Y., Chen, M.-D., Huang, Y.-C., Lin, C.-Y., and Chang, J.-H. (2017). Association between smartphone use and musculoskeletal discomfort in adolescent students. *J. Commun. Health* 42, 423–430. doi: 10.1007/s10900-016-0271-x

Zeidner, M., and Olnick-Shemesh, D. (2010). Emotional intelligence and subjective well-being revisited. *Pers. Individ. Differ.* 48, 431–435. doi: 10.1016/j.paid.2009.11.011

Zhang, M. W. B., and Ho, R. C. M. (2017). Smartphone applications for immersive virtual reality therapy for internet addiction and internet gaming disorder. *Technol. Heal. Care* 25, 367–372. doi: 10.3233/THC-161282

Zhang, M. W. B., Tran, B. X., Huong, L. T., Hinh, N. D., Nguyen, H. L. T., Tho, T. D., et al. (2017). Internet addiction and sleep quality among Vietnamese youths. *Asian J. Psychiatr.* 28, 15–20. doi: 10.1016/j.ajp.2017.03.025

Zimmerman, B. J. (2002). *Achieving Self-Regulation: The Trial and Triumph of Adolescence*. California: ERIC.

Zimmermann, P., and Iwanski, A. (2014). Emotion regulation from early adolescence to emerging adulthood and middle adulthood: age differences, gender differences, and emotion-specific developmental variations. *Int. J. Behav. Dev.* 38, 182–194. doi: 10.1177/0165025413515405

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 31 March 2020
doi: 10.3389/fpsyg.2020.00565

# Remembering, Reflecting, Reframing: Examining Students' Long-Term Perceptions of an Innovative Model for University Teaching

*Giuseppe Ritella[1]*, Rosa Di Maso[2], Katherine McLay[3], Susanna Annese[4] and Maria Beatrice Ligorio[4]*

*[1] Department of Education, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, [2] Department of Humanities, Social Sciences and Cultural Industries, University of Parma, Parma, Italy, [3] School of Education, The University of Queensland, Brisbane, QLD, Australia, [4] Department of Educational Science, Psychology, and Communication Science, University of Bari, Bari, Italy*

OPEN ACCESS

***Edited by:***
*Maria Luisa Lorusso, Eugenio Medea (IRCCS), Italy*

***Reviewed by:***
*Zuhoor Abdullah Al-Khanjari, Sultan Qaboos University, Oman*
*Ioannis Magnisalis, International Hellenic University, Greece*

****Correspondence:***
*Giuseppe Ritella*
*giuseppe.ritella@helsinki.fi*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *29 November 2019*
***Accepted:*** *10 March 2020*
***Published:*** *31 March 2020*

***Citation:***
*Ritella G, Di Maso R, McLay K, Annese S and Ligorio MB (2020) Remembering, Reflecting, Reframing: Examining Students' Long-Term Perceptions of an Innovative Model for University Teaching. Front. Psychol. 11:565. doi: 10.3389/fpsyg.2020.00565*

This article presents a follow-up examination of 10 iterations of a blended course on educational psychology and e-learning carried out at the University of Bari. All iterations of the course considered in this study were designed using the constructive and collaborative participation (CCP) model. Our main research questions are: What are the students' long lasting memories of this course? How do the students use the skills and the competences acquired through the course across an extended period of time? In line with these research questions, the aims of this investigation can be summarized as follows: (i) to understand the students' perceptions and long lasting memories of the course and (ii) to investigate the transfer of skills and knowledge across an extended period of time, based on a self-reported survey. The analysis was carried out by administering the survey to all 196 students who took part in the course in the 2005–2015 decade. 96 participants responded to the survey. The survey is designed to collect data in two areas. First, the memories related to the course and second, the way skills and content knowledge acquired during the course have been transferred to and used in other contexts after the course ended. The data were analyzed using a mixed methods approach, which revealed trends in the responses across the decade. In general, participants remembered the teaching methodology and often recalled specific activities such as Role Taking and the creation of products through group-work. These activities and approaches seemed to provide significant learning opportunities for the students. Several students also recalled key concepts and content knowledge acquired during the course. In relation to transfer of skills, participants tended to reuse mostly transversal skills, such as communicative and organizational skills, especially in work contexts. Further, about half of the respondents reused the content knowledge of the

course. This analysis is valuable because it allows us to understand the aspects of the model that are significant for the students in the long term, and to discover and interrogate the acquisition and transfer of skills useful for the students' personal and professional lives beyond the academy.

**Keywords: innovative model, blended learning, collaborative learning, transfer, collaborative and constructive participation model**

## INTRODUCTION

The advent of digital technology is contributing to profound transformations in many spheres of life, requiring new literacy frameworks, and the development of novel practices of teaching and learning. Nevertheless, educational institutions have often failed to fruitfully transform their practices of teaching and learning (Hakkarainen, 2009) and to align with the evolving needs of the digital society. Some authors argue that this claim is also compelling for higher education, where universities and students have been facing challenges "in making 'good' use of digital technologies" (Henderson et al., 2015, p. 2). One example is the case of interactive whiteboards, which in many schools and universities are part of the normal technological toolset available for teachers and students. However, research shows that their usage is often limited to traditional learning practices that could be carried out on normal boards (Gursul and Tozmaz, 2010; Ritella and Sansone, 2020). Therefore, the capacity of educational systems to prepare learners for complex twenty-first century life and work is under increasing scrutiny. The importance of this point is emphasized by research showing significant interconnections between educational achievements, occupational success and well-being (e.g., Samuel et al., 2013). In 2017, The European Commission affirmed the need to strengthen European identity through education and culture, foregrounding the responsibility of educational systems for students' development of the necessary tools to thrive in the new paradigm of a knowledge society. These tools, which are often defined in terms of transversal competences or career and life skills (CLS), include: flexibility and adaptability; initiative and self-direction; social and cross-cultural skills; productivity and accountability; and leadership and responsibility (Kivunja, 2015). However, whether universities are employing teaching practices that help students to develop the transferable skills and competencies needed in a rapidly evolving society is questionable.

To address this issue, we argue that it is crucial to examine student perspectives. Indeed, students' interpretations and sensemaking concerning educational activities are crucial both for academic achievement (Schneider and Preckel, 2017) and for the transfer of skills and knowledge (Engle, 2006). For example, research shows that discursively framing a learning task in connection to broader contexts where students use what they learn may support the transfer of learning (Engle, 2006). In particular, the space-time context in which a learning task is situated seems to play a role in mediating the students' interpretations concerning learning situations (Ritella et al., 2017). Beyond the interpretations of single learning situations, students' interpretations of the implications of their academic activities for life after graduation need to be considered. A few large-scale studies show that most students often feel that their academic studies have not prepared them well for their professional life. For example, a McGraw-Hill Education (2016) of a large and heterogeneous sample of American university students found that only 21% feel very prepared to start their professional career. In Italy, the situation is similar. The Eighth Eurostudent Survey (2016–2018), focused on Italian university students, provides a problematic assessment of students' professional knowledge acquisition. While four out of five are satisfied with their theoretical preparation, less than 50% feel professionally prepared. Thus, there is a need for higher education to better fulfill its *trait d'union* function between educational and professional spheres, helping students to develop "the skills most in demand in the 21st century workplace" (Kivunja, 2015, p. 1). This involves redefining the aims of higher education and reevaluating strategies to accomplish these aims. This is a complex and multidimensional task that arguably involves challenging and undoing decades of what many scholars characterize as misguided policies that have, across the globe, cultivated homogeneity, standardization and "testing over teaching" (Zhao, 2015, p. 129) – qualities that are at odds with the CLS necessary to flourish in twenty-first century life and work.

Effective psycho-pedagogical models are widely regarded as key to enhancing learning outcomes for contemporary university students (Schleicher, 2011). This involves "creating environments and feedback mechanisms and systems to allow students' views, learning experience, and their performance to be taken into account" (European Commission, 2013, p. 28). For example, this might be the case in blended and flipped approaches (Wanner and Palmer, 2015); cooperative learning in small groups (e.g., Smith et al., 2005; Gillies, 2007); and in approaches that actively involve students in teaching and learning processes, encouraging student contributions to coursework planning (see Cook-Sather, 2002; European Association for Quality Assurance in Higher Education [ENQA], 2005, 2015). Kelly et al. (2014) contend that learning and teaching processes should be developed in parallel, looking for intersection points between how to teach and how to learn. Furthermore, tools and technologies deployed in the classroom are important; an effective teacher can cultivate a rich, stimulating and appropriate environment for students by knowing how to choose the right tools and methods (Lucena et al., 2018). The role played by digital tools in collaborative tasks is particularly important as mediating tools can significantly affect the collaborative sensemaking of the groups (Ritella and Ligorio, 2016) and, through participation, the identity construction of learners (Annese et al., 2010). However, while it has long been recognized that the new paradigm of twenty-first

century life and work demands approaches that provide learners with more autonomy (Steeples et al., 1994) such as those described here, educational innovation in university settings is not easy to implemente.

Tadesse and Gillies (2015) argue that innovating to achieve "instructional conditions that promote quality learning [is] challenging for many higher education teachers" (p. 1) for whom traditional lecturing remains the most common approach. Some research indicates that innovation is only possible when both political and corporate stakeholders are involved and aligned (Etzkowitz, 2008; Mowery et al., 2015). Others contend that university teaching can be improved by focusing on training skills and supporting teaching with purposely trained tutors (Muukkonen et al., 2005; Sansone et al., 2016). Through a systematic review of 38 meta-analyses investigating 105 correlates of achievement, Schneider and Preckel (2017) discussed the variables that are associated with achievement in higher education. In their account, the mere presence of technology has little effect compared to other variables such as social interaction, meaningful learning, and assessment. This reflects research showing that digital tools do not automatically improve educational practices, nor affect learning by themselves (Säljö, 2016). The effects of technology on education also depend on how tools are integrated into practice. In particular, assessment practices are crucial for any significant shifts in university teaching structure. Indeed, when educational practices change, a robust evaluation system is needed, able to account for different dimensions, including the various elements of a course (Gatignon et al., 2002). However, the typical elements that are usually considered when discussing teaching methods are not able to ensure high quality teaching practices. Indeed, one of the most interesting findings of the review conducted by Schneider and Preckel (2017) is that strong moderator effects were found for all the teaching methods considered. This indicates that how a method is implemented impacts on achievement. For example, "[t]eachers with high-achieving students invest time and effort in designing the microstructure of their courses, establish clear learning goals, and employ feedback practices" (Schneider and Preckel, 2017, p. 565). Going beyond examining general features of the teaching methods and technological aspects is therefore valuable because interrogating the details of course design (e.g., orchestration of activities, tools, and strategies) can support deeper understanding and more successful implementation of future iterations.

In response to these challenges, this paper first describes in detail the organization of a course based on the CCP model. The perceptions and memories of students, especially concerning the capacity of the course to foster the transfer of CLS across multiple professional and personal contexts, is then explored. In particular, the analytical focus is on the long-lasting memories and long term effects of the course concerning the transfer of competences and skills. In other words, our goal is to discuss if and how the course under scrutiny has had any significant impact on the students' lives in the long term, and how the students remember it several years after completion. We argue that current literature on the topic tends to focus on the short term and is mainly based on data collection carried out immediately after the completion of a learning experience. To our knowledge, there are not existing surveys specifically designed to examine these issues on timescales beyond a few months. For example, research on students evaluation of teaching (SET) is usually based on surveys carried out immediately after the end of the course. This approach is not well suited to examining the long lasting memories that the students can retrieve several years after completion of a course, nor the long term transfer of competences. Similarly, the survey developed by Maul et al. (2017), which allows examination of how students perceive their school experience as "connected" with their interests and with their life out of school, is meant to be administered when the students are participating in the learning programs investigated. In addition, although this survey addresses issues relevant to our study, it is not designed for higher education. Our review of the existing literature led us to develop a survey specifically designed for our study (**Supplementary Appendix 1**).

In order to address this research gap identified in the literature, our analysis examines the elements of the course that the students remember as most significant several years after the completion of the course and investigates the role of the learning environment in developing and supporting the re-use of both soft and professional skills beyond graduation. The elements retained by the students, as well as the self-reported transfer of skills to other contexts in the longer term, can be considered crucial dimensions to evaluate course design based on the CCP model.

## THE MODEL OF CONSTRUCTIVE AND COLLABORATIVE PARTICIPATION

The students under scrutiny in this research have participated in a university course titled "Psychology of education and e-learning," offered at the University of Bari, in Italy. The course takes a blended approach using the collaborative and constructive participation (CCP) model, developed over more than 10 years of consecutive application to several higher education courses (see Ligorio and Sansone, 2009; Ligorio and Annese, 2010; Ligorio and Cucchiara, 2011). The CCP model conceives learning as the co-construction of knowledge and aims to support students to develop new ideas through the creation of both individual and group products (Cucchiara et al., 2014). Drawing on social constructivism (Kelly, 1955; Berger and Luckman, 1966; Shotter, 1993; Potter, 1996; Gergen, 1999, 2001; Scardamalia and Bereiter, 2006), this course requires students to build knowledge through actively producing meaning, products and forms of interaction, negotiation, and social collaboration. Thinking is not considered as a private or individual process; rather, it is distributed or "stretched across" people and the environment, artifacts, and technological tools mediate the relationship between individuals and the learning context (Suchman, 1987; Lave, 1988; Hutchins, 2001; Ritella and Hakkarainen, 2012). Learning is a complex system in which the relationship between the subject and the object is mediated by artifacts (Engeström, 1987) and by social factors such as teamwork and collaboration (Dillenbourg, 1999) that occur both online and offline (Graham, 2006). The social dimension is crucial to this process; people learn through

interaction with other members of a community (Brown and Campione, 1990; Wenger, 1998). Grounded in these theoretical underpinnings, the course is comprised of alternating online and offline activities, which are distributed across five or six modules, lasting an average of about 10 days each. Each module begins with face-to-face lectures, during which the teacher introduces the content of the module, and ends with the students jointly negotiating a research question that guides all subsequent module activities. The modular structure allows for easy comparison across cohorts, with each module operating as a milestone within the course. In the design, implementation and evaluation of the course, the model considers the intervention of purposely trained professional tutors, who act as mediators between the teacher and students. The course activities can be summarized as follows.

## Independent Individual Activities

These are activities that students can perform alone, without the support of other peers. Individual activities support self-evaluation strategies, including metacognitive reflection upon what has been done and what to do next. Two activities are included in this category:

- An individual e-portfolio which includes personal information and material such as photos, reflections, links to Facebook pages and blogs. At the end of each module, students include this material in their e-portfolio, record what they feel they have learned, and outline their goals for the next module. These two latter aspects of the e-portfolio are based on the concepts of the actual zone of development and the proximal zone of development conceived by Vygotsky (1986). This tool is discussed in more detail in a later section.
- Compilation of a self-assessment grid. At the end of each module, students identify the specific and transversal skills acquired. The grid is comprised of questions aimed at developing critical self-assessment and recognition of the skills learned.

## Individual Interdependent Activities

These are activities performed individually within a group. They are designed to support individual responsibility within a social context and provide a structure for student social participation and include:

- Writing a review. Groups of students called "expert groups" are formed. Within these groups, to each student is assigned a specific learning material (e.g., a book chapter, a scientific article, a website, etc.). Students can discuss the material but are ultimately individually responsible for writing a critical review (e.g., summary of key information, strengths and weaknesses, etc.) to inform the subsequent group discussion. Each review is an individual but interdependent activity because, once the reviews are complete, the expert groups are dissolved and new groups are formed. Students from the different expert groups now form a new group where they use the reviews to build a shared answer to a previously negotiated research question. This activity is inspired by the Jigsaw model (Aronson et al., 1978; Aronson and Patnoe, 1997) and it is appropriately adapted to the blended nature of this course by including web-forum discussions and online group work.
- Role Taking. Inspired by the work of Strijbos and Weinberger (2010), Role Taking requires each group member to take on a specific tasks and responsibilities (Hare, 1994), aimed at supporting individuals to achieve a shared objective (Topping, 2005). Student participation is clearly structured to (i) improve individual satisfaction (Zigurs and Kozar, 1994); (ii) empower students with a sense of individual responsibility; (iii) support group cohesion; (iv) stimulate awareness of the interational processes (Mudrack and Farrell, 1995); and (v) support group dynamics. Examples of roles implemented within the course and assigned in turn to students are:
    - E-tutor: The e-tutor coordinates the group discussion, manages times and spaces (when and where the group will meet), and monitors the development of the other roles. The student who performs this role must have a clear understanding of the objectives of any group discussion and of the related tasks. The e-tutor becomes the temporary leader of the group and must deploy suitable communication strategies to stimulate collaboration. This role is designed to (i) keep discussion focused on common objectives; (ii) monitor deadlines; (iii) be aware of the functions and affordances of the virtual space; (iv) regulate possible conflicts between group members; (v) manage unexpected events; and (vi) balance focus on the task with attention to relationships.
    - Synthesizer (S): The synthesizer's role is to summarize the group discussion. This metacognitive role requires the student to analyze and describe the group dynamics and methods of discussion rather than engaging with subject matter. In particular, the synthesizer considers how discussion progresses, from facts and data to ideas and knowledge building. Students who take on this role develop the skills to carefully and critically review discussion, as well as the capacity to identify and manage the dimensions that can help or hinder the group's progress.
    - Product Manager: The student in this role manages and monitors the process of building the collaborative products (further discussed in the following section). Taking this role requires the student to develop their capacity to coordinate and supervise group work to successfully develop a collaboratively designed and built product.

Students rotate through these roles from one module to the next so that all students experience as many different roles as possible. Through Role-Taking, ways of participating that would not otherwise be experienced are encouraged. Students are assigned to the roles randomly.

## Small Group Activities

These activities are organized so that success can only be achieved if students work collaboratively in their small groups. There are two types of activities in this category:

- Jigsaw group activities. Students are first required to read all the critical reviews produced by the participants and then to discuss them, searching for connections to their research question. Discussion usually occurs online via web-forum but can also be interspersed with face-to-face discussion, if the group prefers. The progressive inquiry model (PIM) (Hakkarainen and Sintonen, 2002) guides both expert and Jigsaw group discussions. Within this model, learning is conceived as a process of investigation, which begins with a large and general question – in our case, the research question underpinning the module. From here, the focus shifts to critical assessment of the various dimensions of the issue under discussion. To this end, students are encouraged to search for further material that will develop and deepen the problems that have emerged. Critical thinking is encouraged by comparing different ideas and divergent material. The aim is to distribute cognition among all participants.
- At the end of the discussions within the Jigsaw groups, students are required to collaboratively create a joint product. This can be a text summary, a concept map or any product that the group considers suitable for the content.

## Plenary Activities

Plenary activities aim to involve all course participants by interconnecting the groups to produce a collective product. All groups are required to build an object that synthesizes what that group has learned. To do this, we often used a grid of indicators that made visible the salient features of e-learning courses. First, all the groups jointly singled out the crucial dimensions of e-learning (i.e., modalities of collaboration, features of the platforms, etc.). After that, each group takes up one dimension and identifies aspects of the course in which that dimension is evident. Finally, all the indicators are compared and organized into a grid that guides observations of e-learning courses in ways that support understanding of its features.

Other approaches to a final product could include a text, a concept map, or a multimedia product summarizing the content of the course. In any case, the plenary work should allow the transfer of knowledge and skills acquired within the smaller groups to larger ones, creating dynamics of collaboration, and meta-reflection on activities, as well as creating space for students to grapple with diverse perspectives.

# ONGOING ASSESSMENT OF THE COURSE

As discussed above, assessment is a core feature of innovative educational practice, which has a strong impact on students' achievement. Therefore, the CCP model involved a careful design of the assessment process. A substantial proportion of research into assessment calls for students to be active participants so they can understand what and how they learn, and to test the efficacy of teaching. To achieve both these aspects, two types of assessment have been considered in the design of the course: assessment for learning (AfL) – also known as formative assessment, assessment as learning, and learning-oriented assessment (Carless, 2016) – and self-evaluation. AfL can be distinguished from assessment of learning or summative assessment, which focuses on grading or marking student work. Promoting student learning rather than grading is the first priority of AfL; this is achieved by helping students reflect on what they know and what they can do, then use this understanding to identify gaps in their knowledge, connect concepts, and face new problems (van Dinther et al., 2015). AfL is a multi-dimensional vision of assessment that conceives cognitive, emotional, affective and social aspects as integral elements of the learning process, and seeks to make visible students' thinking and reasoning skills using a variety of tools (Astleitner, 2018). Two of the most widely used tools for formative assessment are e-portfolios and case-based assessment.

The second type of assessment, self-evaluation, has been the object of some scholarly disagreement. According to Topping (2003), self-evaluation should be considered a formative assessment technique, while others conceive self-evaluation as a tool for encouraging students to take responsibility for their own learning (Brown and Harris, 2014). Either way, self-evaluation requires students to assess aspects of the learning process (for example, as members of a group) they are involved in and of the products they build (for example, the task they are working on). Zimmerman (2001) considers self-evaluation a key component of the broader ability to self-regulate. Self-evaluation is closely connected to the capacity to manage learning processes, as well as to meta-cognitive and motivational aspects through which personal skills are deployed to control learning outcomes. Our contribution takes into account both approaches. In the following section, we describe and analyze data generated through longitudinal follow-up study "cohorts" – groups enrolled in the same course but in different academic years.

Given these premises and definitions, the CCP model approaches assessment to sustain students' reflection upon their own learning journey and on how they learn most effectively. We already reported how during the course, self-assessment and peer-assessment (Topping, 2005) are facilitated through specific tools and interactive moments. The goal is to equip students with critical skills in terms of their learning methods, participation in group-work and, ultimately, their results. The main goal is to cultivate self-monitoring – a critical twenty-first century competency that both empowers and supports life-long learning. The tools used are:

- The e-portfolio. As described earlier, at the end of each module students are asked to review what they learnt and set personal goals for the upcoming modules. At the end of the course, the e-portfolio provides students a concrete artifact of their learning journey and helps them identify knowledge, skills, and attributes for inclusion in their curriculum vitae or professional social networks

profiles, such as LinkedIn. The e-portfolio is deliberately structured to promote self-regulation and self-assessment. Cultivating student awareness of their current and proximal zones of development promotes autonomy and a sense of agency (Bruner, 1996). This helps students focusing less on performative aspects, like summative assessment, and more on evaluating the strategies they have deployed to achieve their goals. Further, explaining the criteria they used to select skills and competences for inclusion in the e-portfolio supports self-assessment (Brown and Harris, 2014).

- A self-assessment record. At the end of each module, students, teacher and tutors record all course activities (individual reviews, collaborative products, role taking, and discussions, etc.). Students must score each activity on a scale from 1 to 5 in relation to (i) how much they believe the activity supported their learning of content and skills, and (ii) the factors they believe contributed to the success of the activities. The aim is to stimulate students' meta-cognitive processes of critical reflection on their own abilities and performance. The teacher and the tutors also fill in the same sections, so students can compare their perspectives with their own self-assessment. By looking at the score obtained across the modules, students can trace their evolution throughout the course.

At the end of each module, students are invited to read the assessment record and comment on it in a dedicated forum. This forum is an opportunity to "wonder," reflect, and appreciate achievements as well as to express doubts and ask questions in an extended discussion between peers, professional tutors and the teacher. This type of discussion aims to help students improve self-assessment, adjust their learning and participation strategies, taking them closer to self-regulation.

## RESEARCH QUESTIONS

Our main research questions are:

1. What are the students' long lasting memories related to this course?
2. How do the students use the skills and the competences acquired through the course across an extended period of time?

In line with these research questions, the aims of this investigation can be summarized as follows: (i) to understand student perceptions and long lasting memories related to the course; and (ii) to investigate the transfer of professional skills and knowledge across an extended period of time, based on a self-report survey. We were also interested in whether the students' perceptions vary over time. To this end, all the students who took the course in the decade 2005–2015 were considered. During this decade, the structure of the course remained almost unchanged, making these iterations of the course comparable. The decision to include several iterations of the course across a decade also supported answering our research questions. Interrogating students' long lasting memories and the transfer of learning over an extended period of time justifies the inclusion of students who had completed the course three to 13 years before data collection. We considered the iterations of the course held at least 3 years before because we expected that this span of time would allow the consolidation of long term memories and it would also allow the students to experience several educational or professional contexts in which to use the skills, knowledge and competences acquired during the course. In addition, the longitudinal dimension helps us to examine the long term effects of some actions related to group management (discussed in the section "Domain 2: Skills" of this article) that were implemented during the course to improve group dynamics.

## PARTICIPANTS AND DATA COLLECTION

Data collection was carried out through an online questionnaire using Google forms (see **Supplementary Appendix 1**). The questionnaire link was sent to all 196 students who had completed the course. 96 students (49% of the sample) answered the questionnaire. Of these, 81 were female and 15 were male. This composition reflects the gender distribution within the course, which is part of a Master of Psychology that, in the Italian context, attracts more females than males. The unbalanced gender distribution did not represent a concern for the design of our study because we did not plan to examine gender differences. The average age was 28 years. The number of participants over the 10-year period has steadily increased each academic year, as shown in **Table 1**.

The questionnaire was constructed in stages. First, a draft was prepared by the lecturer in consultation with two experts in designing questionnaires. The draft version was administered to 10 students randomly chosen from those enrolled in the course in the last three academic years. Immediately after completing the questionnaire, the teacher and experts met one-on-one with these students to seek their feedback. Specifically, participant students were asked to explain how they interpreted each item and whether they could suggest improvements to the questionnaire. After collecting this feedback, a second version of the questionnaire was prepared in consultation with the two experts. The second version was also shared with the course

**TABLE 1 |** Number of participants and respondents per academic year.

| Academic year | Number of participants | Number of respondents |
|---|---|---|
| 2005/2006 | 10 | 2 |
| 2006/2007 | 10 | 3 |
| 2007/2008 | 16 | 6 |
| 2008/2009 | 14 | 9 |
| 2009/2010 | 25 | 10 |
| 2010/2011 | 19 | 8 |
| 2011/2012 | 16 | 7 |
| 2012/2013 | 17 | 7 |
| 2013/2014 | 39 | 30 |
| 2014/2015 | 30 | 14 |

tutors. This resulted in further comments and suggestions that led to the final version of the questionnaire.

The final version of the survey consisted of two domains: The first focused on memories of the course, and the second focused on the reuse of the skills and knowledge acquired during the course in other contexts. To investigate memories, an open question was used to encourage free expression. In order to investigate to what degree skills and knowledge had been reused, we constructed 27 items to which participants responded on a three-point Likert scale (1 = little, 2 = enough, and 3 = a lot). Finally, to investigate the contexts in which skills and knowledge were reused, we deployed two multiple-choice questions asking participants to indicate the online and the offline contexts where they had mostly re-used the skills and knowledge acquired through the course.

## DATA ANALYSIS

The questionnaire responses were first analyzed considering the whole corpus of data; that is, all the academic years involved. This provided a broad overview of the students' perceptions. Subsequently, the questionnaires were grouped into the following three clusters: (i) academic years 2005–2006, 2006–2007, and 2007–2008; (ii) academic years 2008–2009, 2009–2010, and 2010–2011; (iii) academic years 2011–2012, 2012–2013, 2013–2014, and 2014–2015.

A mixed methods approach was deployed. Multiple-choice and Likert scale items were quantitatively analyzed by calculating frequencies and percentages. For this purpose, the 27 items about skills and content have been grouped into the six following categories:

- Organizational skills (how to work toward a common goal; how to manage work deadlines).
- Communication skills (how to communicate effectively during collaborative work; how to participate in the creation of collaborative products).
- Managing group dynamics (how to negotiate between different points of view; how to observe group dynamics).
- Academic skills (how to write academic texts; how to find useful and reliable material online).
- Content about e-learning (how to assess online courses; how to operationalize theoretical constructs related to e-learning).
- Self-assessment skills (how to enhance their own skills; how to assess their own learning content and strategies).

The open question on memories was analyzed by building a system of categories using the Grounded Theory methodology (Glaser and Strauss, 1967). First, all responses were analyzed by two researchers who worked independently to code the data without predetermined categories. After independent analysis, the coders compared the categories that they identified and 90% agreement was reached. The remaining categories were discussed with a third researcher until 100% agreement was reached. Each memory has been segmented by isolating specific units of content and the coding was applied to each segment. In this way, more than one category could be assigned to each answer.

The categories identified are described in **Table 2**, including examples from the data corpus for each category. The three independent researchers reached unanimous agreement to confirm each code assignment. Once the system of categories was defined, an analysis of the frequencies and percentages on the whole sample was carried out. Later, the three clusters previously formed were compared to understand differences between them.

## RESULTS

The analysis of students' memories is useful to identify key aspects of the course that participants recall and the meanings that the students associate to them. The examination of responses concerning skills and knowledge, and the contexts of their reuse, helps us identify participants' perspectives about which skills were transferred to other contexts and where they were re-used. Together, these two aspects provide rich insight

**TABLE 2 |** The category system for the analysis of memories about the course.

| Category | Description | Examples |
|---|---|---|
| Teaching methods | References to the educational models used i.e., activities, group work, individual work, Role Taking, objects produced | *I remember a particularly interactive and dynamic course, where the groups were the privileged places of exchange, comparison and construction of new knowledge through the enhancement of the various roles covered by each member (as tutor, map manager), alternating moments of work online to moments for face-to-face discussion and collaboration* |
| Technological devices | References to the tools and technologies used during the course | *I remember we used any type of online tool; chat, virtual whiteboards, videoconferences, forums, ebooks* |
| Skills | References to soft and professional skills acquired during the course | *In addition, critical thinking, the ability to synthesize, to collaborate and to construct knowledge have become my skills* |
| Group dynamics | Memories related to processes and group dynamics | *Group dynamics, cooperation, collaborative learning, respect for group norms, exchange and relationship* |
| Educational content | Memories of syllabus content | *I remember the definition of e-learning and how to go from theory to practical implications* |
| Generic comments | Generic memories of the course | *Excellent course. Innovative, motivating and creative* |

into the student experience of this course, as well as into the perceived effectiveness of the course design in fostering the transfer of knowledge and skills. Further, these data can also be used to improve the next iterations of the course. In the following sections, we first present results concerning participants' memories of the course. In the subsequent section, we report the data in relation to learning, skills and re-use. In both sections, we consider the totality of the sample in the first instance, followed by an analysis of the three clusters. For an overview of the data see **Supplementary Appendix 2**.

## Domain 1: Participants' Memories of the Course

The analysis of the participants' memories of the course shows that, beyond a series of generic comments on the course (26%), the participants mostly remember the teaching methodology (35%), and the educational content (17%). The technological tools used (7%), the skills acquired (7%) and the group dynamics (8%) (see **Figure 1**) reached lower percentages and are more or less at the same level. Within the category "Teaching Methodology," many have qualified the method as innovative and exciting, as we can see from the following excerpts:

> Excerpt n. 1: "*The learning modality: absolutely innovative, interactive, stimulating. The content were learned using them in the online platform context*."
> Excerpt n. 2: "*I remember the innovative way of conducting the course and the exam, the collaborative atmosphere, the blended approach, the scaffolding of the tutors and the challenging objectives of the course."*

Considering that this was the first experience of blended courses during their academic studies, this is not surprising. Nevertheless, the sense of innovation that has been perceived by the students might have contributed to their learning. Indeed, it is known that novelty can have multiple effects on cognition, including the enhancement of learning (Schomaker and Meeter, 2015). A close analysis of this category indicates that the activities of Role Taking (RT) (27%) and the creation of products through group work (22%) were particularly memorable to participants. With regard to RT, this approach appears to have elicited emotional and cognitive involvement that stimulated students to activate new forms of reasoning and interaction, as we can see in the following extracts:

> Excerpt n. 3: "*I perfectly remember the importance of the roles that, in my opinion, have made this course closer to us students making us feeling really part of a community. Having covered several roles has allowed me to experience the different dynamics that can emerge within an online group and to understand both the negative and the positive aspects."*
> Excerpt n. 4: "*I remember that there was the role of the critical friend, whom I liked very much and it had greatly stimulated my reflection."*

The role of the Critical Friend, mentioned in excerpt n. 4, has the task of offering constructive criticism on another student's products. This role is designed to promote critical thinking and argumentation, as well as relational skills and the capability to empathize with the author of the product. We suggest that the design of tasks based on the RT approach might have been one of the elements characterizing the course design for the students and contributing to the students' learning achievement. In particular, the memories of the students might signal that RT supported the development of a range of inter and intrapersonal skills that are still considered relevant and useful for them at the moment of the data collection, which took place several years after the end of the course.

Concerning the collaborative activities, the participants most often cited the construction of the conceptual maps. Participants saw this activity as valuable because it promoted skills related to effective summarizing and synthesis:

> Excerpt n. 5. "*The construction of maps, the importance of syntheses and discussions with others."*
> Excerpt n. 6. "*I remember we were working on collaborative conceptual maps in every module, which for me was a very important activity for summing up the content of the module."*

When we compared the results across the three different clusters, we detected a similar trend in all clusters although there were slight differences, as shown in **Figure 2**. From this figure, it is clear that the course content was best recalled by the 2012–2015 cluster, probably because this is the most recent one. Nevertheless, the memories reported by the students from the clusters 2005–2008 and 2009–2011 show that at least some of them have a clear memory of some of the concepts that were addressed during the course, such as virtual communities, knowledge building, interactivity, blended learning, etc. Below we report two excerpts that were categorized with this code:

> Excerpt n. 7. "*I remember the concept of collaborative learning and knowledge building."*
> Excerpt n. 8. "*Virtual communities of learning, interactivity, learning by doing, multimedia, blended learning, virtual classrooms, semantics web."*

## Domain 2: Skills

The analysis of skills and knowledge acquired during the course, reported in **Figure 3**, shows that on average 70% or more of the students re-used communicative, organizational and self-assessment skills, while the transfer of academic skills and skills related to collaboration and group dynamics is reported as high by approximately 66% of the students. Instead, the 47% of the students declare to have reused knowledge and content concerning e-learning. The answers in this domain suggest that the course provided to a high percentage of occasions for developing transversal skills that they have found useful in their subsequent career after graduation, while the specific content of the course has been reused by about half of the students. This might also be related to the different career paths started by the students, since some of them might have chosen careers in other fields of psychology where e-learning is not necessarily relevant.

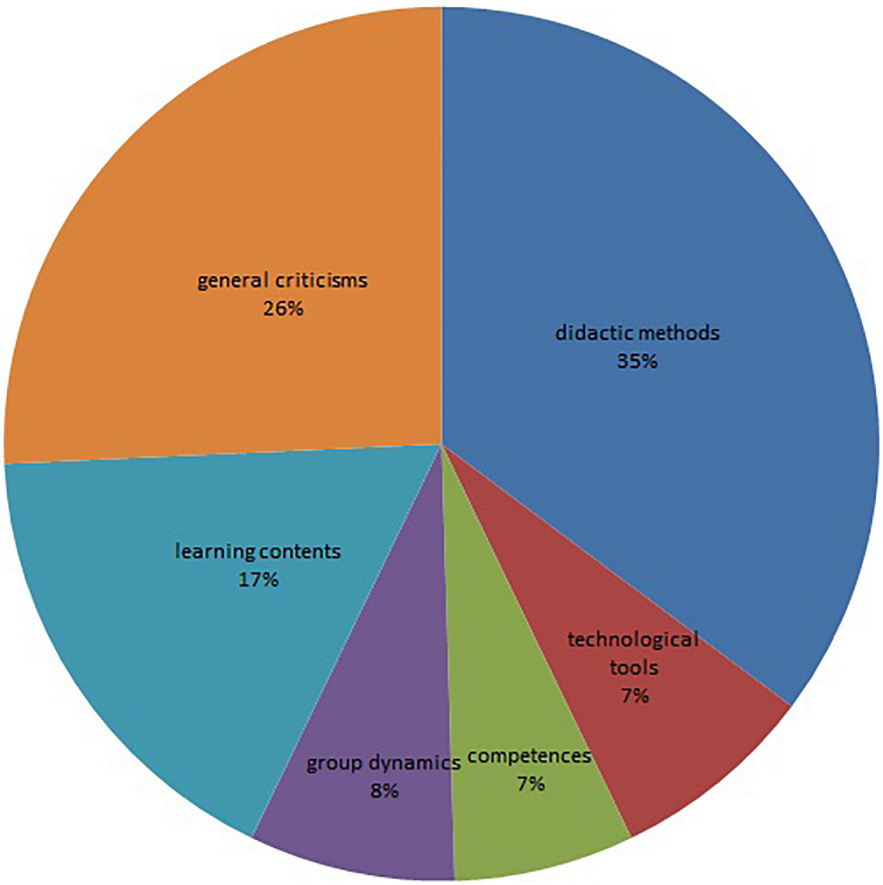


**FIGURE 1 |** Distribution among the whole sample. Categories of course memories.

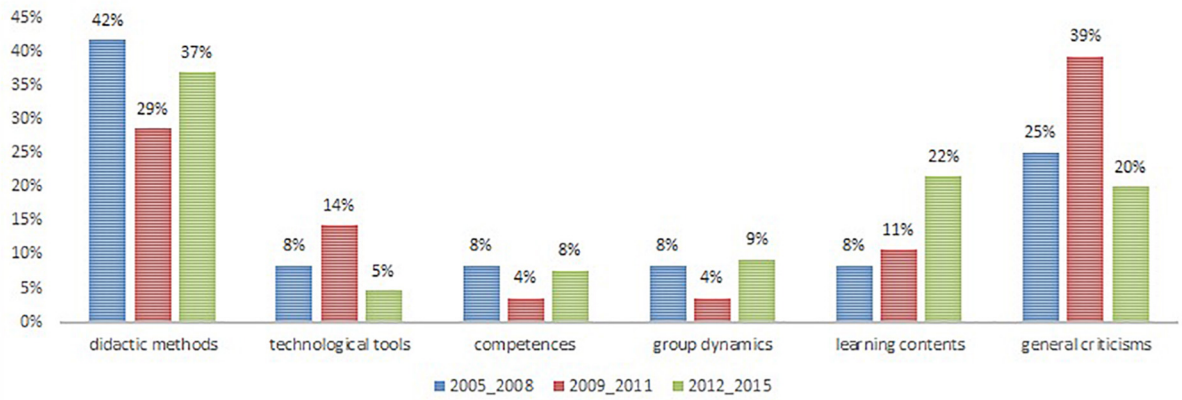


**FIGURE 2 |** Distribution comparing clusters. Categories of memories across the clusters.

The skills with the highest percentages of responses relate to effective group communication (73%) and to the ability to participate effectively to create collaborative products (72%). A close examination of these categories reveals that most participants have re-used the ability to be flexible during problematic situations that arise in the group (73%), which is a subset of communication skills considered in the survey. The organizational skills that are reused most frequently are related to the organization of group work toward a common goal (76%). Finally, the self-assessment skill with high frequency of transfer concerns the ability to know how to exploit and maximize one's own skills (69%).

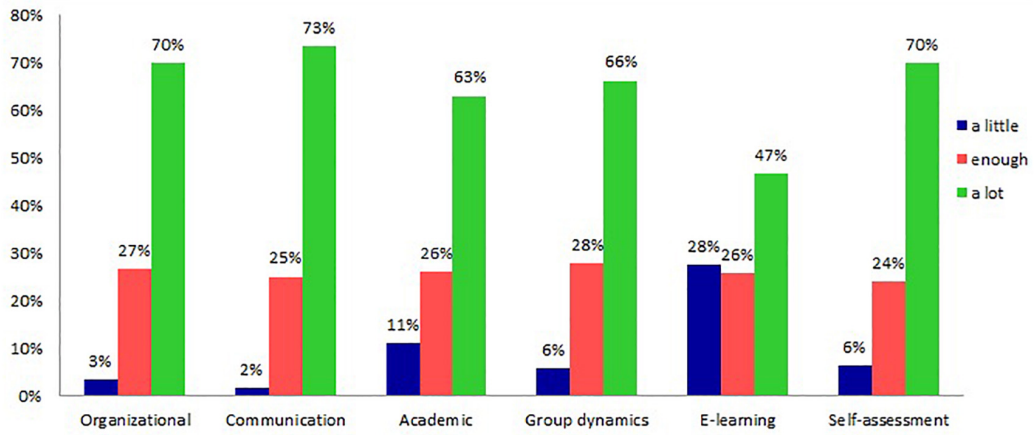


**FIGURE 3 |** Distribution of categories with higher score. Reused skills and knowledge acquired during the course.

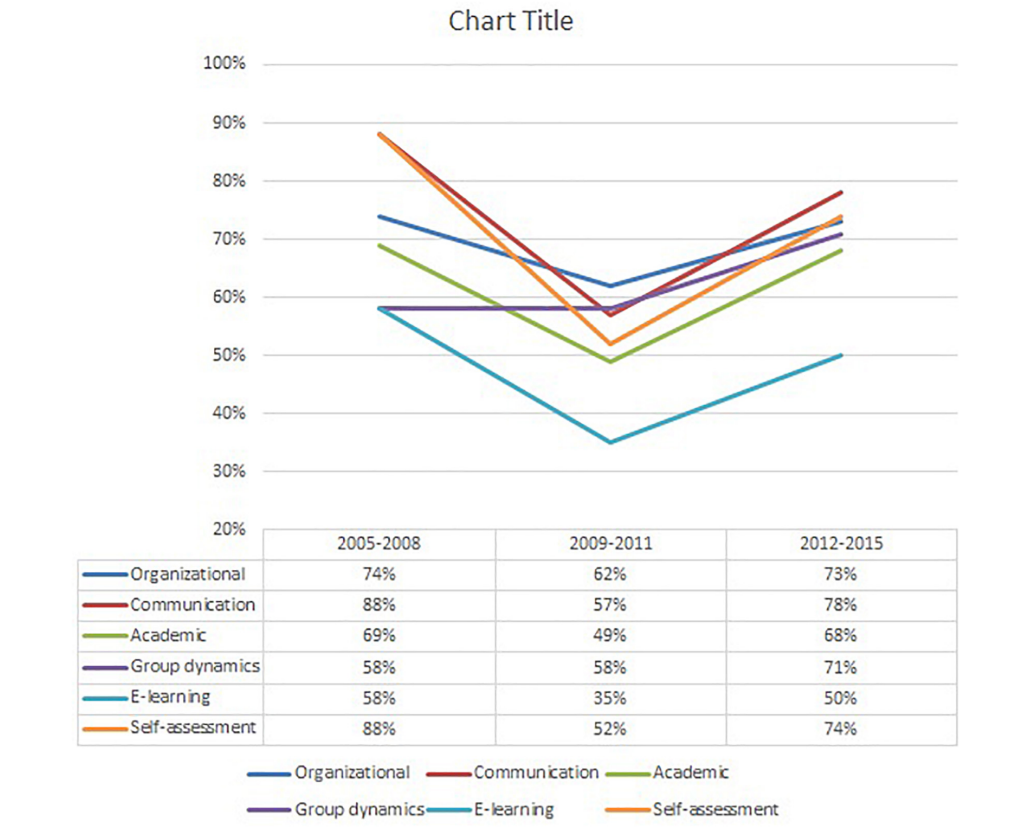


**FIGURE 4 |** Distrubution of the higher score across clusters. Skills reused divided by clusters.

**Figure 4** represents the percentage of respondents who answered "a lot" for each type of competences across the three clusters (2005–2008, 2009–2011, and 2012–2015). As shown in this figure, the participants in the first two clusters reported Organizational, Communicative and Self-assessment skills as the most often reused; this is in line with what emerged from the analysis of the whole corpus of data. The second cluster seems to show a drop in all dimensions apart from the skills related to Management of group dynamics, which in the third cluster increased to 71%. One possible interpretation of this result is that throughout the 10 iterations considered in this study, the teacher carried out specific actions to improve group management, based on previous experience (Annese and Traetta, 2012). Indeed, the management of group dynamics was not easy in the first

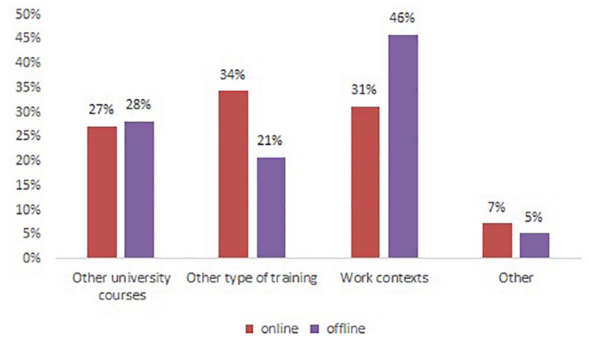


**FIGURE 5 |** Skills reuse comparing online and offline contexts. Contexts of use of skills.

iterations of the course and the teacher and tutors dedicated extra effort and time to understand how to improve this aspect in future courses. This is well reflected in the data.

Looking at the contexts where the students reported to have reused these skills, almost half the participants (46%) reused them within offline work contexts (**Figure 5**). This finding seems to suggest that this course allows a learning experience highly relevant for work contexts. Instead, only 34% of the students reused them in other contexts of online training.

Comparing the data across the three clusters (see **Figure 6**), it became evident that for the first and second cluster the acquired skills were mostly reused in work contexts, while for the last cluster the most relevant contexts for re-use were other university courses. This might be interpreted considering that the students belonging to the first two clusters have completed their academic studies earlier than the others and probably had more occasions to reuse their skills at work. The students belonging to the last cluster instead might be still enrolled in a university degree or searching for their first job. Indeed, in the local context of South Italy – characterized by high unemployment rate among young people, even those with a master's degree – the transition from higher education to the working career might take several years. Therefore, the differences among the clusters seem to capture different dimensions: the first two clusters allow the examination of the transfer of learning from academia to professional contexts, while the last cluster seems to generate more insights about the transfer across multiple learning contexts experienced by the students. Unfortunately, we did not ask to the students any question about their working career, thus this hypothesis needs to be confirmed by future research. Moreover, for the first and third cluster, the reuse of competences takes place more offline (58 and 36%, respectively), than to online courses (41 and 18%). For the second cluster, online and offline working contexts show similar percentages of reuse.

## DISCUSSION

This paper reports our analysis of students *post hoc* perceptions and memories related to a blended university course based on the constructive and collaborative participation (CCP) model. The goal of this research was to understand the students' perceptions of the transfer of skills learned through the course across an extended period of time, as well as to identify the elements of the course considered significant by the students and thus retained across time. To do this, we sought feedback from former students who had undertaken the course over a 10-year period. The results were generated through a mixed methods approach, in line with other research studies that deploy a mixed methodology for research into blended learning communities (Annese and Traetta, 2011, 2018). The CCP model is based on principles of socio-constructivism and applies a knowledge building approach to teaching and learning. The CCP

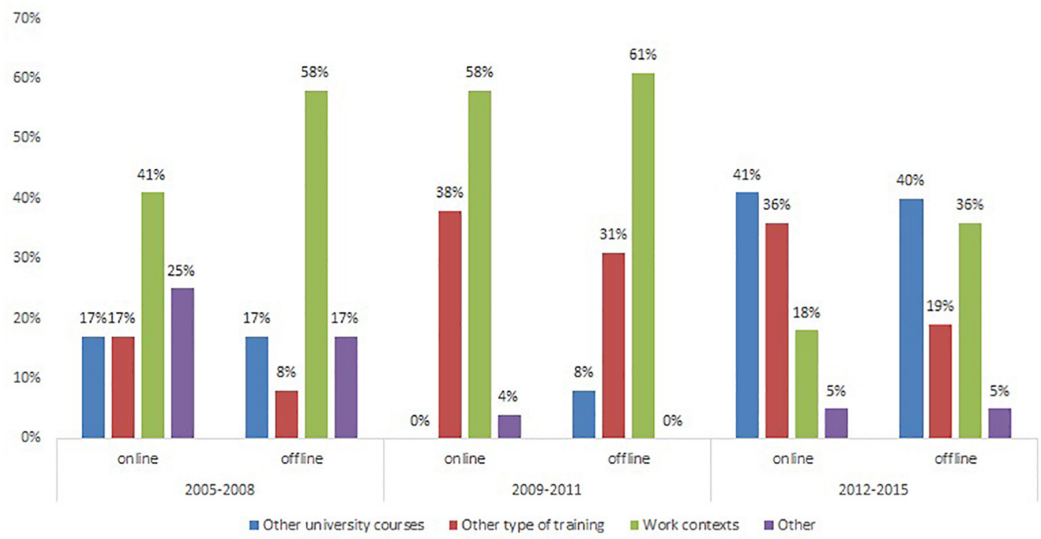


**FIGURE 6 |** Skills reuse comparing online and offline contexts across clusters. Contexts of re-use divided by cluster.

model also aims to support students to acquire CLS through productive use of several technological tools. These skills include not only academic abilities but also non-cognitive and personal skills. In this course, academic skills such as analyzing and constructing texts are supported through the task of producing critical reviews. However, non-cognitive skills, such as flexibility and adaptability, initiative and self-direction, social and cross-cultural skills, productivity and accountability, and leadership and responsibility (Kivunja, 2015) are equally important. The course aims to support the development of these skills by requiring students to work effectively and productively in both expert and Jigsaw groups and to produce collaboratively designed products, such as concept maps. Additionally, meta-reflection is cultivated through self-assessment opportunities as well as the e-portfolio, approaches that support assessment *for* and not only *of* learning. All these experiences are embedded in the context of belonging to a community, which simultaneously encourages and values each individual's contribution by exploiting Role Taking theory that, as we experienced, facilitates social inclusion. The results of this study are promising, as a high percentage of participants have reported to reuse these skills in other contexts in the long term. In particular, the fact that 46% of the students transferred some skills to offline work contexts and 31% re-used them in online work contexts shows that they perceive the design of the course as relevant for their professional career. This finding is at odds with the large-scale studies showing that most students feel that their academic studies have not prepared them well for their professional life.

In particular, the analysis of the area of skills in this study illustrates how the application of the CCP model allowed students to experiment with and enact different types of skills in learning situations that in some ways resonates with the work contexts experienced afterward. This alignment between the learning context and the professional situations could have fostered the reuse of transversal skills after graduation, when the students entered professional life. Analysis of these data indicate that the communicative, organizational, and self-assessment skills have been most often reused in professional contexts.

Our results reflect other research into student satisfaction with e-learning courses (Eom et al., 2006) indicating that students highly value interactions between both students and teacher and among peers through self-assessment and teacher feedback. The CCP model supports these interactions by means of several interconnected activities and tools. For example, discussions aimed at the construction of a collaborative product, and meta-reflection on the roles performed and the e-portfolio. Through these varied interactions, students appear to have developed some transversal skills that are relevant and valuable in diverse contexts – educational, professional, and personal. Also Lai et al. (2016) have found that online interaction (between teacher and students and among students) improves the effectiveness of the course in terms of consolidation of knowledge, causing students to "rethink" what and how they have learned.

Considering the significant reuse of organizational skills, the ability to work in a group to achieve a joint goal, and the ability to be flexible during group work, we suggest that the design of the course based on the CCP model has the potential to powerfully cultivate productive attitudes to collaborative work. By focusing on relational competences and group dynamics the CCP model seeks to enhance students' ability to create and maintain positive social relationships, and in turn to promote students' psycho-social wellbeing. In our approach, this is supported also by the opportunities for critical reflection and self-assessment; for instance through the e-portfolio, the self-assessment record, the grid of e-learning characteristics, and by supporting and scaffolding individual participation in collaborative groups through Role Taking. In this way, students learn to be flexible and adaptable by drawing on communication skills in their peer interactions, to demonstrate initiative and self-direction, to be productive and accountable for their work, and to be meta-cognitive about their own contributions and approaches to learning. This often leads students to radically reconsider what they know and what they can do, as well as re-evaluating their goals and personal potentiality.

The sequencing and progression of modules, together with the repetition of activities (but in relation to different content), might have helped students consolidate their learning, and envision a personal trajectory. Students gradually became more self-regulated and their capacity to self-assess improved. Certainly, these claims are not new. For example, Eom et al. (2006) found that self-assessment allows students to become "responsible" for their own learning, for instance. Our research adds strength to the contention that the transversal skills learned in the academic context can be retained and transferred to other contexts. Furthermore, by recognizing the relevance of social interaction based on group work, a qualitative change is induced. Students become more capable to face life experiences in general and their well-being is reinforced based on the interlacement between learning and social processes.

In general, the structure of this course has supported not only the acquisition of skills and their reuse, but also allowed the recall of vivid memories of the teaching methodology and the learning activities. Based on the results presented here, we suggest it is reasonable to claim that teaching strategies, such as Role Taking and the creation of group products, can stimulate the student to go beyond the simple acquisition of knowledge by requiring and supporting them to be active participants in learning. Taking on specific roles encourages both cognitive and intra and inter-personal development, while the collaborative creation of group products makes learning a concrete experience and, in our case, seems to support transition to professional contexts. The perception of innovativeness might play a role in this process.

## CONCLUSION

The CCP model presented in this paper holds promise as an example of innovative teaching for higher education that effectively combines a range of approaches to teaching and learning, sharing a collaborative and constructive vision of learning aimed at supporting students' development of skills and competences relevant for the students' lives after graduation. Furthermore, by contributing to significant learning for the

workplace, the implementation of course design based on the CCP model can enhance students' sense of efficacy when facing novel working situations and thus contributes to their wellbeing in the long term. Indeed, as claimed by Salanova (2004), efficacy beliefs are associated with increased psycho-social wellbeing and performance.

Of course, this research is not without limitations. The number of participants in some clusters was lower than others, and there was a higher percentage of female than male participants in all the cohorts, restricting the opportunity to run more advanced statistical tests and to compare gendered experiences. Nevertheless, we argue that the descriptive quantitative analysis combined with the qualitative analysis of the answers to the open-ended questions, allowed us to explore this overlooked topic of investigation and glean valuable insights concerning students' perspectives over a long-term timeframe.

In relation to the evaluation process of higher education courses, having a proper evaluation tool is essential (Kirkpatrick, 1998; Owen and Rogers, 1999). In the Italian higher education context, SET questionnaires are the only tool available and they are administered right at the end of the courses or at most, a few weeks later. In those questionnaires, students cannot comment on their learning processes and it is impossible to assess the learning impact of the courses across extended periods of time. Therefore, additional tools are needed to examine the multifaceted experience of students over longer timeframes; these are often overlooked in current research. The CCP model, including teaching strategies such as Role Taking and group work, is expected to trigger specific learning processes having long-term effects in terms of CLS development. Therefore, there is a need to develop strategies to examine such long-term effects. The mixed methodological approach used in this paper allows for assessment of the "historical effect" (Green, 1977) of the course; that is, the long term impact of the learning process.

Although our study does not allow for generalizability of findings due to the limitations discussed above, it does support discussion of the long term effects of the learning model for the participating students. Further, this research provides feedback to different stakeholders of the educational context: teachers, learners, and policy makers.

First, this type of research allows teachers to reflect on their professional decisions, to identify strong and critical aspects of the psycho-pedagogical design, and to adjust the learning setting in response. Although this approach generates data years after the completion of the course, and does not allow to generate immediate feedback for teachers, we argue that the long term impact of higher education courses should not be ignored. Interrogating student perspectives over time can provide crucial information about the impact of an educational psycho-pedagogical design on the students' future personal and professional learning and lives. Furthermore, considering that the career of a teacher might last several decades, we argue that teachers might be able to use such feedback in a productive way while progressively fine-tuning their professional intervention over time.

Second, from the learners' point of view, responding to the survey we developed in relation to a course they attended several years before allows them to reflect upon their learning and to become aware of their development. In particular, including information about the transfer of skills acquired during the course represents an opportunity for significant reflection on their long term learning.

Third, in relation to stakeholders such as policy makers, within the Italian national context, higher education assessment tools aimed at specific evaluation of single courses are conspicuously absent. Recent reforms have introduced several assessment tools, but these always target short term effects and the overall evaluation of the whole set of courses comprising the bachelor or master degrees. Tools aimed at assessing long terms effects are still needed. A research approach such as the one proposed in this article triggers specific reflection on the long-term effects of individual courses as well as gleaning insight into the types of courses that most contribute to students' learning and wellbeing after graduation.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

ML contributed to the design and conception of the study. ML and RD carried out the data collection and contributed to the analysis and interpretation of the data, and wrote the first draft of the manuscript. GR outlined the theoretical framework. GR, KM, and SA contributed to the interpretation and discussion of the data, and reviewed the manuscript contributing significantly to improve it, in some cases re-writing some sections. All authors contributed to multiple iterations of manuscript revision. Finally, GR took responsibility to finalize the manuscript and act as corresponding author.

## FUNDING

The authors thank the University of Helsinki Library for funding the open access publication of this manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00565/full#supplementary-material

## REFERENCES


Annese, S., and Traetta, M. (2011). "A methodological approach for blended communities: social network analysis and positioning network analysis," in *Handbook of Research on Methods and Techniques for Studying Virtual Communities: Paradigms and Phenomena*, ed. B. K. Daniel (Hershey: IGI Global), 103–121. doi: 10.4018/978-1-60960-040-2.ch006

Annese, S., and Traetta, M. (2012). Distributed participation in blended learning communities: actors, contexts and groups. *Int. J. Web Based Commun.* 8, 422–439.

Annese, S., and Traetta, M. (2018). "A dialogical approach for learning communities between positioning and reformulation," in *The Dialogical Self Theory in Education: A Multicultural Perspective*, eds F. Meijers and H. Hermans (Cham: Springer), 189–209. doi: 10.1007/978-3-319-62861-5_13

Annese, S., Traetta, M., and Spadaro, P. F. (2010). "Blended learning communities: relational and identity networks," in *Interpersonal Relations and Social Patterns in Communication Technologies: Discourse Norms, Language Structures and Cultural Variables*, eds J. Park and E. G. Abels (Hershey: IGI Global), 256–276. doi: 10.4018/978-1-61520-827-2.ch014

Aronson, E., and Patnoe, S. (1997). *The Jigsaw Classroom: Building Cooperation in the Classroom*, 2nd Edn. New York, NY: Addison Wesley Longman.

Aronson, E., Stephan, C., Sikes, J., Blaney, N., and Snapp, M. (1978). *The Jigsaw Classroom*. Beverly Hills, CA: Sage Pubblications, Inc.

Astleitner, H. (2018). Multidimensional engagement in learning–an integrated instructional design approach. *J. Instr. Res.* 7, 6–32.

Berger, P., and Luckman, T. (1966). *The Social Construction of Reality: A Treatise it's the Sociology of Knowledge*. New York, NY: Anchor Books.

Brown, A. L., and Campione, J. C. (1990). "Communities of learning or a content by any other name," in *Contribution to Human Development*, ed. D. Kuhn (New York, NY: Oxford University Press), 108–126. doi: 10.1159/000418984

Brown, G. T., and Harris, L. R. (2014). The future of self-assessment in classroom practice: reframing self-assessment as a core competency. *Frontline Learn. Res.* 2, 22–30. doi: 10.14786/flr.v2i1.24

Bruner, J. (1996). *The Culture of Education*. Cambridge, MA: Harvard University Press.

Carless, D. (2016). "Scaling up assessment for learning: progress and prospects," in *Scaling up Assessment for Learning in Higher Education*, eds D. Carless, S. Bridges, C. Chan, and R. Glofcheski (Singapore: Springer), 3–17. doi: 10.1007/978-981-10-3045-1_1

Cook-Sather, A. (2002). Authorizing students' perspectives: toward trust, dialogue, and change in education. *Educ. Res.* 31, 3–14. doi: 10.3102/0013189x031004003

Cucchiara, S., Ligorio, M. B., and Fujita, N. (2014). "Understanding online discourse strategies for knowledge building through social network analysis," in *Innovative Methods and Technologies for Electronic Discourse Analysis*, eds H. Lim and F. Sudweeks (Hershey, PA: Information Science Reference), 42–62. doi: 10.4018/978-1-4666-4426-7.ch003

Dillenbourg, P. (1999). *What do you Mean by Collaborative Learning? Cognitive and Computational Approaches*. Oxford: Elsevier, 1–19.

Engeström, Y. (1987). *Learning by Expanding*. Helsinki: OrientaKonsultit.

Engle, R. A. (2006). Framing interactions to foster generative learning: a situative explanation of transfer in a community of learners classroom. *J. Learn. Sci.* 15, 451–498. doi: 10.1207/s15327809jls1504_2

Eom, S. B., Ashill, N., and Wen, J. (2006). The determinants of students' perceived learning outcomes and satisfaction in University online education: an empirical investigation. *Decis. Sci. J. Innov. Educ.* 4, 215–235. doi: 10.1111/j.1540-4609.2006.00114.x

Etzkowitz, H. (2008). *The Triple Helix: University-Industry-Government Innovation in Action*. Abingdon: Routledge.

European Association for Quality Assurance in Higher Education [ENQA] (2005). *Standards and Guidelines for Quality Assurance in the European Higher Education Area*. Available online at: https://media.ehea.info/file/ENQA/05/3/ENQA-Bergen-Report_579053.pdf (accessed March 25, 2018).

European Association for Quality Assurance in Higher Education [ENQA] (2015). *Standards and Guidelines for Quality Assurance in the European Higher Education Area*. Available online at: http://www.enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf (accessed March 25, 2018).

European Commission (2013). *Report to the European Commission on Improving the Quality of Teaching and Learning in Europe's Higher Education Institutions*. Luxembourg: Publications Office of the European Union.

Gatignon, H., Tushman, M. L., Smith, W., and Anderson, P. (2002). A structural approach to assessing innovation: construct development of innovation locus, type, and characteristics. *Manag. Sci.* 48, 1103–1122. doi: 10.1287/mnsc.48.9.1103.174

Gergen, K. J. (1999). *An Invitation to Social Construction*. London: Sage.

Gergen, K. J. (2001). Psychological science in a postmodern context. *Am. Psychol.* 56, 803–813. doi: 10.1037/0003-066x.56.10.803

Gillies, R. M. (2007). *Cooperative Learning: Integrating Theory and Practice*. Thousand Oaks, CA: SAGE Publications.

Glaser, B. G., and Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies of Qualitative Research*. Chicago: Aldine.

Graham, C. R. (2006). "Blended learning systems," in *Handbook of Blended Learning: Global Perspectives, Local Designs*, eds C. J. Bonk and C. R. Graham (San Francisco, CA: Pfeiffer Publishing), 3–21.

Green, L. W. (1977). Evaluation and measurement. some dilemmas for health education. *Am. J. Public Health* 67, 155–161. doi: 10.2105/ajph.67.2.155

Gursul, F., and Tozmaz, G. B. (2010). Which one is smarter? Teacher or board. *Proc. Soc. Behav. Sci.* 2, 5731–5737. doi: 10.1016/j.sbspro.2010.03.936

Hakkarainen, K. (2009). A knowledge-practice perspective on technology-mediated learning. *Int. J. Comput. Support. Collab. Learn.* 4, 213–231. doi: 10.1007/s11412-009-9064-x

Hakkarainen, K., and Sintonen, M. (2002). Interrogative model of inquiry and computer-supported collaborative learning. *Sci. Educ.* 11, 25–43.

Hare, A. P. (1994). Types of roles in small groups: a bit of history and a current perspective. *Small Group Res.* 25, 443–448.

Henderson, M., Selwyn, N., Finger, G., and Aston, R. (2015). Students' everyday engagement with digital technology in university: exploring patterns of use and 'usefulness'. *J. High. Educ. Policy Manag.* 37, 308–319. doi: 10.1080/1360080x.2015.1034424

Hutchins, E. (2001). "Distributed cognition," in *International Encyclopedia of the Social & Behavioral Sciences*, eds N. J. Smelser and P. B. Baltes (Amsterdam: Elsevier), 2068–2072.

Kelly, A., Lesh, R., and Baek, J. (eds) (2014). *Handbook of Design Research Methods in Education*. New York, ny: Routledge.

Kelly, G. A. (1955). *The Psychology of Personal Constructs*. New York, NY: Norton.

Kirkpatrick, D. L. (1998). *Evaluation Training Programs: The Four Level*. San Francisco, CA: Berrett-Kohler.

Kivunja, C. (2015). Teaching students to learn and to work well with 21st century skills: unpacking the career and life skills domain of the new learning paradigm. *Int. J. High. Educ.* 4, 1–11.

Lai, M., Lam, K. M., and Lim, C. P. (2016). Design principles for the blend in blended learning: a collective case study. *Teach. High. Educ.* 21, 716–729. doi: 10.1080/13562517.2016.1183611

Lave, J. (1988). *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life*. New York, NY: Cambridge University Press.

Ligorio, M. B., and Annese, S. (2010). "Blended activity design approach: a method to innovate e-learning for higher education," in *Psychology Research*, eds A. Blachnio, A. Przepiorka, and T. Rowiński (Warsaw: CSWU Press), 165–188.

Ligorio, M. B., and Cucchiara, S. (2011). Blended collaborative constructive participation (BCCP): a model for teaching in higher education. *eLearning Papers Presented at 27th Transforming Education Through Technology*, Spain.

Ligorio, M. B., and Sansone, N. (2009). "Structure of a Blended University course: applying constructivist principles to a blended course," in *Information Technology and Constructivism in Higher Education: Progressive Learning Frameworks*, ed. C. R. Payne (London: IGI Global), 216–230. doi: 10.4018/978-1-60566-654-9.ch014

Lucena, R., Gitirana, V., and Trouche, L. (2018). "Meta-orchestration instrumental for teacher education," in *Proceedings of the Re (s) sources 2018 International Conference*, Lyon, 300–303.

Maul, A., Penuel, W. R., Dadey, N., Gallagher, L. P., Podkul, T., and Price, E. (2017). Measuring experiences of interest-related pursuits in connected learning. *Educ. Technol. Res. Dev.* 65, 1–28. doi: 10.1007/s11423-016-9453-6

McGraw-Hill Education (2016). *Workforce Readiness Survey*. New York, NY: McGraw-Hill.

Mowery, D. C., Nelson, R. R., Sampat, B. N., and Ziedonis, A. A. (2015). *Ivory Tower and Industrial Innovation: University-Industry Technology Transfer Before and After the Bayh-Dole Act*. Palo Alto, CA: Stanford University Press.

Mudrack, P. E., and Farrell, G. M. (1995). An examination of functional role behavior and its consequences for individuals in group settings. *Small Group Res.* 26, 542–571. doi: 10.1177/1046496495264005

Muukkonen, H., Lakkala, M., and Hakkarainen, K. (2005). Technology-mediation and tutoring: how do they shape progressive inquiry discourse? *J. Learn. Sci.* 14, 527–565. doi: 10.1207/s15327809jls1404_3

Owen, J. M., and Rogers, P. J. (1999). *Program Evaluation : Forms and Approaches*. St. Leonards: Allen & Unwin.

Potter, J. (1996). "Representing reality. discourse, rhetoric and social construction," in *Handbook of Qualitative Research Methods for Psychology and Social Sciences*, ed. J. E. Richardson (Leicester: British Psychological Society), 136–178.

Ritella, G., and Hakkarainen, K. (2012). Instrumental genesis in technology-mediated learning: from double stimulation to expansive knowledge practices. *Int. J. Comput. Support. Collab. Learn.* 7, 239–258. doi: 10.1007/s11412-012-9144-1

Ritella, G., and Ligorio, M. B. (2016). Investigating chronotopes to advance a dialogical theory of collaborative sensemaking. *Cult. Psychol.* 22, 216–231. doi: 10.1177/1354067x15621475

Ritella, G., Ligorio, M. B., and Hakkarainen, K. (2017). Interconnections between the discursive framing of space-time and the interpretation of a collaborative task. *Learn. Cult. Soc. Interact.* 20, 45–57. doi: 10.1016/j.lcsi.2017.08.001

Ritella, G., and Sansone, N. (2020). Transforming the space-time of learning through interactive whiteboards: the case of a knowledge creation collaborative task. *Qwerty* 15. doi: 10.30557/QW000022

Salanova, M. P. (2004). Engagement and burnout: analysing their associated patterns. *Psychol. Rep.* 94, 1048–1050. doi: 10.2466/pr0.94.3.1048-1050

Säljö, R. (2016). "Apps and learning," in *Apps, Technology and Younger Learners: International Evidence for Teaching*, eds N. Kucirkova and G. Falloon (Milton Park: Taylor & Francis).

Samuel, R., Bergman, M. M., and Hupka-Brunner, S. (2013). The interplay between educational achievement, occupational success, and well-being. *Soc. Indic. Res.* 111, 75–96. doi: 10.1586/ern.11.50

Sansone, N., Ligorio, M. B., and Buglass, S. (2016). Peer e-tutoring: effects on students' participation and interaction style in online course. *Innov. Educ. Teach. Int.* 55, 13–22. doi: 10.1080/14703297.2016.1190296

Scardamalia, M., and Bereiter, C. (2006). "Knowledge building: theory pedagogy, and technology," in *Cambridge Handbook of the Learning Sciences*, ed. K. Sawyer (New York, NY: Macmillan), 97–118.

Schleicher, A. (2011). *Building a High-Quality Teaching Profession: Lessons from Around the World*. Paris: Organization for Economic Cooperation and Development (OECD).

Schneider, M., and Preckel, F. (2017). Variables associated with achievement in higher education: a systematic review of meta-analyses. *Psychol. Bull.* 143, 565–600. doi: 10.1037/bul0000098

Schomaker, J., and Meeter, M. (2015). Short-and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neurosci. Biobehav. Rev.* 55, 268–279. doi: 10.1016/j.neubiorev.2015.05.002

Shotter, J. (1993). *Cultural Politics of Everyday Life: Social Constructionism, Rhetoric, and Knowing of the third Kind*. Milton Keyne: Open University Press.

Smith, K. A., Sheppard, S. D., Johnson, D. W., and Johnson, R. T. (2005). Pedagogies of engagement: classroom-based practices. *J. Eng. Educ.* 94, 87–101. doi: 10.1002/j.2168-9830.2005.tb00831.x

Steeples, C., Goodyear, P., and Mellar, H. (1994). *'Flexible Learning in Higher Education: The use of Computer-Mediated Communications', Computers and Education*. Oxford: Pergamon Press.

Strijbos, J. W., and Weinberger, A. (2010). Emerging and scripted roles in computer-supported collaborative learning. *Comput. Hum. Behav.* 26, 491–494. doi: 10.1016/j.chb.2009.08.006

Suchman, L. (1987). *Plans and Situaded Actions*. Cambridge: Cambridge University Press.

Tadesse, T., and Gillies, R. M. (2015). Nurturing cooperative learning pedagogies in higher education classrooms: evidence of instructional reform and potential challenges. *Curr. Issues Educ.* 18, 1–18.

Topping, K. (2005). Trends in peer learning. *Educ. Psychol.* 25, 631–645. doi: 10.1080/01443410500345172

Topping, K. J. (2003). "Self and peer assessment in school and university: reliability, validity and utility," in *Optimizing New Modes of Assessment: In Search of Qualities and Standards*, eds M. Segers, F. Dochy, and E. Cascallar (Dordrecht: Kluwer Academic), 55–87. doi: 10.1007/0-306-48125-1_4

van Dinther, M., Dochy, F., and Segers, M. (2015). The contribution of assessment experiences to student teachers' self-efficacy in competence-based education. *Teach. Teach. Educ.* 49, 45–55. doi: 10.1016/j.tate.2015.02.013

Vygotsky, L. S. (1986). *Thought and Language*. Cambridge, MA: MIT Press.

Wanner, T., and Palmer, E. (2015). Personalising learning: exploring student and teacher perceptions about flexible learning and assessment in a flipped university course. *Comput. Educ.* 88, 354–369. doi: 10.1016/j.compedu.2015.07.008

Wenger, E. C. (1998). *Communities of Practice. Learning, Meaning, and Identity*. New York, NY: Cambridge University Press.

Zhao, Y. (2015). A world at risk: an imperative for a paradigm shift to cultivate 21st century learners. *Society* 52, 129–135. doi: 10.1007/s12115-015-9872-8

Zigurs, I., and Kozar, K. A. (1994). An exploratory study of roles in computer-supported groups. *MISQuarterly* 18, 277–297.

Zimmerman, B. J. (2001). "Theories of self-regulated learning and academic achievement: an overview and analysis," in *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*, eds B. J. Zimmerman and D. H. Schunk (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 1–37.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 23 April 2020
doi: 10.3389/fpsyg.2020.00751

# The Use of New Technologies for Improving Reading Comprehension

*Agnese Capodieci[1]*, Cesare Cornoldi[1], Elizabeth Doerr[1], Laura Bertolo[2]* and Barbara Carretti[1]*

*[1] Department of General Psychology, University of Padova, Padua, Italy, [2] Azienda Sociosanitaria Ligure 5 Spezzino, La Spezia, Italy*

Since the introduction of writing systems, reading comprehension has always been a foundation for achievement in several areas within the educational system, as well as a prerequisite for successful participation in most areas of adult life. The increased availability of technologies and web-based resources can be a really valid support, both in the educational and clinical field, to devise training activities that can also be carried out remotely. There are studies in current literature that has examined the efficacy of internet-based programs for reading comprehension for children with reading comprehension difficulties but almost none considered distance rehabilitation programs. The present paper reports data concerning a distance program *Cloze*, developed in Italy, for improving language and reading comprehension. Twenty-eight children from 3rd to 6th grade with comprehension difficulties were involved. These children completed the distance program for 15–20 min for at least three times a week for about 4 months. The program was presented separately to each child, with a degree of difficulty adapted to his/her characteristics. Text reading comprehension (assessed distinguishing between narrative and informative texts) increased after intervention. These findings have clinical and educational implications as they suggest that it is possible to promote reading comprehension with a distance individualized program, avoiding the need for the child displacements, necessary for reaching a rehabilitation center.

**Keywords: reading comprehension, training, distance rehabilitation program, digital device, Cloze app**

OPEN ACCESS

***Edited by:***
*Ana Lucia Pereira, Universidade Estadual de Ponta Grossa, Brazil*

***Reviewed by:***
*Gaston Saux, National Council for Scientific and Technical Research (CONICET), Argentina*
*Giovanna Bubbico, G. d'Annunzio University of Chieti and Pescara, Italy*

****Correspondence:***
*Agnese Capodieci*
*agnesecapox@gmail.com*
*Laura Bertolo*
*laura.bertolo7@gmail.com*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *20 November 2019*
***Accepted:*** *27 March 2020*
***Published:*** *23 April 2020*

***Citation:***
*Capodieci A, Cornoldi C, Doerr E, Bertolo L and Carretti B (2020) The Use of New Technologies for Improving Reading Comprehension. Front. Psychol. 11:751. doi: 10.3389/fpsyg.2020.00751*

## INTRODUCTION

Reading comprehension is a fundamental cognitive ability for children, that supports school achievement and successively participation in most areas of adult life (Hulme and Snowling, 2011). Therefore, children with learning disabilities (LD) and special educational needs who show difficulties in text comprehension, sometimes also in association with other problems, may have an increased risk of life and school failure (Woolley, 2011). Reading comprehension is, indeed, a complex cognitive ability which involves not only linguistic (e.g., vocabulary, grammatical knowledge), but also cognitive (such as working memory, De Beni and Palladino, 2000), and metacognitive skills (both for the aspects of knowledge and control, Channa et al., 2015), and, more specifically, higher order comprehension skills such as the generation of inferences (Oakhill et al., 2003).

Recently, due to the diffusion of technology in many fields of daily life, text comprehension at school, at home during homework, and at work is based on an increasing number of

digital reading devices (computers and laptops, e-books, and tablet devices) that can become a fundamental support to improve traditional reading comprehension and learning skills (e.g., inference generation).

Some authors contrasted in children with typical development the effects of the technological interface on reading comprehension vs printed texts (Kerr and Symons, 2006; Rideout et al., 2010; Mangen et al., 2013; Singer and Alexander, 2017; Delgado et al., 2018). Results were consistent and showed a worse comprehension performance in screen texts compared to printed texts for children (Mangen et al., 2013; Delgado et al., 2018) and adolescents who nonetheless showed a preference for digital texts compared to printed texts (Singer and Alexander, 2017). Regarding children with learning problems, only few studies considered the differences between printed texts and digital devices (Chen, 2009; Gonzalez, 2014; Krieger, 2017) finding no significant differences, suggesting that the use of compensative digital tools for children with a learning difficulty could be a valid alternative with respect to the traditional written texts in facilitating their academic and work performance. This conclusion is also supported by the results of a meta-analysis (Moran et al., 2008), regarding the use of digital tools and learning environments for enhancing literacy acquisition in middle school students, which demonstrates that technology can improve reading comprehension.

Different procedures and abilities are targeted in the international literature concerning computerized training programs for reading comprehension. In particular, various studies include activities promoting cognitive (e.g., vocabulary, inference making) and metacognitive (e.g., the use of strategies, comprehension monitoring, and identification of relevant parts in a text) components of reading comprehension. **Table 1** reports the list of papers proposing computerized training programs with a summary of the findings encountered. Participants involved cover different ages and school grades, the majority belonging to middle school and high school. The general outcome of the studies is positive due to a significant improvement in comprehension skills after the training program with long-lasting effects also during follow-up; indeed, the majority of participants involved in training programs outperformed their peers assigned to comparison groups and maintained their improvements. Specifically, several studies (O'Reilly et al., 2004; Magliano et al., 2005; McNamara et al., 2006) used the iSTART program with adolescents and young adults. This program promotes self-explanation, prior knowledge and reading strategies to enhance understanding of descriptive scientific texts. Results demonstrated that students who followed the iSTART program received more benefits than their peers, improving self-explanation and summarization. Additionally, strategic knowledge was a relevant factor for the outcome in comprehension tasks including multiple choice questions: students who already possessed good strategic knowledge improved their accuracy when answering to bridging inference questions, whereas students with low strategic knowledge became more accurate with text-based questions. Another program, ITSS, was used with younger students (Meyer et al., 2011; Wijekumar et al., 2012, 2013, 2017), with the objective to support activities based on identifying main parts and key words in a text and classifying information in a hierarchical order. Positive outcomes were found also with such program since students who followed the ITSS program significantly improved text comprehension compared to their peers in the control group.

Although most of the literature deals with typical development, also cases of students with learning difficulties were considered. For example, Potocki et al. (2013) (see also Potocki et al., 2015) examined the effects of two different computerized programs with specific aims: one focusing on comprehension features, such as inference making and the analysis of text structure, the other considering decoding skills. Both training programs brought some benefits to reading comprehension, however larger effects were found with the program focused on comprehension with long-lasting effects in listening and reading comprehension (see also Kleinsz et al., 2017). Studies by Johnson-Glenberg (2005) and Kim et al. (2006), using respectively the programs 3D Readers and CACSR, were able to promote reading comprehension abilities in middle school students through metacognitive activities. Thanks to these programs students also became more aware of reading strategies and implemented them more successfully during text comprehension. In particular, a study by Niedo et al. (2014), obtained positive results on silent reading in a small group of children struggling with reading using the "cloze" procedure. This procedure proposes exercises in which parts of a text, typically words, are missing and participants are required to complete the text guessing what is missing.

Thus, computerized programs generally seem to improve reading comprehension skills. However, it should be noticed that, in most cases, students were trained at school, without the personalized support of a clinician taking into consideration the cognitive and psychological needs of the child. In particular, to our knowledge, no program examined the effects of an internet-based distance reading comprehension program which allows the child to be trained at home in a personalized way. A useful aspect of an internet-based distance training is that the psychologist can monitor with the application (*app*) the child's results and activities and write him/her some motivational messages, reducing the attritions present in programs carried out at home with the only supervision of parents. Literature concerning distance trainings is still rare, however, some evidence suggests that these programs may represent a good integration to other types of intervention, usually carried out at school, in a rehabilitation center or at home (e.g., Mich et al., 2013).

Therefore, despite still preliminary, we think that it is relevant to present data about a distance program developed in Italy named Cloze (Cornoldi and Bertolo, 2013), devised for rehabilitation purposes but with potential implication also for educational contexts. *Cloze* has been developed to promote inferential abilities both at a sentence- and discourse-level using the "cloze" procedure. Several findings in the literature demonstrate that abilities, such as anticipating

**TABLE 1 |** Synthesis of the main results of the computerized training programs on comprehension present in the literature.

| Author/s | Computerized training program | Type of control group | Trained reading comprehension components | School grade | Measures | Efficacy (during/after post-test) |
|---|---|---|---|---|---|---|
| **Typical Development** | | | | | | |
| O'Reilly et al. (2004) | iSTART | Active | Cognitive (textual information, connecting text sentences)<br>Metacognitive (self-explanation, prior knowledge, reading strategies) | End of 7th–8th | Reading<br>Prior Knowledge<br>Reading strategy knowledge | I > C<br>Effects on comprehension<br>Improvement in comprehension questions regardless of the initial strategy knowledge level |
| Magliano et al. (2005) | iSTART | Active (live training) | Cognitive (textual information, connecting text sentences)<br>Metacognitive (self-explanation, prior knowledge, reading strategies) | Undergraduate | Reading<br>Self-explanation | Skilled and less skilled readers improved with the program<br>Less skilled readers from the live training improved more in text-based question, whereas skilled readers improved with inference questions<br>Both treatments brought similar effects |
| McNamara et al. (2006) | iSTART | Active | Cognitive (textual information, connecting text sentences)<br>Metacognitive (self-explanation, prior knowledge, reading strategies) | 8th–9th | Reading strategy knowledge<br>Self-explanation tests | I > C<br>Effects on comprehension and self-explanation<br>Correlation between self-explanation and comprehension questions was stronger in the intervention group |
| Meyer et al. (2011) | ITSS | No control (comparison between standardized or tailored version of ITSS). | Cognitive (recall specific text structure, identify key words, classification of information) | 5th | Reading comprehension | Both groups improved<br>More improvement with the individualized version |
| Wijekumar et al. (2012) | ITSS | Active | Cognitive (recall specific text structure, identify key words, classification of information) | 4th | Reading comprehension | I > C<br>Small but significant effects on comprehension |
| Wijekumar et al. (2013) | ITSS | Active | Cognitive (recall specific text structure, identify key words, classification of information) | 4th–5th | Silent Reading | I > C;<br>Small but significant effects on comprehension |
| Ortlieb et al. (2014) | myON® | 3 Groups<br>Print-based<br>Hybrid<br>myON | Cognitive (online reading and text analysis) | 4th | Reading<br>Comprehension | All groups improved in comprehension<br>myON group was outperformed by the other two when using printed material<br>myON group improved reading comprehension in the digital format |
| Wijekumar et al. (2017) | ITSS | Active | Cognitive (text structure recognition, hierarchical classification of information) | 7th | Reading comprehension | I > C (in all measures)<br>Small but significant effects on comprehension |
| **Learning Difficulties** | | | | | | |
| Leong (1992) | DECtalk | Active | Cognitive (vocabulary, text-to-speech skills) | 6th–8th | Vocabulary<br>Reading comprehension | C > I (when combining the program with online reading explanations) |
| Johnson-Glenberg (2005) | 3D Readers | Within groups design with control conditions | Metacognitive (comprehension monitoring, graphic organizers, question answering, question generation, summarization) | 6th–7th | Comprehension assessed with open-ended answers<br>Comprehension assessed with vocabulary gains<br>Rereading assessed with ScrollBacks<br>Question generation | Higher performance in comprehension and longer rereading time in the metacognitive condition |

*(Continued)*

**TABLE 1** | Continued

| Author/s | Computerized training program | Type of control group | Trained reading comprehension components | School grade | Measures | Efficacy (during/after post-test) |
|---|---|---|---|---|---|---|
| Kim et al. (2006) | CACSR | Active | Cognitive (recognition or relevant elements of the text, connecting text sentences)<br>Metacognitive (monitoring before, during and after reading, previous knowledge) | Middle-school | Comprehension<br>Writing main ideas from a text<br>Decoding | I > C<br>Positive effects on comprehension<br>Perceived improvement from students and teachers |
| Potocki et al. (2013) | LoCoTex (reading comprehension, intervention group) v.s. Chassymo (decoding, control group) | Active | Cognitive (inference making, text analysis) | 2nd | Word reading<br>Non-verbal intelligence<br>Word span;<br>Updating<br>Listening and reading comprehension<br>Vocabulary;<br>Comprehension monitoring | I > C<br>Positive effects on reading and listening comprehension<br>Effects on monitoring and vocabulary (significative only until follow-up) |
| Niedo et al. (2014) | RAP | Active | Cognitive (sentence logic, Cloze technique, paragraph understanding) | 4th | Word reading<br>Pseudoword reading<br>Reading comprehension accuracy<br>Silent word reading<br>Silent sentence reading<br>Working memory<br>Attention and hyperactivity rating | Both groups improved<br>Attention and working memory predicted post-test accuracy (silent reading comprehension) |
| Cullen et al. (2014) | Headsprout Comprehension Program | No control | Cognitive (vocabulary, text analysis, answering questions) | 3rd and 5th | Reading comprehension | Substantial improvement in reading comprehension |
| Potocki et al. (2015) | LoCoTex vs Chassymo | 4 Groups<br>normal readers<br>poor decoders<br>poor comprehenders<br>general poor readers | Cognitive (inference making, text analysis) | 6th–7th | Silent word reading<br>Fluency<br>Listening comprehension | Improved fluency with decoding training<br>Improvement in word recognition, listening and reading comprehension with comprehension training |
| Kleinsz et al. (2017) | LoCoTex vs Chassymo | 3 Groups:<br>Specific decoding difficulty<br>Specific comprehension difficulty<br>General reading difficulty | Cognitive (inference making, text analysis) | 2nd | Written word recognition<br>Listening and reading comprehension<br>Decoding<br>Phonological skills<br>Decoding fluency<br>Vocabulary<br>Comprehension monitoring<br>Working memory<br>Non-verbal reasoning | Difficulties in working memory during testing phases in all three groups<br>Improvements in word decoding and phonological awareness with decoding training<br>Improvements in vocabulary, comprehension monitoring, decoding fluency and accuracy, word-recognition with comprehension training |

*I > C, intervention group outperformed control group; iSTART, Interactive Strategy Training for Active Reading and Thinking; ITSS, Intelligent Tutoring for the Structure Strategy; DECtalk, Text-to-speech program; CASR, Computer-Assisted Collaborative Strategic Reading; PILE, ProgramaInformatizado de LeituraEstratégica; RAP, Rapid Accelerated-reading Program. Active control: group not involved in the experimental training but receiving some kind of treatments (for example other exercises or games).*

**TABLE 2 |** Main characteristics of the sample in terms of reading and cognitive abilities.

| Task | Mean | Standard deviation |
|---|---|---|
| **Reading assessment** | | |
| Text reading speed (syllables per seconds) | 2.70 | 0.81 |
| Text reading speed (syllables per seconds) *z* score | −0.81 | 0.77 |
| Text reading errors | 4.32 | 3.7 |
| Words reading (syllables per seconds) | 2.37 | 0.75 |
| Words reading (syllables per seconds) *z* score | −0.75 | 1.03 |
| Words reading errors | 4.79 | 2.3 |
| Non-word reading (syllables per seconds) | 1.73 | 0.66 |
| Non-word reading (syllables per seconds) *z* score | −0.13 | 1.14 |
| Non-word reading (syllables per seconds) errors | 7.04 | 3.13 |
| **Cognitive assessment** | | |
| Full Scale IQ (FSIQ) | 97.61 | 10.24 |
| Verbal Comprehension Index (VCI) | 103.04 | 11.29 |
| Perceptual Reasoning Index (PRI) | 103.29 | 11.52 |
| Working Memory Index (WMI) | 90.56 | 9.80 |
| Processing Speed Index (PSI) | 92.44 | 11.46 |

text parts and inference making, bring improvements in text comprehension (e.g., Yuill and Oakhill, 1988) and it has been shown that one way to promote inferential competences is to improve the ability to predict parts of the text that are missing or that follow, considering the available information: the "cloze" technique appears to be one of the most successful ways for this purpose (e.g., Greene, 2001).

In the current study the effectiveness of this training program has been tested on a clinical population who exhibited, for various reasons, difficulties in reading comprehension. Participants were 28 children (16 male and 12 female) attending a private practice for learning difficulties in the city of La Spezia, in the north-west of Italy, from 3rd to 6th school grade (5 of 3rd, 9 of 4th, 11 of 5th and 3 of 6th grade), with a mean age of children of $M$ = 9.79 years (SD = 1.03). Seventeen children had a current or past speech disorder: of these children 10 also had a LD (Learning Disabilities) and one was bilingual (speech problems were not due to bilingualism). The other 11 children had a LD or important learning difficulties, and one of them had also ADHD (Attention Deficit/Hyperactivity Disorder). For the goals of the study, all these children were considered together as they all presented a severe reading comprehension difficulty as reported by parents and teachers and confirmed by the initial assessment.

All children had received a comprehensive psychological assessment (see **Table 2**), adapted to their particular needs and ages. In particular all children had an IQ >80 assessed with the Wechsler Intelligence Scale for Children-IV (WISC-IV; Wechsler, 2003) and did not have anxiety disorders, mood affective disorders or other developmental disorders, with the exception of the cases with language disorder and the case with ADHD. Children were not receiving any additional treatment, including medication. Written consent was obtained from the children's parents in the context of the private practice.

# MATERIALS AND METHODS

## Pre-/Post-test Assessment and Procedure of the Training

Each child started a training program through the distance rehabilitation platform Ridinet, using the Cloze app, after the assessment of learning and cognitive abilities, including comprehension assessment with two texts, one narrative and one informative (Cornoldi and Carretti, 2016; Cornoldi et al., 2017). Connection to the Ridinet web site was required in order to access to the app, three or four times a week for more or less 15/20 min. The period of use was of 3 months for 6 children and 4 months for 22 children. After this period children's comprehension was assessed again. Additionally, some questions were asked to parents and children about the app's utility and pleasantness. In particular, children were asked: "Do you think the program helped you improve your text comprehension skills?," "Did you like doing this program instead of the same exercises on paper?"; and parents were asked: "Was it difficult to start the Cloze activities on days when it had to be done?," "Compared to the beginning of the treatment, how do you currently judge the ability of your child to understand the texts?". For all questions, except the last one, the answer had to be given on a 5-point scale with 1 = not at all, 2 = a little, 3 = enough, 4 = very, 5 = very much. For the last question the answer changed on a 4-point scale with 1 = got worse, 2 = unchanged, 3 = slightly improved, and 4 = greatly improved.

## Comprehension Tasks

Reading comprehension was assessed with two texts, the first narrative and the other informative, taken from Italian batteries for the assessment of reading (Cornoldi and Carretti, 2016; Cornoldi et al., 2017). The texts range between 226 and 455 words in length, and their length increases with school grade (in order to have texts and questions matching the degrees of expertise at different grades the batteries include a different pair of texts for each grade). Students read the text in silence at their own pace, then answer a variable number of multiple-choice questions (depending on school grade), choosing one of four possible answers. There is no time limit, and students can reread the text whenever they wish. The final score is calculated as the total number of correct answers for each text. Alpha coefficients, as reported by the manuals, range between 0.61 and 0.83. For the purposes of the study we decided to use the same two comprehension texts, at pre-test and post-test, as the procedure offered the opportunity of directly examining and showing to parents changes in comprehension and previous evidence had shown the absence of relevant retest effects with this material in a retest carried out after 3 months (Viola and Carretti, 2019).

## Distance Rehabilitation Program: Cloze

Cloze (Cornoldi and Bertolo, 2013) is an app for the promotion of text comprehension with the specific aim to recover processes of lexical and semantic inference. At each work session the

child works with texts that lack words and must complete the empty spaces by choosing the correct alternative from those automatically proposed by the app, so that the text becomes congruent. The program is adaptive, as text complexity and proportion of missing words vary according to the previous level of response, and is designed for children who have weaknesses in written text comprehension, mainly due to poor skills in lexical and semantic inferential processes. The app also allows to enhance a set of language skills (phonology, syntax, semantics) which contribute to ensuring the fluidity of text and production processing. The recommended age range for the use of this program is between 7 and 14 years. In this study the semantic mode (only content words may be missing and no syntactic cues can be used for deciding between the alternatives) was proposed to 21 children and the syntactic mode (where all words may be missing) to 7 children. The mode type selected for each child depends from the performance at pre-test and diagnosis. A clinician, co-author of the present study (LB), monitored the child's results and activities with the app and sent him/her from time to time some motivational messages. The motivational messages were typically sent once a week for congratulating with children for the work done and check with him/her possible problems emerged. Training lasted from 3 to 4 months and involved between 3 and 4 sessions of 15–20 min per week. The variation in duration depended on the decision of each individual family. In fact, children were required to use the software for about 4 months or in any case for a minimum period of 3 months (choice made by six families).

# RESULTS

## Effects on Reading Comprehension of *Cloze* Training

All analyses were carried out with SPSS 25 (IBM Corp, 2017). A preliminary analysis found that all the examined variables met the assumptions of normality (K-S between 0.106 and 0.143, $p > 0.05$). Then, we compared the reading comprehension performance of children before and after the computerized training with *Cloze*. For this analysis, a repeated measure Analysis of Variance (ANOVA) was conducted on comprehension scores to examine the differences in the whole group of children between the scores obtained before and after the training. A significant difference was found for both comprehension texts [$F(1,27) = 22.37$, $p < 0.001$, $\eta^2_p = 0.453$ and $F(1,27) = 38.90$, $p < 0.001$, $\eta^2_p = 0.599$, respectively]. Possible differences between the two training modalities (semantic vs syntactic) and between different training periods (3 months vs 4 months) were then analyzed; no significant differences emerged between groups in both cases [$F(1,27) < 1$].

Secondly, to analyze the role of individual differences at pre-test, the standardized training gain score (STG; Jaeggi et al., 2011) – computed by subtracting post-test score minus pre-test score, divided by the SD of the pre-test – was calculated for the two texts comprehension.

**TABLE 3 |** Changes in performance in relations to norms (provided by the manual) after the training program Cloze.

| | Pre-test | | | Post-test | | |
|---|---|---|---|---|---|---|
| | $z < -2$ | $-2 < z < -1$ | $z > -1$ | $z < -2$ | $-2 < z < -1$ | $z > -1$ |
| Text 1 (narrative) | 4 | 13 | 12 | 1 | 3 | 23 |
| Text 2 (informative) | 6 | 16 | 5 | 1 | 7 | 19 |

Pearson correlations were computed between the STG and the variable collected at pre-test (reading speed and errors, WISC IV – Full scale IQ, Verbal Comprehension, Perceptual Reasoning, Working Memory and Processing Speed indexes). The only significant correlation was between STG of the narrative text and Verbal Comprehension Index of the WISC-IV Scale ($r = 0.38$, $p = 0.048$). Finally, individual improvements from pre- to post-test were also confirmed considering changes in performance in terms of standard deviation in relations to norms (provided by the manual). **Table 3** shows the number of children for each comprehension text who improved their performance moving from a performance at least 2 standard deviations or between 1 and 2 negative standard deviations under the mean to a performance above one negative standard deviation.

## Perceived Utility, Pleasantness, Parents and Child's Improvements of *Cloze*

Results concerning the answers of parents and children about utility, pleasantness and self-perceived efficacy of the app, were also analyzed. At the first question, addressing children's perceived improvement in comprehension skills, more than half of the sample chose the alternatives "very" or "very much" (15 "very" and 5 "very much"), only 1 child answered "a little" and the others chose "enough." At the second question, about the pleasure of doing this kind of activity instead of pen and paper activities, all children answered "very" or "very much." Concerning parents' questions, at the first question about the difficulty to start the *Cloze* activity, only one parent answered "enough," a quarter of the sample chose "a little" (seven families) and all the other 20 families chose the alternative "not at all." At the last question about the perceived training efficacy on their child's performance, the large majority of the families chose "slightly improved" or "greatly improved" and only three parents thought their children's ability had remained unchanged. However, no correlations between parents and child's perceived improvements and STG in reading comprehension were found.

# CONCLUSION

The present study examined the effects of the use of *Cloze*, a distance rehabilitation program focused on inference skills, for improving reading comprehension, on the basis of the hypothesis that, being inference making related to reading comprehension

at different ages (e.g., Oakhill and Cain, 2012), positive effects of the training activities on reading comprehension should be found.

Concerning the efficacy of computer-assisted training programs, literature highlights that many training programs are devised for an educational context. Results are generally encouraging with positive effects on reading comprehension, measured with materials different from those practiced during the training. However, few studies analyzed the efficacy in children with specific reading comprehension problems, and no studies considered the possibility of carrying out a training at home under the distance supervision of an expert. The latter characteristics are those that make the *Cloze* peculiar compared to the existent literature. *Cloze* is indeed based on a rehabilitation online platform which allows the child to complete personalized training activities several times a week, without moving from his/her home, and concurrently enabling the clinician to monitor the child's progress or manage activities' characteristics. The advantage of this procedure is twofold: on one hand it increases the potential number of training sessions per week, on the other hand it permits to save the necessary time to reach the center for rehabilitation and to reduce the costs of the intervention.

The preliminary data on *Cloze* were generally positive: children, working on either two slightly different versions of the same program, showed a generalized improvement in reading comprehension tasks and, together with their families, expressed appreciation for the pleasantness and the efficacy of the program. Encouraging results emerged also from the analysis of individual improvements referring to normative scores, as reported in **Table 3**: most of the children's performance migrated from a highly negative level to an average level.

It is noticeable that the efficacy of the training was assessed with materials different from those practiced during the training sessions, since reading comprehension tasks required to read a paper text and complete a series of multiple-choice questions. In future studies it would be interesting to analyze the effects of the program on skills known to be related to text comprehension, such as vocabulary or comprehension monitoring, for example. There is good reason to believe that since these variables are highly predictive of comprehension skills (and given that training in these skills sometimes improve comprehension; e.g., Beck et al., 1982; see also Hulme and Snowling, 2011), training that specifically targets comprehension might, in turn, lead to improvements in vocabulary or comprehension monitoring skills. Further studies are needed to explore this hypothesis.

A second relevant finding of the present study is the presence of a positive correlation between the gain obtained in one of the reading comprehension text (the narrative one) and the Verbal Comprehension Intelligence Quotient (VCIQ) index of the WISC-IV battery, showing that children who started with more resources in verbal intelligence achieved greater improvements in text comprehension at least with one type of text through the *Cloze*. The activities probably required to develop some kind of strategies, and for this reason students with larger verbal intellectual resources, who were presumably more able to develop new strategies, were more advantaged. Indeed, this amplification effect is usually found when training activities require the development of strategies (von Bastian and Oberauer, 2014). Such result has clinical and educational implications, inviting professionals and teachers to consider children's starting resources and, if necessary, to combine activities conducted through distance rehabilitation programs with personal intervention sessions that could teach strategies and promote a metacognitive approach to reading comprehension. However, some limitations of the present study must be acknowledged. Firstly, study did not include a control group, therefore findings should be taken with caution, although normative data and previous results obtained with the same test offer support to the robustness of our results and the use of normative data offers a control measure of how reading comprehension skills are acquired in typically developing children without specific training, therefore functioning as a sort of passive control group. Secondly, the treated group, although characterized by a common reading comprehension difficulty, was partly heterogeneous, as children attended different grades and could have different diagnoses. Unfortunately, the limited number of subjects, with the consequence that it was not possible to form groups defined both by the grade and the diagnosis, did not permit to make analyses taking into account the grade and the diagnosis as between-subjects factors. Future studies should examine a more homogeneous population or consider a larger sample of children, giving more information about the efficacy of training in different children population. Additionally, the fact that the treatment was concluded with the post-training assessment did not offer the opportunity to further examine the procedure and maintenance effects with a follow-up. Despite the limitations, this study offers evidence concerning the efficacy of new methods, based on computer-assisted training programs that could be beneficial in training high-level skills such as comprehension and inference generation. Such tools can be extremely worthwhile for struggling readers who may need to receive further attention in mastering higher level reading comprehension.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent

to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

AC, CC and BC contributed to the design and implementation of the research. LB provided the data. BC organized the database. AC performed the statistical analysis. ED did the literature research and wrote the section about the review of the literature. AC and BC wrote the other sections. CC contributed to the manuscript revision, read and approved the submitted version.

## FUNDING

The present work was carried out within the scope of the research program Dipartimenti di Eccellenza (art.1, commi 314-337 legge 232/2016), which was supported by a grant from MIUR to the Department of General Psychology, University of Padua and partially supported by a grant (PRIN 2015, 2015AR52F9_003) to Cesare Cornoldi funded by the Italian Ministry of Research and Education (MIUR).

## REFERENCES

Beck, I. L., Perfetti, C. A., and McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *J. Educ. Psychol.* 74, 506–521. doi: 10.1037/0022-0663.74.4.506

Channa, M. A., Nordin, Z. S., Siming, I. A., Chandio, A. A., and Koondher, M. A. (2015). Developing reading comprehension through metacognitive strategies: a review of previous studies. *Eng. Lang. Teach.* 8, 181–186. doi: 10.5539/elt.v8n8p181

Chen, H. (2009). Online reading comprehension strategies among fifth- and sixth-grade general and special education students. *Educ. Res. Perspect.* 37, 79–109.

Cornoldi, C., and Bertolo, L. (2013). *Cloze Ridinet*. Bologna: Anastasis.

Cornoldi, C., and Carretti, B. (2016). *Prove MT-3-Clinica*. Firenze: Giunti Edu.

Cornoldi, C., Carretti, B., and Colpo, C. (2017). *Prove MT-Kit Scuola. Dalla valutazione degli Apprendimenti di Lettura E Comprensione Al Potenziamento. [MT-Kit for the Assessment In The School. From Reading Assessment To Its Enhancement]*. Firenze: Giunti Edu.

Cullen, J. M., Alber-Morgan, S. R., Schnell, S. T., and Wheaton, J. E. (2014). Improving reading skills of students with disabilities using headsprout comprehension. *Remed. Spec. Educ.* 35, 356–365. doi: 10.1177/0741932514534075

De Beni, R., and Palladino, P. (2000). Intrusion errors in working memory tasks: are they related to reading comprehension ability? *Learn. Individ. Differ.* 12, 131–143. doi: 10.1016/s1041-6080(01)00033-4

Delgado, P., Vargas, C., Ackerman, R., and Salmerón, L. (2018). Don't throw away your printed books: a meta-analysis on the effects of reading media on reading comprehension. *Educ. Res. Rev.* 25, 23–38. doi: 10.1016/j.edurev.2018.09.003

Gonzalez, M. (2014). The effect of embedded text-to-speech and vocabulary eBook scaffolds on the comprehension of students with reading disabilities. *Intern. J. Spec. Educ.* 29, 111–125.

Greene, B. (2001). Testing reading comprehension of theoretical discourse with cloze. *J. Res. Read.* 24, 82–98. doi: 10.1111/1467-9817.00134

Hulme, C., and Snowling, M. J. (2011). Children's reading comprehension difficulties: nature, causes, and treatments. *Curr. Direct. Psychol. Sci.* 20, 139–142. doi: 10.1177/0963721411408673

IBM Corp (2017). *IBM SPSS Statistics for Windows, Version 25.0*. Armonk, NY: IBM Corp.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., and Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10081–10086. doi: 10.1073/pnas.1103228108

Johnson-Glenberg, M. C. (2005). Web-based training of metacognitive strategies for text comprehension: focus on poor comprehenders. *Read. Writ.* 18, 755–786. doi: 10.1007/s11145-005-0956-5

Kerr, M. A., and Symons, S. E. (2006). Computerized presentation of text: effects on children's reading of informational material. *Read. Writ.* 19, 1–19. doi: 10.1007/s11145-003-8128-y

Kim, A. H., Vaughn, S., Klingner, J. K., Woodruff, A. L., Klein Reutebuch, C., and Kouzekanani, K. (2006). Improving the reading comprehension of middle school students with disabilities through computer-assisted collaborative strategic reading. *Remed. Spec. Educ.* 27, 235–249. doi: 10.1177/07419325060270040401

Kleinsz, N., Potocki, A., Ecalle, J., and Magnan, A. (2017). Profiles of French poor readers: underlying difficulties and effects of computerized training programs. *Learn. Individ. Differ.* 57, 45–57. doi: 10.1016/j.lindif.2017.05.009

Krieger, R. (2017). *The Effect of Electronic Text Reading on Reading Comprehension Scores of Students with Disabilities*. Master thesis, Governors State University, Park, IL.

Leong, C. K. (1992). Enhancing reading comprehension with text-to-speech (DECtalk) computer system. *Read. Writ.* 4, 205–217. doi: 10.1007/bf01027492

Magliano, J. P., Todaro, S., Millis, K., Wiemer-Hastings, K., Kim, H. J., and McNamara, D. S. (2005). Changes in reading strategies as a function of reading training: a comparison of live and computerized training. *J. Educ. Comput. Res.* 32, 185–208. doi: 10.2190/1ln8-7bqe-8tn0-m91l

Mangen, A., Walgermo, B. R., and Brønnick, K. (2013). Reading linear texts on paper versus computer screen: effects on reading comprehension. *Intern. J. Educ. Res.* 58, 61–68. doi: 10.1016/j.ijer.2012.12.002

McNamara, D. S., O'Reilly, T. P., Best, R. M., and Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Intern. J. Educ. Res.* 34, 147–171. doi: 10.2190/1ru5-hdtj-a5c8-jvwe

Meyer, B. J., Wijekumar, K. K., and Lin, Y. C. (2011). Individualizing a web-based structure strategy intervention for fifth graders' comprehension of nonfiction. *J. Educ. Psychol.* 103, 140–168. doi: 10.1037/a0021606

Mich, O., Pianta, E., and Mana, N. (2013). Interactive stories and exercises with dynamic feedback for improving reading comprehension skills in deaf children. *Comput. Educ.* 65, 34–44. doi: 10.1016/j.compedu.2013.01.016

Moran, J., Ferdig, R. E., Pearson, P. D., Wardrop, J., and Blomeyer, R. L. Jr. (2008). Technology and reading performance in the middle-school grades: a meta-analysis with recommendations for policy and practice. *J. Liter. Res.* 40, 6–58. doi: 10.1080/10862960802070483

Niedo, J., Lee, Y. L., Breznitz, Z., and Berninger, V. W. (2014). Computerized silent reading rate and strategy instruction for fourth graders at risk in silent reading rate. *Learn. Disabil. Q.* 37, 100–110. doi: 10.1177/0731948713507263

Oakhill, J. V., and Cain, K. (2012). The precursors of reading ability in young readers: evidence from a four-year longitudinal study. *Sci. Stud. Read.* 162, 91–121. doi: 10.1080/10888438.2010.529219

Oakhill, J. V., Cain, K., and Bryant, P. E. (2003). The dissociation of word reading and text comprehension: evidence from component skills. *Lang. Cogn. Process.* 18, 443–468. doi: 10.1080/01690960344000008

O'Reilly, T. P., Sinclair, G. P., and McNamara, D. S. (2004). "iSTART: A web-based reading strategy intervention that improves students' science comprehension," in *Proceedings of the IADIS International Conference Cognition and Exploratory*

*Learning in Digital Age: CELDA 2004*, eds D. G. Kinshuk and P. Isaias (Lisbon: IADIS), 173–180.

Ortlieb, E., Sargent, S., and Moreland, M. (2014). Evaluating the efficacy of using a digital reading environment to improve reading comprehension within a reading clinic. *Read. Psychol.* 35, 397–421. doi: 10.1080/02702711.2012.683236

Potocki, A., Ecalle, J., and Magnan, A. (2013). Effects of computer-assisted comprehension training in less skilled comprehenders in second grade: a one-year follow-up study. *Comput. Educ.* 63, 131–140. doi: 10.1016/j.compedu.2012.12.011

Potocki, A., Magnan, A., and Ecalle, J. (2015). Computerized trainings in four groups of struggling readers: specific effects on word reading and comprehension. *Res. Dev. Disabil.* 45, 83–92. doi: 10.1016/j.ridd.2015./break07.016

Rideout, V. J., Foehr, U. G., and Roberts, D. F. (2010). *Generation M 2: Media in the Lives of 8-to 18-Year-Olds*. San Francisco, CA: Henry J. Kaiser Family Foundation.

Singer, L. M., and Alexander, P. A. (2017). Reading on paper and digitally: what the past decades of empirical research reveal. *Rev. Educ. Res.* 87, 1007–1041. doi: 10.3102/0034654317722961

Viola, F., and Carretti, B. (2019). Cambiamentinelleabilità di letturanelcorso di unostesso anno scolastico [Changes in readingskillsduring the sameschoolyear]. *Dislessia* 16, 147–159. doi: 10.14605/DIS1621902

von Bastian, C. C., and Oberauer, K. (2014). Effects and mechanisms of working memory training: a review. *Psychol. Res.* 78, 803–820. doi: 10.1007/s00426-013-0524-526

Wechsler, D. (2003). *Wechsler Intelligence Scale For Children–Fourth Edition (WISC-IV)*. San Antonio, TX: The Psychological Corporation.

Wijekumar, K. K., Meyer, B. J., and Lei, P. (2012). Large-scale randomized controlled trial with 4th graders using intelligent tutoring of the structure strategy to improve nonfiction reading comprehension. *Educ. Technol. Res. Dev.* 60, 987–1013. doi: 10.1007/s11423-012-9263-4

Wijekumar, K. K., Meyer, B. J., and Lei, P. (2013). High-fidelity implementation of web-based intelligent tutoring system improves fourth and fifth graders content area reading comprehension. *Comput. Educ.* 68, 366–379. doi: 10.1016/j.compedu.2013.05.021

Wijekumar, K. K., Meyer, B. J., and Lei, P. (2017). Web-based text structure strategy instruction improves seventh graders' content area reading comprehension. *J. Educ. Psychol.* 109, 741–760. doi: 10.1037/edu0000168

Woolley, G. (2011). "Reading comprehension," in *Reading Comprehension: Assisting Children With Learning Difficulties*, ed Springer Science+Business Media (Dordrecht, NL: Springer), doi: 10.1007/978-94-007-1174-7_2

Yuill, N., and Oakhill, J. (1988). Effects of inference awareness training on poor reading comprehension. *Appl. Cogn. Psychol.* 2, 33–45. doi: 10.1002/acp.2350020105

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 30 June 2020
doi: 10.3389/fpsyg.2020.01363

# Using Mobile Devices in Teaching Large University Classes: How Does It Affect Exam Success?

*Tommaso Feraco[1,2]*, Nicole Casali[1], Carla Tortora[1], Cecilia Dal Bon[3], Donatella Accarrino[3] and Chiara Meneghetti[1]**

*[1] Department of General Psychology, University of Padua, Padua, Italy, [2] Pentathlon Srl, Naples, Italy, [3] Digitial Learning Center, University of Padua, Padua, Italy*

The aim of this study was to examine the role of mobile-based student response systems in teaching to improve university students' academic outcomes. Mobile devices can be useful tools for conveying content to large classes, with a potential impact on academic outcomes. This study involved a total of 294 undergraduates taking a psychology course. The course involved lessons in the classroom, which included answering quizzes (quiz activities) and activities such as preparing reports and laboratory experiences (out-of-class activities). Quizzes were administered using a mobile technology system. Data were collected on the motivational beliefs (theory of intelligence) and self-regulated learning strategies of students who voluntarily completed the online questionnaires. The results of the linear models showed that using the quizzes positively affected the performance in the final exams (involving closed and open questions). The same was true for the out-of-class activities. The motivation and strategy scores correlated moderately with out-of-class activities, but not with quiz activities. These results offer insight on the use of technology during lessons and other course-related activities to promote academic achievement.

**Keywords: student response systems, technology-enhanced teaching, innovative teaching, academic achievement, large classes, Italian university**

OPEN ACCESS

***Edited by:***
*Maria Luisa Lorusso, Eugenio Medea (IRCCS), Italy*

***Reviewed by:***
*Heinrich Söbke, Bauhaus-Universität Weimar, Germany*
*Christopher A. Was, Kent State University, United States*

****Correspondence:***
*Tommaso Feraco*
*tommaso.feraco@phd.unipd.it*
*Chiara Meneghetti*
*chiara.meneghetti@unipd.it*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *25 March 2020*
***Accepted:*** *22 May 2020*
***Published:*** *30 June 2020*

***Citation:***
*Feraco T, Casali N, Tortora C, Dal Bon C, Accarrino D and Meneghetti C (2020) Using Mobile Devices in Teaching Large University Classes: How Does It Affect Exam Success? Front. Psychol. 11:1363. doi: 10.3389/fpsyg.2020.01363*

## INTRODUCTION

One of the main aims of higher education systems is to enable students to learn effectively. Teaching modalities and practices can play an important part in academic achievement throughout a student's university career (Caird and Lane, 2015). Among several teaching approaches, technology systems based on mobile devices are potentially powerful learning tools. The European Commission is increasingly advocating integration between technology and teaching to increase efficiency in higher education systems (e.g., Digital Education Action Plan; European Commission, 2018).

The interest in innovative teaching methods is nourished by a growing body of research showing the benefits of exploiting technologies during lectures at a university, as well as in various other higher-education settings, to learning and learning-related outcomes (Sung et al., 2016). These benefits are particularly important in contexts where university classes are large and/or attendance is not mandatory, making it more challenging for lecturers to encourage students' participation in class (Hunsu et al., 2016).

Among technology systems, mobile-based student response systems (SRS, also known as audience/classroom response systems or "clickers") are attracting particular attention as a way for teachers to interact actively with students (Kay and LeSage, 2009). SRS allow students to answer closed (multiple-choice or true/false) questions displayed on a screen during lessons by pressing a button on a keypad, mobile phone screen, or web-based interface (hence the name "clickers") and to receive immediate feedback. The literature on this topic is flourishing, with three recent meta-analyses on the influence of these methods on various learning domains, such as academic performance (Castillo-Manzano et al., 2016; Sung et al., 2016), and on related learning outcomes like engagement, which means students' involvement or participation in class (Hunsu et al., 2016). On average, SRS have shown a small-to-medium positive effect on achievement [$0.21 < g < 0.37$ in Hunsu et al. (2016), Castillo-Manzano et al. (2016), and Sung et al. (2016)] and students' participation in class ($g = 0.19$, Hunsu et al., 2016).

The positive effect of quizzes, administered through a SRS, on academic performance can be due to optimizing the learning process and stimulating cognitive mechanisms involved in learning, such as testing effect and feedback. The testing effect (e.g., Roediger and Karpicke, 2006; Roediger and Butler, 2011) is a well-known cognitive phenomenon according to which taking practice tests promotes learning and long-term retention better than other study strategies, such as re-reading. A recent meta-analysis (118 studies; Adesope et al., 2017) found that testing positively influences achievement ($g = 0.60$ at university level) and is more effective than re-reading ($g = 0.51$). Interestingly, the testing effect was not moderated by the presence or the absence of feedback, although there is evidence of immediate feedback yielding greater retention benefits than delayed feedback (Butler et al., 2007), and corrective feedback seems to enhance the testing effect (Rowland, 2014). The efficacy of SRS quizzes on achievement may therefore be explained by both testing effect and immediate feedback. Most studies on testing effect have focused on regular-sized classrooms, with fewer than 100 students (e.g., Roediger and Butler, 2011), while none—to our knowledge—have specifically investigated whether the testing effect is preserved in large classrooms. A moderator analysis from the meta-analysis by Hunsu et al. (2016) on SRS and achievement suggests that the positive effect of using clickers on performance declines with increasing classroom size (small effect with 100–200 students, $g = 0.10$, compared to $g = 0.58$ with 21–30 students). The testing effect (believed to explain the efficacy of SRS quizzes) may therefore be reduced in large university classes, too.

The present study aims to analyze the effect on academic performance of using interactive teaching practices in large classes where attendance is not mandatory.

Concerning interactive technologies, among several SRS-based quizzes available in the market (e.g., Kahoot!, Socrative, and Learning Catalytics), Top Hat[1] is one of the systems most often used in higher education nowadays. The Italian university where our study was conducted has been supporting its use for numerous courses (182 courses, involving a total of 12,612 students in 2018–2019). Top Hat has proved particularly effective and been much appreciated by students (Neilson et al., 2016; Kawash and Collier, 2018). It assures anonymity, allowing students to receive feedback without feeling exposed to the judgment of peers or teachers (Salzer, 2018). It reduces the use of smartphones and laptops in class for distraction purposes (Rozgonjuk et al., 2018) and enhances engagement in lectures (Kawash and Collier, 2018). Top Hat is also able to integrate quizzes (and other activities) with lecture content (such as PowerPoint presentations) with more refined features than other SRS systems, making quizzes a complementary part of the lesson in real time.

The present study specifically examined whether using SRS-based quizzes during university lectures can positively affect the final exam performance. If answering quizzes helps students to memorize information better than by just listening to lectures, as seen in studies on the testing effect (Roediger and Karpicke, 2006; Roediger and Butler, 2011), and providing immediate feedback on their answers can consolidate their learning (Rowland, 2014), then we can expect the quiz activities during lessons to have an impact on the final exam performance.

The use of quizzes together with other teaching methods was examined, given the potential efficacy of combining several types of activities (Sung et al., 2016). Among the various potentially activating teaching methods, we opted for out-of-class activities that further elaborated on the course content, such as laboratory experiences and preparing reports. These activities are designed to make learning of the course material more effective, with an impact on academic performance. To our knowledge, the effect of such complementary activities on academic performance has yet to be thoroughly examined, although there is some evidence supporting their usefulness (Heng, 2014; Ko et al., 2015; Ribeiro et al., 2019). For instance, Heng (2014) found a positive contribution of out-of-class educational activities (e.g., time spent on course-related tasks, reading before lectures) to undergraduates' academic performance, explaining around 20% of the variance in achievement (exam performance).

Participation in quiz activities and out-of-class activities was not mandatory, and a student's decision to exploit these opportunities might be a sign of more active engagement. Students' participation in class is an indicator of behavioral engagement, part of a multi-dimensional construct that includes cognitive (e.g., the cognitive aspects involved and the ability to regulate them) and affective (e.g., the type of emotion and motivation experienced) components [see Bond et al. (2020) for a review]. Therefore, since strategic and motivational aspects are considered as proxies of cognitive and emotional engagement (Bond et al., 2020), we examined whether students' self-regulated learning strategies (e.g., Zimmerman, 2000) and motivational beliefs, such as incremental theories of intelligence (Dweck, 2000), are related to their use of quizzes and out-of-class activities. We also investigated whether study strategies and motivation were positively associated with final exam performance [as seen by Lei et al. (2018), Mega et al. (2014), and Richardson et al. (2012), for instance]. To achieve this purpose, we invited students to answer online questionnaires.

[1] https://tophat.com/

# METHOD

## Participants

The sample consisted of 294 undergraduates (72 males; age range, 20–25 years) from two cohorts (153 from the 2018 cohort and 141 from the 2019 cohort) enrolled on a psychology course for a Bachelor's Degree in Psychology (for which 150 students had enrolled). The course lasts 42 h (11 weeks, 21 lessons, with two weekly sessions lasting 90 min each). Since we could not plan the number of participants in advance, we ran a retrospective power analysis *via* simulation on the number of participants collected to test the reliability of our results (Altoè et al., 2020). Hypothesizing a plausible correlation of 0.30 between exam results and the use of quizzes and out-of-class activities (Castillo-Manzano et al., 2016; Hunsu et al., 2016; Sung et al., 2016; Lei et al., 2018), 294 participants yielded a power of 0.84, with a plausible magnitude error of 0.72 and no sign error. The study was approved by the Ethics Committee for Research in Psychology at the University of Padova (N. 3519).

## Materials

### Activities

#### *Quiz activities*

Using Top Hat, a total of seven quiz sessions were planned, with three to 12 questions each ($M$ = 6.57, SD = 3.26), consisting of multiple-choice (77%), true/false (18%), or click-on-target (touching part of a figure, 5%) questions. During the lessons, the teacher launched Top Hat (using the teacher's credentials), and the students accessed the quiz on their own devices (smartphones, tablets, and laptops) using their academic credentials. Each question was shown on their screens together with the possible answers. After they had selected their answers, the right answer and the collective frequencies of right, wrong, and n/a answers were shown. All students' answers were recorded. The number of sessions where a student answered at least one question was used as a final measure (maximum: seven).

#### *Out-of-class activities*

Three other types of activity were proposed: (1) prepare a group presentation on a topic related to the course content (based on research or reviews published in scientific journals) and present it in class (group work), (2) participate in one or two laboratory studies (lab experiences), and (3) describe a person's psychological profile starting from their scores in personality questionnaires and then have the resulting description blind-judged by two classmates on clarity and correctness, all under the teacher's supervision (case report). These activities were presented during the first lesson and were not mandatory, but participation was rewarded with bonus points (one point for group work, one point for the first lab experience, one point for the second lab experience, and one point for case reports) on the students' final exam score. The number of activities that the students engaged in was counted as a final measure (maximum: four).

### Motivation and Strategy Questionnaires

#### *Theory of Intelligence Questionnaire*

This Theory of Intelligence Questionnaire (TIQ;Dweck, 2000; De Beni et al., 2014) consists of eight items that measure incremental and static theories of intelligence, i.e., beliefs about whether or not one's own intelligence can be modified (e.g., "You can learn new things, but you can't change your intelligence"). The answers are given on a six-point Likert scale (1 = "completely agree" to 6 = "completely disagree"). The internal consistency is good ($\alpha$ = 0.88).

#### *Self-Regulated Strategy Questionnaire*

This Self-Regulated Strategy Questionnaire (SRSQ; De Beni et al., 2014) contains 50 items that assess self-regulated learning strategies, i.e., organization, elaboration, self-evaluation, preparing for exams, and metacognition (e.g., "I like to think about how my mind works"). The answers are given on a five-point Likert scale (1 = "completely agree" to 5 = "completely disagree"). The internal consistency is good ($\alpha$ = 0.76).

### Exam Performance

#### *Exam*

The exam consists of two parts: 30 closed questions (one point for each correct answer, 0 point if left blank, and minus one for wrong answers; maximum = 30) that differ from those used in the quizzes, plus four mandatory open questions (7.5 points for each answer; total 30). The exam also includes a list of five optional open questions on specific course content that can earn up to six bonus points. The questions are selected from a large pool already available and double-checked for clarity by other professors. The mean score for the two parts is calculated as the student's exam performance. When students obtain a raw score $\geq$33, they are awarded the maximum possible score of 30 with honors. The data were only collected for the students' first exam attempt. The bonus points awarded for the out-of-class activities were not considered in the analyses.

## Procedure

The psychology course was held in the first semester of the academic year, from October to December. During the first lesson, the teacher (the last author) presented quiz activities and out-of-class activities, inviting students to register in the SRS. She explained that the impact of their use on exam performance would be the object of a study. All the students signed the consent form. The quizzes were administered in one of the two weekly lessons using the Top Hat system, which was integrated with the use of slides, presenting the course content, that were interspersed with the quizzes. After each question was answered, the feedback was projected on the students' personal devices and on the classroom projector screen, indicating the percentages of the respondents' right and wrong answers. Each 90-min lesson involving quizzes included one to three sets of mostly multiple-choice questions (3 to 12 questions in all) presented at the beginning, in the middle, and/or at the end of the lesson. Each question took up 2 to 5 min, considering the time taken to answer and then to project the feedback (i.e., 1 min per question) and the time spent discussing the answers. The time spent in each

lesson on the quizzes varied from around 6–15 to 24–60 min. The students in class completed the quizzes without any difficulty, although technical problems could sporadically occur (when logging out of the system). During the course, the students were also invited to complete the SRSQ and the TIQ using the university online platform (Moodle). This was not mandatory and was done by the students in their own time. At the end of the course, the students could decide when to take the exam on one of five dates in January, February, June, July, or September.

# RESULTS

## Preliminary Analysis

Of the 294 students, 187 participated in both the quiz and the out-of-class activities (63.60%), 55 were only involved in the latter (18.71%), and 52 in neither (17.69%); none of the students participated in quiz activities only. Concerning the out-of-class activities, of the 242 students taking part, 87 completed all four activities, 34 only produced a case report, 21 only participated in the group work, and only eight were involved in lab experiences, while 92 engaged in two or three different activities. The results of a *t*-test for the independent samples showed that the students who did not participate in any of these activities had lower final exam performance than those who did [$t(292) = 3.81$, $p < 0.001$; exam score: no activity, $N = 52$; $M = 20.01$; SD = 4.98; participating group, $N = 242$, $M = 22.86$, SD = 4.86). **Table 1** shows the means, standard deviations, and correlations between all measures.

## Effect of Quiz and Out-of-Class Activities on Exam Performance

A two-step mixed linear model was run with R (R Core Team, 2019), using the lmerTest package (Kuznetsova et al., 2017), to analyze the effects of quiz activities (number of Top Hat sessions attended; maximum: seven) and out-of-class activities (number of activities; maximum: four) on exam performance (i.e., mean raw scores for closed and open questions without bonus points). Given that different questions were used for different exam sessions, we added exam session as a random factor to control for general differences due only to random differences in the sessions' difficulty.

In the first step, the model was run adding only quiz activities as a predictor of exam performance. The quizzes had a significant effect ($\beta = 0.62$; $p < 0.001$) on exam results, explaining 9% of the variance (marginal). In the second step, we added out-of-class activities to the predictors of exam performance. The results showed a significant effect ($p < 0.001$) of both out-of-class activities ($\beta = 0.77$) and quiz activities ($\beta = 0.42$). The model explained 13% of the variance (marginal), adding 4% of variance compared with the first regression model. Given that the two predictors are not orthogonal, part of the variance explained is common between them.

To ensure a similar impact on the different types of exam question, the same analysis was run again twice, once considering the scores for the closed questions as dependent variables and once considering the scores for the open questions. The results showed a similar trend (see **Figure 1**). Although we were not interested in the students' accuracy in completing the quizzes (since they were used primarily to stimulate interaction with students), its effect (the percentage of correct answers) on exam performance (total, closed, and open questions) was ascertained and the results remained comparable with those presented above.

## Relationship Between Quiz Activities, Out-of-Class Activities, Exam Performance, Theories of Intelligence, and Self-Regulated Strategy Questionnaires

We analyzed the correlations among SRSQ, TIQ, quiz, and out-of-class activities. The two questionnaires were not mandatory, and correlations for the SRSQ and TIQ could be calculated on 134 and 103 students, respectively. The results (see **Table 1**) showed that the quiz activities did not significantly correlate with SRSQ and TIQ scores, while the out-of-class activities showed a tendency toward significance with both the SRSQ and the TIQ scores. Exam performance significantly correlated with SRSQ score (i.e., self-regulated strategies; $r = 0.21$).

**TABLE 1 |** Means, standard deviations, and correlations for all the variables included in the study.

| | M | SD | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|
| Exam performance | 22.36 | 5.00 | – | | | | |
| Quiz activities | 2.31 | 1.42 | 0.30*** | – | | | |
| Out-of-class activities | 2.38 | 2.48 | 0.32*** | 0.45*** | – | | |
| Theory of Intelligence Questionnaire | 28.19 | 2.31 | 0.11 | 0.07 | 0.19° | – | |
| Self-Regulated Strategy Questionnaire | 185.23 | 15.93 | 0.21* | 0.14 | 0.17° | −0.07 | – |

***$p < 0.001$, *$p < 0.05$, °$p = 0.06$.

# DISCUSSION

The results of the regression models showed that answering quizzes positively affected the final exam performance (9% of marginal variance explained). This is in line with the literature showing a positive effect on the academic performance of using SRS during lessons (Hunsu et al., 2016; Sung et al., 2016). The variance explained in the model is in line with the small effect of SRS on academic achievement in large classes ($g = 0.10$; Hunsu et al., 2016).

The effect of administering quizzes during lessons, supported by the use of an advanced SRS, can be explained by the testing effect, i.e., repeated quizzes in class make it easier to memorize the course content and foster its recall over time (Karpicke and Roediger, 2007). In the present study, the testing effect refers to the fact that being repeatedly tested on course content (through

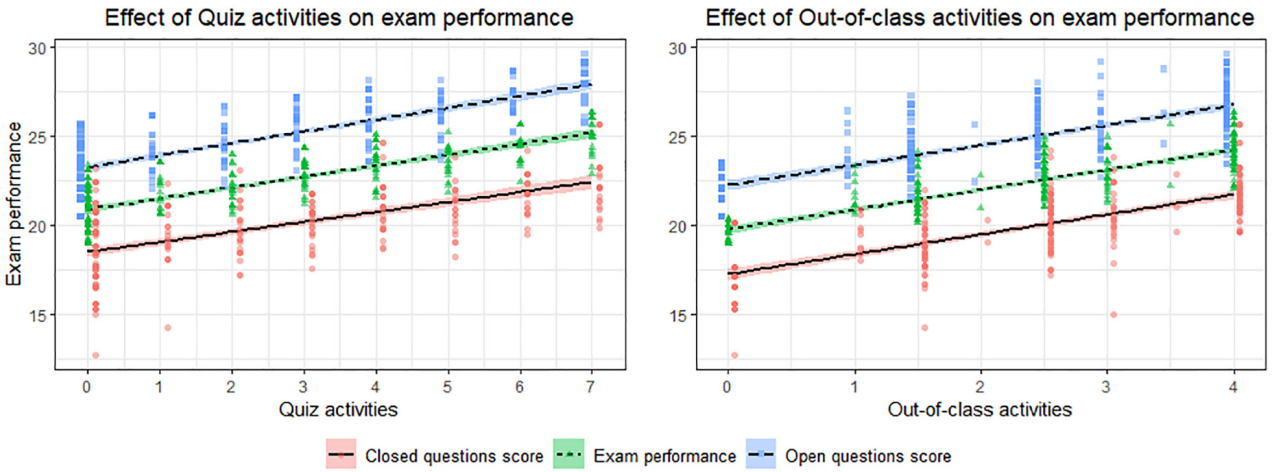


**FIGURE 1 |** Effects of quiz activities and out-of-class activities on exam performance by question type.

sample exam questions) during the lessons may have increased the retention of study material and support the performance during the final exam. The positive effect of quizzes may also be supported by feedback given after each answer (as in Rowland, 2014). It should be noted that this feedback was followed by class discussion to link the question to the content of the lesson. While this combination (quiz feedback and discussion) can be a useful strategy, it makes it difficult to disentangle whether and to what degree the positive effect of quizzes on exam performance was due to a specific role of feedback and/or related discussion.

The quizzes during lessons with SRS can be used to optimize the learning process by activating the students in class and enabling them to memorize what they learn more efficiently. In such contexts, where many students are grouped in the same room to hear a lecture and attendance is not mandatory, the use of technologies like SRS to administer quizzes during lessons appears to be beneficial to exam performance. Using a mobile-based technology to administer quizzes seems to have some advantages over the more traditional ways of asking questions in class (including anonymity and immediate feedback) or other technologies. It may have refined functions that enable the quizzes to be integrated with course material and make students more comfortable about following the lesson and answering questions (Kawash and Collier, 2018). Although the evidence emerging from our study cannot be generalized to mobile device use in lectures, it does support the benefits of administering quizzes, possibly *via* the students' own mobile devices, during lessons.

Our results concerning out-of-class activities that are proposed to the students also deserve to be discussed. Additional course-related activities improved the students' final exam performance and seemed to be just as effective as quizzes (as shown in **Figure 1**). The experience of preparing reports and participating in lab experiences had a positive effect on the students' academic achievement, confirming the importance of such complementary activities (as suggested by Heng, 2014) in the learning process. A further indication of the relevance of participating in an out-of-class activity comes from the correlations with the motivational and SRSQs: significant correlations emerged not only between exam performance and self-regulated strategies (in line with previous evidence; Richardson et al., 2012; Mega et al., 2014) but also with out-of-class activities. On the other hand, completing the quiz activities did not relate to such functional motivational beliefs and self-regulated learning strategies, possibly because attending classes and completing the quizzes remain important regardless of the students' self-reported strategies and motivational beliefs, whereas participating in out-of-class activities seemed more related to effective learning motivation and strategies. It may be that such additional activities (which demand extra time and effort) were more likely to be chosen by more motivated and strategic students. It is also plausible that only the more motivated students answered the non-mandatory questionnaires since motivation is often considered an antecedent of affective student engagement (Bond et al., 2020). As for using quizzes during lessons, this could have helped to foster active learning in class regardless of the students' motivation by enhancing information retention and, ultimately, performance (Roediger and Butler, 2011). Although the value of out-of-class activities was not the main focus of this paper, our intriguing findings warrant further study.

More research is needed to better understand these findings and overcome some limitations of our study. Among them, there is the fact that our sample could be biased, even if the large class where attendance is not mandatory represents an ecological setting and a common condition at a university. Since attendance and involvement in the quizzes and the out-of-class activities were voluntary, it may be that only students who were already motivated took part. It is worth noting that the quizzes seemed equally appealing to students with different levels of motivation, including the more "reluctant" ones (Graham et al., 2007). Although our sample may be biased *per se* and there was no control group of other students, a *post hoc* analysis on the students who did not take part at the proposed activities showed

that their average exam grades were lower than those of the students who did take part at least once. This suggests that quizzes and/or out-of-class activities can benefit performance in the final exams. Another limitation of our study concerns the fact that only two types of activating teaching method were investigated (i.e., quizzes and out-of-class activities) among the many available. In addition, the use of mobile-based quizzes is just one of a variety of applications for mobile devices in class (although it is one of the most advanced), so our results cannot be extended to their use for teaching in general.

As a final consideration, proposing quizzes and out-of-class activities seemed to promote class attendance as around 64% of students attended the lessons (registering with the Top Hat system) and took part in the out-of-class activities—a satisfactory proportion for a course for which attendance is not mandatory. This percentage is similar to that of other psychology courses (on average 65%) that also incentivize attendance. To some degree at least, this could be seen as an objective indicator of behavioral engagement (Ribeiro et al., 2019; Bond et al., 2020). Previous studies also found SRS use and out-of-class activities to be positively associated with engagement (Heng, 2014; Bond et al., 2020) and, in turn, with achievement (Lei et al., 2018). The use of SRS may promote active engagement with learning material, thereby indirectly influencing achievement (Blasco-Arcas et al., 2013; Hunsu et al., 2016). Whether and how quizzes and additional course-related activities encourage engagement through active participation deserves to be systematically investigated to shed light on the role of individual characteristics (such as motivational/strategic approach), technology use (i.e., SRS), and different types of additional activity in promoting various types of student engagement (not only behavioral but also cognitive and affective; Bond et al., 2020).

## CONCLUSION

Our study contributes to expanding what is known about the use of different teaching methods in higher education systems with large cohorts of students and voluntary attendance, as in this Italian experience included in this Special Issue. Although more systematic evidence will be needed, these results suggest that using technology to improve learning in class, based on quizzes during lessons and out-of-class activities, has a positive effect on the students' exam performance. These results suggest that a combination of activating teaching modalities such as those used in the present study can be a feasible way to foster academic achievement in higher education.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories: doi: 10.6084/m9.figshare.11984361.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee for Research in Psychology (University of Padova). The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

DA and CD provided the technical support on the use of Top Hat/SRS. TF performed the statistical analysis. TF and CT organized the database. NC wrote the first draft of the manuscript. TF, CM, and CT contributed to the sections of the manuscript. All authors contributed to the conception and the design of the study, contributed to revising the manuscript, and have read and approved the submitted version.

## FUNDING

The present work was conducted as part of the Dipartimenti di Eccellenza research program (art.1, commi 314-337 legge 232/2016), supported by a grant from MIUR to the Department of General Psychology, University of Padua.

## REFERENCES

Adesope, O. O., Trevisan, D. A., and Sundararajan, N. (2017). Rethinking the use of tests: a meta-analysis of practice testing. *Rev. Educ. Res.* 87, 659–701. doi: 10.3102/0034654316689306

Altoè, G., Bertoldo, G., Zandonella Callegher, C., Toffalini, E., Calcagnì, A., Finos, L., et al. (2020). Enhancing statistical inference in psychological research via prospective and retrospective design analysis. *Front. Psychol.* 10:2893. doi: 10.3389/fpsyg.2019.02893

Blasco-Arcas, L., Buil, I., Hernández-Ortega, B., and Sese, F. J. (2013). Using clickers in class. The role of interactivity, active collaborative learning and engagement in learning performance. *Comput. Educ.*62, 102–110. doi: 10.1016/j.compedu.2012.10.019

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., and Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: a systematic evidence map. *Int. J. Educ. Technol. Higher Educ.* 17:2. doi: 10.1186/s41239-019-0176-8

Butler, A. C., Karpicke, J. D., and Roediger, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *J. Exp. Psychol.* 13, 273–281. doi: 10.1037/1076-898x.13.4.273

Caird, S., and Lane, A. (2015). Conceptualising the role of information and communication technologies in the design of higher education teaching models used in the UK. *Br. J. Educ. Technol.* 46, 58–70. doi: 10.1111/bjet.12123

Castillo-Manzano, J. I., Castro-Nuño, M., López-Valpuesta, L., Sanz-Díaz, M. T., and Yñiguez, R. (2016). Measuring the effect of ARS on academic performance: a global meta-analysis. *Comput. Educ.* 96, 109–121. doi: 10.1016/j.compedu.2016.02.007

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R foundation for Statistical Computing.

De Beni, R., Zamperlin, C., Meneghetti, C., Cornoldi, C., Fabris, M., Tona, G. D. M., et al. (2014). *Test AMOS-Abilitá e Motivazione allo Studio: prove di Valutazione e Orientamento per la Scuola Secondaria di Secondo Grado e L'universitá: Nuova edizione*. Trento: Edizioni Centro Studi Erickson

Dweck, C. S. (2000). *Self-Theories: Their Role in Motivation, Personality, and Development*. New York, NY: Psychology Press.

European Commission (2018). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the Digital Education Action Plan. Available online at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A22%3AFIN

Graham, C. R., Tripp, T. R., Seawright, L., and Joeckel, G. (2007). Empowering or compelling reluctant participators using audience response systems. *Act. Learn. Higher Educ.* 8, 233–258. doi: 10.1177/1469787407081885

Heng, K. (2014). The relationships between student engagement and the academic achievement of first-year university students in Cambodia. *Asia Pac. Educ. Res.* 23, 179–189. doi: 10.1007/s40299-013-0095-8

Hunsu, N. J., Adesope, O., and Bayly, D. J. (2016). A meta-analysis of the effects of audience response systems (clicker-based technologies) on cognition and affect. *Comput. Educ.* 94, 102–119. doi: 10.1016/j.compedu.2015.11.013

Karpicke, J. D., and Roediger, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *J. Mem. Lang.* 57, 151–162. doi: 10.1016/j.jml.2006.09.004

Kawash, J., and Collier, R. D. (2018). "On the use of classroom response systems as an integral part of the classroom," in *Proceedings of the 10th International Conference on Computer Supported Education*, (Setúbal: SciTePress), 38–46. doi: 10.5220/0006668500380046

Kay, R. H., and LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: a review of the literature. *Comput. Educ.* 53, 819–827. doi: 10.1016/j.compedu.2009.05.001

Ko, J. W., Park, S., Yu, H. S., Kim, S.-J., and Kim, D. M. (2015). The structural relationship between student engagement and learning outcomes in Korea. *Asia Pac. Educ. Res.* 25, 147–157. doi: 10.1007/s40299-015-0245-2

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Lei, H., Cui, Y., and Zhou, W. (2018). Relationships between student engagement and academic achievement: a meta-analysis. *Soc. Behav. Personal.* 46, 517–528. doi: 10.2224/sbp.7054

Mega, C., Ronconi, L., and De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *J. Educ. Psychol.* 106, 121–131. doi: 10.1037/a0033546

Neilson, M. M., Mafi, G., Pfeiffer, K., VanOverbeke, D., and Ramanathan, R. (2016). Students' perceptions regarding the use of tophat as an interactive tool in meat science class. *Meat Sci.* 112, 117–118. doi: 10.1016/j.meatsci.2015.08.043

Ribeiro, L., Rosário, P., Núñez, J. C., Gaeta, M., and Fuentes, S. (2019). First-year students background and academic achievement: the mediating role of student engagement. *Front. Psychol.* 10:2669. doi: 10.3389/fpsyg.2019.02669

Richardson, M., Abraham, C., and Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol.Bull.* 138, 353–387. doi: 10.1037/a0026838

Roediger, H. L., and Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends Cogn. Sci.* 15, 20–27. doi: 10.1016/j.tics.2010.09.003

Roediger, H. L., and Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249–255. doi: 10.1111/j.1467-9280.2006.01693.x

Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol. Bull.* 140, 1432–1463. doi: 10.1037/a0037559

Rozgonjuk, D., Saal, K., and Täht, K. (2018). Problematic smartphone use, deep and surface approaches to learning, and social media use in lectures. *Int. J. Environ. Res. Public Health* 15:92. doi: 10.3390/ijerph15010092

Salzer, R. (2018). Smartphones as audience response systems for lectures and seminars. *Anal. Bioanal. Chem.* 410, 1609–1613. doi: 10.1007/s00216-017-0794-8

Sung, Y. T., Chang, K. E., and Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: a meta-analysis and research synthesis. *Comput. Educ.* 94, 252–275. doi: 10.1016/j.compedu.2015.11.008

Zimmerman, B. J. (2000). "Attaining self-regulation: a social cognitive perspective," in *Handbook of Self-Regulation*, eds M. Boekaerts, P. R. Pintrich, and M. Zeidner (Cambridge, MA: Academic Press), 13–39. doi: 10.1016/b978-012109890-2/50031-7

**Conflict of Interest:** TF was employed by the company Pentathlon Srl.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 16 July 2020
doi: 10.3389/fpsyg.2020.01702

# A Novel Approach to Measure Executive Functions in Students: An Evaluation of Two Child-Friendly Apps

*Valeska Berg[1*†], Shane L. Rogers[1], Mark McMahon[1], Michael Garrett[1,2] and Dominic Manley[2]*

[1] School of Arts and Humanities, Edith Cowan University, Perth, WA, Australia, [2] Cinglevue International, Perth, WA, Australia

Interest in measurement of children's executive functions has shown a major increase over the past two decades. The present study evaluates two new apps (EYT and eFun) for measuring executive functions in children. The results of this study show that children (aged 5–8) enjoy executive function assessment in the form of games on an iPad. However, only one executive function, EYT working memory, showed significant positive correlations with several types of grades (e.g., English and maths) in primary school students. New, self-assessed, child-friendly executive function measurement tools have the potential to provide future possibilities for teachers to integrate information on cognitive ability into student learning plans.

**Keywords: executive functions, cognitive assessment, educational technology, students, classroom, cognitive functions, cognition, assessment**

OPEN ACCESS

***Edited by:***
*Jie Chi Yang,*
*National Central University, Taiwan*

***Reviewed by:***
*Gail C. Tiemann,*
*University of Kansas, United States*
*Nicos Souleles,*
*Cyprus University of Technology, Cyprus*

****Correspondence:***
*Valeska Berg*
*v.berg@ecu.edu.au;*
*Valeska.berg1@gmail.com*

***†ORCID:***
*Valeska Berg*
*orcid.org/0000-0002-2156-0969*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *07 April 2020*
***Accepted:*** *22 June 2020*
***Published:*** *16 July 2020*

***Citation:***
*Berg V, Rogers SL, McMahon M, Garrett M and Manley D (2020) A Novel Approach to Measure Executive Functions in Students: An Evaluation of Two Child-Friendly Apps. Front. Psychol. 11:1702. doi: 10.3389/fpsyg.2020.01702*

## INTRODUCTION

Over the past two decades, researchers from cognitive psychology, neuroscience and education have increasingly shown interest in measuring executive functions (EFs) in children (e.g. Miyake et al., 2000; Diamond, 2016; Zelazo et al., 2016). Recent attempts have been made to design child-friendly tools to measure EFs in children but a number of challenges persist (Zelazo, 2015; Howard and Melhuish, 2017; Józsa et al., 2017; Obradović et al., 2018; Holmboe et al., 2019; Willoughby et al., 2019). The current study presents the first step toward validating newly developed executive function tools. This study evaluates two novel child-friendly EF measurement tools specifically developed to assess EFs in the classroom. These are the Early Years Toolbox (EYT) developed by Howard and Melhuish (2017), and a new set of executive function assessment tasks called eFun (Berg et al., 2019). Both use technology-enhanced tasks rather than traditional paper-based assessments. A key consideration of our research was the children's experience with these tasks as fun and enjoyable activities since it has been shown that enjoyment with a task can lead to greater task performance (Schukajlow and Krug, 2014).

To evaluate these tasks the following research questions were investigated: (a) Do children find the EF tasks enjoyable? (b) Does task performance across and within the two apps show positive associations? (c) Does the performance on the EF tasks positively associate with primary school grades?

The term "executive functions" is used to describe a set of interrelated cognitive processes that enable us to accomplish tasks and pursue goals by controlling cognition and behavior in a goal-directed manner (e.g., Gioia et al., 2001). The core triad of EFs are working memory, inhibition and cognitive flexibility (Miyake et al., 2000; Lehto et al., 2003; Diamond, 2013). Working memory (WM) is the ability to hold information in mind and manipulate it (Baddeley, 1992). WM is essential for remembering task requirements and organizing action plans (Diamond, 2013). Inhibition refers to the ability to deliberately stop or inhibit dominant or automatic behaviors and/or thoughts (e.g., Mäntylä et al., 2010). Inhibition is required to withhold inappropriate responses or thoughts and sustain attention to the task at hand. Cognitive flexibility is the ability to flexibly switch between and apply new and existing rules (Zelazo, 2015). This core EF is active when a situation or context requires the application or adaptation of a (new) set of rules. While these EFs are typically discussed as if they are relatively distinct, it is widely accepted in the literature that in everyday life activities EFs work together (Diamond, 2016; Zelazo et al., 2016).

When working in concert, these core executive functions allow a person to engage in the more complex cognitive processes necessary to solve everyday problems, such as planning and evaluating (Zelazo et al., 1997; Diamond, 2016). For example, when solving a task in school, rules or guidelines are kept in working memory and applied where needed. Additionally, distractors that interfere with a task are inhibited and flexible thinking is applied when switching between rules appropriate to the task at hand.

Indeed, executive functions build a foundation for learning and academic success (Posner and Rothbart, 2007; Zelazo et al., 2016). Executive functions predict math, English and science achievements in school (Bull and Scerif, 2001; St Clair-Thompson and Gathercole, 2006; Memisevic et al., 2018; Usai et al., 2018). Research has shown that teachers value their students' EF capacities. For example, following classroom instructions while inhibiting distractions has been identified by teachers as a key element in a successful classroom (Lin et al., 2003). When not developed properly, EFs can result in learning difficulties. Executive function deficiencies can turn into severe behavioral issues including aggression, emotional disturbance, and criminality (Broidy et al., 2003; Denson et al., 2011). Therefore, it is not surprising that EFs are associated with overall quality of life (Moffitt et al., 2011).

Given the tremendous influence EFs have on success in school and life, EF levels need to be identified and addressed in a valid and reliable manner, with tools that are appropriate for use with children. The information gained from appropriate EF measurement tools can potentially provide teachers with a deeper understanding of students' learning skills, which could, in turn, form the basis for individual learning plans in the future (Bierman et al., 2008; Flook et al., 2010; Diamond, 2012).

There are two main challenges for scholars who are interested in developing EF measurement tools for children. Firstly, EF tasks need to be child-friendly. The original executive function tasks designed for adults do not take into account the level of reading and writing ability a child possesses, or the limited attention span of a child. For example, one version of the traditional Go/No Go Task for adults requires the participant to respond to 600 stimuli, which can take up to 30 min to complete (Hackley et al., 1990). Other EF tasks, like the Stroop task (MacLeod, 1991) require the participants to read words. Therefore, researchers have recently tried to design new EF tasks to be more child-friendly by shortening task length and adapting the design and delivery method (e.g. Howard and Melhuish, 2017; Holmboe et al., 2019). Task instructions can be tailored for children by using child-friendly language to make it easier to understand and more engaging for children. Furthermore, tablets/computerized tasks allow for verbal (standardized) instructions given via headphones alongside visual interactive instructions on the screen. This eliminates instructor bias and limits the cognitive demands associated with social interactions. Nevertheless, although recent EF tasks attempt to be child-friendly, there is a lack of studies that evaluate how the child experiences the tasks.

If children enjoy playing the task it is more likely that they pay attention to the task, which can influence their performance on the task. In order to capture children's attention and measure their full potential on the EF tasks, the eFun tasks were designed to be an enjoyable experience that the children like to play. Research has found that students that show enjoyment and interest in performance tasks score higher on the performance tasks (Schukajlow and Krug, 2014). Furthermore, task enjoyment has been found to be positively associated with attention and task persistence (Reeve, 1989; Engelmann and Pessoa, 2014), which leads to enhanced performance (Engelmann et al., 2009; Pessoa and Engelmann, 2010). Thus, for researchers designing new EF tasks, the challenge is not just the design of the task, but also the evaluation of children's experience with the task. This lack of evidence to support assertions of child-friendliness is a common issue among tasks that have been recently designed for children (Cianchetti et al., 2007; Zelazo, 2015; De Greeff et al., 2016; Howard and Melhuish, 2017; Dawson and Guare, 2018; Holmboe et al., 2019; Willoughby et al., 2019).

The second challenge is that EF tasks need to be modified to suit *non-clinical* populations. EF tasks were originally developed to diagnose a small number of people with severe cognitive dysfunctions in a clinical context (e.g., Otto et al., 1991; Gold et al., 1997). However, identifying EF levels in *typically developing* children has recently attracted interest in research that aims to support children's cognitive development (Howard and Melhuish, 2017; Holmboe et al., 2019). Therefore, there is a need to adjust the difficulty levels of the tasks to capture varying levels of EFs rather than only capturing severe executive dysfunction. Additionally, the initial clinical EF assessment tasks were originally designed to be conducted in decontextualized clinical settings that do not reflect how EFs operate in the everyday life of a child (Wallisch et al., 2018). This is problematic not only because the ecological validity of the tasks is low, but also because a child might not feel comfortable in an unfamiliar environment with one examiner assessing the child. This environment can induce stress or (test) anxiety in a child which may affect test performance (Shute et al., 2016).

A few recent attempts have been made to design child-friendly tools in order to measure EFs in children (Kado et al., 2012;

Diamond and Wright, 2014; Zelazo, 2015; Howard and Melhuish, 2017; Józsa et al., 2017; Holmboe et al., 2019; Willoughby et al., 2019). Cognitive demands and assessment methods have been adjusted to make tasks more appropriate for children. For example, slight variations have been made to the stimuli and administration procedures and the length of EF tasks have been reduced to account for children's limited attention span and (e.g., Howard and Melhuish, 2017).

To make EF tasks more appealing to children, several researchers have decided to use tablets instead of computers or physical tasks (Zelazo, 2015; Howard and Melhuish, 2017; Józsa et al., 2017; Holmboe et al., 2019; Willoughby et al., 2019). Using a tablet instead of a computer has several advantages (Falloon, 2013). Firstly, tablets require less attentional demands. The response location is on the tablet screen and not the computer keyboard, which means that the participants do not need to reorient their attention away from the computer screen to the keyboard. Reorienting attention can result in both additional time and effort, especially for children (Posner and Cohen, 1984; Hunt and Kingstone, 2003). Research confirms the benefits of this approach, showing that using a tablet instead of a computer is a more reliable measurement method for EF assessment in children, eliciting faster and better performances (Howard and Okely, 2015). Furthermore, tablets are mobile and can therefore be applied to different contexts, and they give the opportunity for self-administered testing, which eliminates instructor bias and costly instructor training. Finally, using an online-connected tool like a tablet enables fast data collection that can be uploaded and analyzed in a more efficient way than the traditional pen and paper recordings (Willoughby et al., 2019). Given the advantages of using a tablet, the current study employed this way of measurement to assess EFs in children.

## The Current Study, Evaluating Two New Child-Friendly Executive Function Measurement Tools

The current study is a beginning set of validity studies that presents the evaluation of two new EF measurement tools, the EYT (Howard and Melhuish, 2017) and eFun (Berg et al., 2019). The validation of educational and psychological test results is an ongoing process that requires multiple sources of evidence, with multiple samples (i.e., replication; Kane, 2013; American Educational Research Association et al., 2014). The present study represents an early stage in the validation assessment of eFun and EYT. Howard and Melhuish's (2017) EYT consists of a group of tasks to measure executive functions in a child-friendly way with tasks that are short and easy to understand for younger children. The EYT consists of several publically available 2D EF apps. To measure the three core executive functions the toolbox has two working memory games, one inhibition task, and one cognitive flexibility task.

The EYT cognitive flexibility task "Card Sorting" is similar to the iPad version of the Dimensional Change Card Sort (DCCS) by Zelazo (2015). However, the EYT Card Sorting task requires less assistance than the DCCS because all instructions are given verbally through the app. In the visual-spatial working memory task called "Mr. Ant" children are asked to remember locations of dots on an ant. Dots are shown on the body of the ant and after a short delay children are asked to replicate the sequence of the previous shown dots. The second working memory task is called "Not This". Children are presented with a number of different shapes that have cartoon faces on them. These vary in shape, size and color (e.g., small red triangle; large green circle). Children are given instructions to point to a shape that does not have a certain features (e.g., pointing to a shape that is not green/large/a circle). In lower levels children are asked to only hold one feature in mind, e.g., "find a shape that is not red". In subsequent higher levels children are required to hold up to three features in mind, e.g., "find a shape that is not small, not blue and not a circle". This task requires close monitoring by instructors, whereas the Mr. Ant WM task does not. The EYT inhibition task is adapted from the original Go/No-Go task (Donders, 1969). Children are either presented with a fish or a shark swimming from left to right on an iPad screen. Children are required to respond to the fish ("catch the fish"; go trial) by tapping the screen and to refrain from responding when the shark is displayed ("avoid catching sharks"; no-go trial). The majority (80%) of stimuli are fish (go trials), to generate a prepotent tendency to respond, while the tendency to respond has to be inhibited when the sharks are presented (the remaining 20%, which are no-go trials), for a detailed description see Howard and Okely (2015).

In the current study we also evaluated a newly developed self-assessed tablet EF measurement tool called "eFun" (Berg et al., 2019). All eFun tasks are based on established EF tasks which provided the foundation for the content of the eFun tasks. The newly built 3D eFun tasks are designed to measure executive functions in an engaging and child-friendly way in typically developing children. To make the eFun narration child-friendly, no numerical or letter knowledge is required. Furthermore, the tasks are brief and include dynamic elements to engage children's attention. The eFun tasks were developed in collaboration with an educational software company using the Unity game engine, deployed on an Apple IOS tablet (iPad). A team of trained software engineers and researchers have collaborated on the eFun app to make it both engaging and based on principles from cognitive psychology.

In order to engage children, the eFun tasks offer advanced design elements with a variety of response mechanics (e.g., swiping, dragging and touching), along with high graphical fidelity and a 3D environment (see **Figure 1**). To eliminate instructor bias the eFun tasks are self-assessed through verbal instructions given via headphones. A narrator called "Owly" guides the child through the eFun winter world. Every task includes a story that outlines the overarching goal of the task. For example, in the inhibition game called "Log Chop" the child is asked to chop firewood to help keep the eFun villagers warm during an icy cold storm that hits the village. The logs/firewood are the "go" stimuli that need to be swiped in the inhibition task (for a detailed explanation of the tasks see the measures section). At the end of each task the children can see that they have achieved their goal via an end game screen (e.g., the villagers sit around the fire made of chopped wood).


**FIGURE 1 |** The eFun map screen showing the progressing through the three games. The Ice Steps game which assesses working memory, the Log Chop game which assesses inhibition, and the Ice Cube Sorting game which assesses cognitive flexibility.

Prior research suggests that providing a narrative has the potential to foster greater involvement with the task (Axelsson et al., 2016), especially if the participants are given a goal to work toward (Dickey, 2006; Lim et al., 2014). Furthermore, having a goal to work toward resembles real-life more closely and aligns with the definition of EFs being goal-directed processes (Gioia et al., 2001). In addition to that, including game elements in EF tasks has the potential to enhance intrinsic motivation. Dörrenbächer et al. (2014) found that adding game elements to computerized task-switching training enhanced intrinsic interest in the task. Thus, providing a narrative with clear goals in a game environment is thought to heighten involvement with tasks and increase ecological validity. However, this has not been previously researched in an executive function assessment context with children. Furthermore, unlike clinical tasks, the eFun games measure executive function in a classroom context that is familiar to the child and resembles real-life situations of the child more closely.

The current study contrasts children's performance and experience of eFun with Howard and Melhuish's (2017) EYT tasks. It was hypothesized that the inclusion of dynamic game elements and the overarching narrative of eFun would result in the participants reporting the eFun tasks as more engaging than the EYT tasks. Based on Howard and Melhuish (2017) findings, moderate to positive associations between matching tests *across* the two apps were expected (i.e., $\text{Inhibition}_{\text{EYT}}$ and $\text{Inhibition}_{\text{eFun}}$, $\text{CF}_{\text{EYT}}$ and $\text{CF}_{\text{eFun}}$, $\text{WM}_{\text{EYT}}$, and $\text{WM}_{\text{eFun}}$). Furthermore, the current study investigates the relationships between the core EFs *within* both the EYT and the eFun tasks. Based on the findings of Howard and Melhuish (2017) it was expected that the three EYT tasks will yield significant moderate inter-correlations among each other. Miyake et al. (2000) and Lehto et al. (2003) have shown that the core EFs are moderately correlated but clearly distinguishable in adults and 8–13 year old children, thus similar results were expected for the three eFun tasks in our sample (Miyake et al., 2000; Diamond, 2016). It was also hypothesized that the core executive functions measured with eFun and the EYT would predict school grades. The literature suggests that EFs positively associate with academic outcomes in school children (Bull and Scerif, 2001; St Clair-Thompson and Gathercole, 2006; Memisevic et al., 2018; Usai et al., 2018). In particular, working memory was expected to show the strongest positive correlation with academic outcomes (Bull and Scerif, 2001; St Clair-Thompson and Gathercole, 2006; Usai et al., 2018).

# MATERIALS AND METHODS

## Participants

Two first and two second-year classes with a total of 81 students (54% girls; $M_{\text{age}}$ = 6.98) from a primary school in Western Australia participated in the current study. Participants were recruited through the school with information letters that were handed out to parents by teachers (see procedure). Seventy-one participants were included in the analysis of the task ratings and 74 were included in the analysis of the task outcomes. Ten cases were excluded due to incomplete data in the EF task outcome data and seven cases were excluded due to incomplete data in the questionnaire data.

## Materials

Executive functions were measured with two test batteries on iPads. Three of the previously developed tasks in the EYT (Howard and Melhuish, 2017) and the newly developed eFun app (including all three tasks), were used in this study. The EYT can be downloaded from the iTunes app store. The EYT tasks have previously been reported to possess good reliability and validity (Howard and Melhuish, 2017). The iPads used in this study were provided by the school. Pen and paper was used to get feedback from the children on the EYT tasks (see the feedback questionnaire in the measures section below), whereas for the eFun tasks the feedback questionnaires were integrated into the app.

## Measures

### EYT Tasks

These tasks are child-friendly executive function tasks that are part of the EYT developed by Howard and Melhuish (2017). For a more detailed description of the tasks beyond what is provided in this paper please see Howard and Melhuish (2017).

### EYT Mr. Ant Task (Working Memory)

This task requires children to remember the spatial locations of dots on a cartoon ant. Working memory capacity is recorded as a point score (Morra, 1994). This is calculated by assigning one point for each consecutive level in which at least two of the three trials were answered correctly and 1/3 of a point for all correct trials thereafter.

### EYT Go/No-Go Task (Inhibition)

This task requires children to respond to fish (go trial) by tapping the screen and to refrain from responding to sharks (no-go

trial) that are swimming from the left to the right side of the screen. For the analyses, trials in which the response is faster than 300 milliseconds were removed, because Howard and Melhuish (2017) suggest that responses that are this fast are not likely to be in response to the target stimulus. This elimination is crucial for this task because it is not required to touch the stimuli on the screen to indicate a response; instead tapping the screen in any location is recorded as a response. Furthermore, if the participant does not respond to the majority of stimuli within one level (go accuracy below 20% and no-go accuracy exceeds 80%), or if the participant responds to all stimuli within one level (go accuracy exceeds 80% and no-go accuracy below 20%) then their data for that level is excluded (Howard and Melhuish, 2017).

### EYT Card Sorting Task (Cognitive Flexibility)

In this task, children are asked to sort cards (e.g., red rabbits sitting on a raft, and blue boats) according to either color or shape into one of two locations: A castle with a flag displaying a blue rabbit, or a castle with a flag displaying a red boat. Scores are based on the number of correct sorts after the pre-switch phase.

### eFun Tasks

These executive function tasks are based on established EF tasks (e.g. Donders, 1969; Milner and Taylor, 1972; Kado et al., 2012) and were newly developed as part of this research (Berg et al., 2019). Three discrete EF tasks were developed, based on the theoretical model by Diamond (2013) who proposes three core executive functions, and research by Miyake et al. (2000) and Lehto et al. (2003) showing that these three core EFs are separable.

### eFun Ice Steps Task (Working Memory)

The Ice Steps task is based on the backward Corsi Block test (Milner and Taylor, 1972) and aims to measure working memory in children. The task starts with a brief introductory story explaining that Pongo the penguin needs to get fish for his chicks, which are on the opposite side of a river. The participant sees the hungry chicks crying for food. Next, the participant sees the penguin crossing a river on ice floats and getting fish for his chicks. The participant is asked to bring the penguin, with the collected fish, back to the chicks on the other side of the river. In order to get back to the other side of the river, the participant must remember the previously shown ice floats (organized in a grid, 3 × 8) in reversed order (see **Figure 2**). Following an explanation of how the game works, the child is asked to start the practice trial by tapping the series of ice floats the penguin previously used to cross the river in reversed order. The game starts with three ice floats to be remembered, increasing by one float to be remembered in each subsequent level. Each level consists of four trials with the same number of floats. In total there are four levels and the highest number of floats to be remembered is six. If the child gets three trials in one level wrong, the game is discontinued with a rewarding screen showing how the chicks get fed with the collected fish. Independent of the child's performance, the task always ends with the rewarding screen.

Working memory is assessed with three measures that are based on scoring procedures of the backward Corsi Block test (Kessels et al., 2008). First, the longest sequence of ice floats that is correctly remembered backward (i.e., the span length; 3–6 floats) is recorded. Secondly, the number of correctly remembered trials (4 levels × 4 trials) is measured. Lastly, these two measures are combined with a product score, which is the span length multiplied by the number of correctly remembered trials.

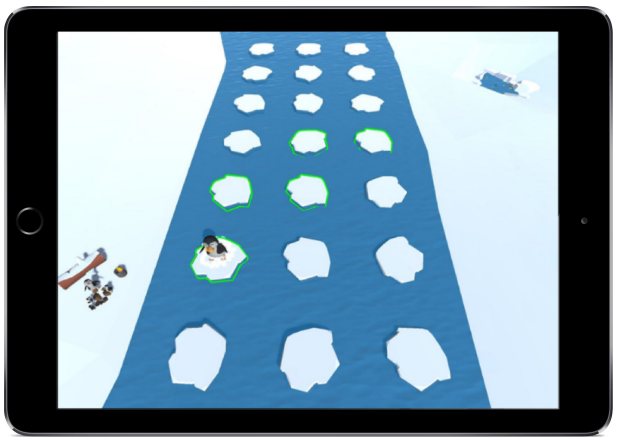

**FIGURE 2 |** The eFun working memory game (Ice Steps) showing how Pongo the penguin crosses the river to feed the hungry chicks.

### eFun Log Chop Task (Inhibition)

The Log Chop task is based on the Go/No-Go tasks (Donders, 1969; Simpson and Riggs, 2006; Wiebe et al., 2012; Howard and Okely, 2015) and aims to measure inhibition in children. The task starts with a brief introductory story explaining that a storm has hit the eFun village and in order to keep the villagers warm the child needs to chop (swipe across) descending fire wood while avoiding reacting to descending icicles. After a practice phase, in which the child learns to respond to the wood but not the icicles, the actual game starts, see **Figure 3**. Independent of the child's performance, the task always ends with a rewarding screen showing the characters around a fire made of the wood that the child had chopped.

The majority of stimuli in the log chop task are "go" stimuli (80% logs) to create a pre-potent tendency to respond. The task consists of three levels with increasing difficulty. Each level consists of 25 stimuli, 80% "go" stimuli (logs) and 20% "no-go" stimuli (icicles) and stimulus presentation (1.5 s) is held constant during all levels (based on Howard and Okely, 2015). The time between stimuli (the interstimulus interval, ISI) decreases from 1.5 s in level one to 1 s in level two and 0.5 s in level three. The decreasing ISIs act to speed up the task to increase difficulty. Additionally, no level starts with a "no-go" stimulus (icicle; Howard and Okely, 2015). The WM load is held constant throughout the game since the same rules apply in all levels and the rules are easy to remember (i.e., chop the logs, avoid the sharp icicles). Furthermore, the logs do not always appear in the same locations, thus responses cannot be anticipated (Aron, 2011; Diamond, 2013). Inhibition is indexed by the product of proportional go accuracy and proportional no-go accuracy (% go

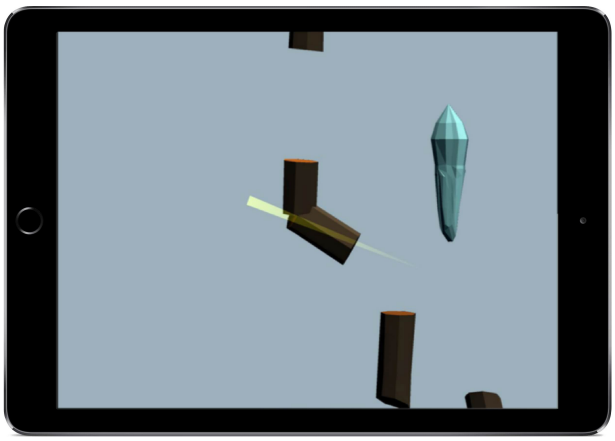

**FIGURE 3 |** The eFun inhibition game (Log Chop) showing how the logs (go stimuli) get chopped.

accuracy x % no-go accuracy; Howard and Melhuish, 2017). This score reflects the participants' ability to withhold their response to the dominant pre-potent response.

### eFun Ice Cube Sorting Task (Cognitive Flexibility)

The Ice Cube Sorting task is based on adapted card sorting tasks (Cianchetti et al., 2007; Kado et al., 2012) and aims to measure cognitive flexibility. The task starts with a brief introductory story explaining that Eski the husky wants to store food for the upcoming winter, since the husky might not be able to leave the house during the cold winter. After a brief practice phase, the child is asked to sort ice cubes containing fruit according to three sorting rules (color, shape, and number) into four tubes displaying four different fruit in four different colors ranging in quantity from one to four (e.g., one red apple, two green pears etc.), see **Figure 4**.

Cognitive flexibility is required because the switching demands in this task are high, while the working memory load and inhibition demand are kept at a constant level. The difficulty of the task increases as the task progresses by introducing more frequent rule switches. This frequency increases, from a new rule after every six correct sorts in level one, to a new rule after every three correct sorts in level two. In order to assist young participants, the rule changes are announced by the explanatory character and children are told that there are three different rules (color, shape, and number) in the beginning of the task. However, the participant does not know which rule needs to be applied next (color, shape or number). In order to find out which rule applies next, the participant needs to remember the previous rule (which cannot be applied again) and then test the two remaining rules. Based on feedback ("that's right" or "that's wrong") the participant can find the correct sorting rule. Cognitive flexibility is measured with a perseveration error rate, which is the number of perseverative errors (PEs) divided by the number of rule changes. PEs occur if the child does not switch to a new rule after the rule switch has been announced, but instead continues to use the previous rule. A higher number of PEs reflects higher cognitive inflexibility (O'Donnell et al., 2017).

### Feedback Questionnaire

After each task, the children were asked to fill out a brief questionnaire with seven questions assessing how enjoyable, fun, exciting, easy, hard, boring and frustrating they found the tasks. The questionnaire is based on the Intrinsic Motivation Inventory (IMI; Deci and Ryan, 2005), which is a multidimensional measurement device intended to assess participants' subjective experience on a target activity. The IMI has been used in the context of intrinsic motivation and self-regulation assessment and includes questions assessing interest and enjoyment. For the purpose of this study with children, questions from the interest/enjoyment scale were adapted and a 4-point response scale was used: "no, not at all", "a little bit", "quite a bit" to "yes, a lot". A very similar type of response scale has previously been shown to be clear and useful for studies involving young children (Rogers et al., 2016). For the EYT tasks, a pen and paper version of the questionnaire was filled out by the participants. The questions and answers were read out to the class to ensure that everyone understood them. For the eFun tasks, the questionnaire was integrated into the apps, therefore responses to the questions were given by tapping on a box underneath the question on an iPad screen. To accommodate non-readers, the questions were verbalized by the explanatory character (owl) and the questions and answers were verbally repeated if the child clicked on them.

### School Grades

Grades ranging from "A" to "D" (for English, Math, HASS, Science, Design and Technology, and Digital Technology) were collected after the testing phase, during the mid-year break.

## Procedure

Before the study commenced, approval from both the Edith Cowan University Ethics Committee and the participating school was sought. The information and consent forms were sent to the school principal, the teachers and parents. The teachers distributed information and consent forms for the children to parents. The information letters that were given out outlined the procedure, possible risks, and purpose of the study. Additionally, the letters informed parents and teachers about a focus group session, in which children could express their opinion about the EF tasks that took place at the end of the study. In consultation with the teachers, appropriate times and dates for the data collection were determined. Before the first study commenced, the researchers were introduced to the students and teachers to familiarize the students with the people assisting the project. Unlike most existing tasks, all tasks were applied in an environment that is familiar to the child (the classroom). All students were tested at the same time in their classroom. Participants were asked to wear their own headphones that were stored at school. The first author (VB) was present throughout all testing sessions to ensure that the children were able to complete the tasks on their own. All instructions were presented verbally through the headphones in addition to being displayed in writing on the screen. Every participant went through two

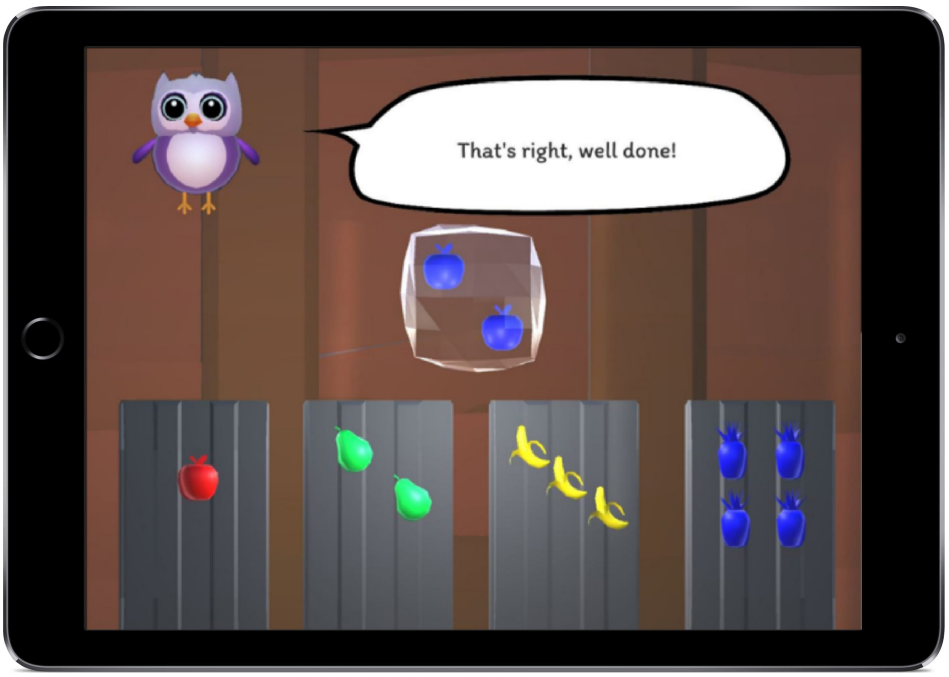


**FIGURE 4 |** The eFun cognitive flexibility game showing the ice cube that needs to be sorted into one of the tubes. If the correct rule was "color" the cube would need to be sorted into the very right tube, for the rule "shape" it would need to be sorted into the very left tube and for the rule "number" it would need to be sorted into the second tube from the left.

testing sessions: one to assess EFs with the EYT; and one to assess EFs with eFun. Both testing sessions were applied on the same day with a minimum break of two hours inbetween (based on the guidelines by Howard and Melhuish, 2017; and Straker et al., 2010). The individual testing sessions took no more than 25 min each (<50 min in total). The collected data was sent to a secure online database, ensuring confidentiality. The physical development guidelines for digital devices use by Straker et al. (2010) was taken into account when conducting this study.

## Research Design and Analysis

A non-experimental correlational design was chosen for this study. This means that no manipulation or selection into groups took place, but all variables of interest were assessed and analyzed. Analyses in SPSS were conducted to examine relationships between working memory, inhibition and cognitive flexibility within each test battery, as well as relationships between the eFun and EYT test batteries to assess convergent validity. Relationships between the variables were assessed using Pearson's correlations. Moderate positive associations between matching tests across the batteries were anticipated (i.e., $Inhibition_{EYT}$ and $Inhibition_{EFun}$, $CF_{EYT}$ and $CF_{EFun}$, $WM_{EYT}$ and $WM_{EFun}$). Repeated measures analysis of variance (ANOVA) was conducted to examine if overall enjoyment of the tasks differed among the tasks.

Furthermore, after the initial measuring phase had ended, a brief focus group with the participants and the teacher took place to yield feedback on the tasks. The feedback informs the design of the eFun tasks for future research. The focus group was not conducted to assess the participants' level of engagement with eFun or to explore the relationship between EFs and the measures within each task, but instead to gain insight into children's opinions on the app design and to debrief children by addressing any issues they might have experienced during or after the testing phase.

# RESULTS

## EF Task Ratings

Answers to the EF questionnaires are presented in **Figure 5**. Overall, students experienced all EF apps as fun and enjoyable with, on average, over 50% of students reporting the tasks to be fun and enjoyable.

With regards to the eFun task ratings, the majority of students reported that they enjoyed the tasks and that they experienced them as fun (**Figure 5**). The eFun working memory and eFun inhibition tasks were reported to be exciting "a lot" by over 55% of students, whereas the eFun cognitive flexibility task was experienced as slightly less exciting (47% "a lot") than the other two tasks. At the same time, over 40% of students reported that the eFun tasks were too easy "a lot" (45–57%). The inhibition task was perceived as the easiest eFun task, while also rated as the most enjoyable (84% "a lot"), the most fun (85% "a lot"), and the most exciting task (70% "a lot") of all EF tasks in this study.

With regards to the EYT task ratings, over 60% of students reported the EYT WM and EYT inhibition task to be enjoyable and fun "a lot". The EYT cognitive flexibility task received slightly lower enjoyment, fun, and excitement ratings than the other EF

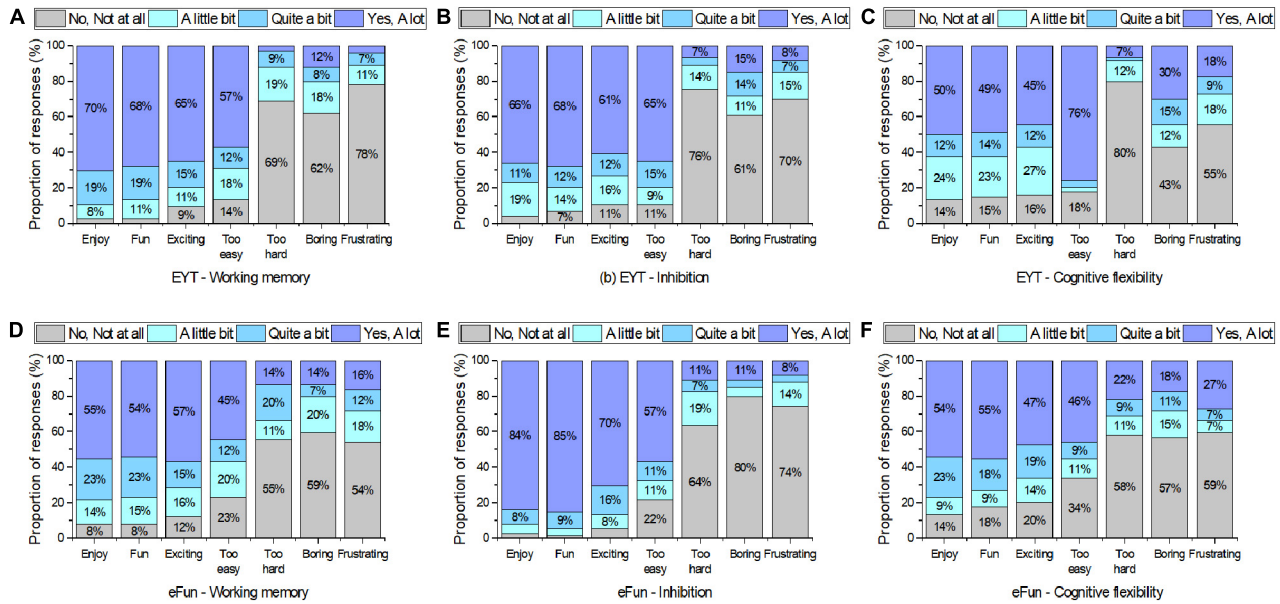


**FIGURE 5 |** Student experience ratings for **(A)** EYT – Working memory, **(B)** EYT – Inhibition, **(C)** EYT – Cognitive flexibility, **(D)** eFun – Working memory, **(E)** eFun – Inhibition, and **(F)** eFun – Cognitive flexibility.

tasks. This task was perceived as the least exciting task (45% "a lot") and the easiest task (76% "a lot"). Overall, the EYT tasks were perceived as too easy by the majority of children (57–76% "a lot").

### EF Task Ratings: ANOVA

In order to compare the enjoyment for each task, the ratings for the adjectives enjoy, fun, exciting, and boring (reverse scored) were combined into an enjoyment score. Inter-correlations among the ratings for the adjectives for each task were consistently moderate to high across all tasks. For example, the average inter-correlation among the eFun CF task adjectives (fun, exciting, boring, enjoy) was 0.71. Therefore, for each task a composite overall enjoyment score was created.

A one-way repeated measures analysis of variance (ANOVA) was applied to compare the student's ratings on the EF tasks from the two test batteries eFun and EYT. The ANOVA results show that the participants enjoyed playing some tasks over others $F(5,365) = 13.32$, $p < 0.001$, $n_p^2 = 0.15$). The follow up pairwise comparison showed that the eFun inhibition task ($M = 3.64$, $SD = 0.55$) was significantly more enjoyed than the EYT inhibition task ($M = 3.30$, $SD = 0.86$), $t(73) = 3.13$, $p = 0.003$, $d = 0.047$. There was no significant difference ($p$-values were greater than 0.05) between the enjoyment rating of the EYT WM task ($M = 3.43$, $SD = 0.79$) and the eFun WM task ($M = 3.23$, $SD = 0.84$) and no significant difference was found between the EYT cognitive flexibility task ($M = 2.87$, $SD = 1.10$) and eFun cognitive flexibility task ($M = 3.10$, $SD = 1.02$).

## Distributions of Performance Scores Among the EF Tasks

The distribution of students' task performance on each task is shown in **Figure 6**. The frequency histograms show that there were distributional issues with all tasks except the EYT WM. The EYT WM shows a normal distribution whereas the other tasks are skewed toward the higher or lower end of the scoring range. Looking over the distributions (see **Figure 6**) it is apparent that overall some tasks were too easy (EYT inhibition, EYT cognitive flexibility, eFun inhibition) while others were too difficult (eFun working memory, eFun cognitive flexibility). Note that higher scores (perseveration error rate) on the eFun cognitive flexibility task indicate worse performance.

## Correlations Between Grades and EF Tasks

Overall, there was one task, the EYT WM, that showed consistent moderate significant correlations with grades. Scores on the EYT WM task were significantly correlated with all subjects (significant correlations ranging between 0.32 and 0.50), see **Table 1**. In addition, scores on the eFun WM task were significantly related to Math grades. However, this correlation ($r = 0.25$, $p < 0.05$) is lower than the correlation between the EYT WM task and grades ($r = 0.34$, $p < 0.05$).

### Correlations Within and Between EYT and eFun

Contrary to expectations, when looking at correlations among working memory, inhibition, and cognitive flexibility *within* each of the apps no significant correlations were observed. There was a single exception where the EYT cognitive flexibility task and EYT inhibition task showed a low correlation of $r$ (69) = 0.24, $p = 0.049$.

With regards to convergent validity *between* the two task batteries, the only significant correlation was found between the eFun and the EYT cognitive flexibility tasks, $r$ (69) = -0.24, $p = 0.041$. The negative correlation can be explained by

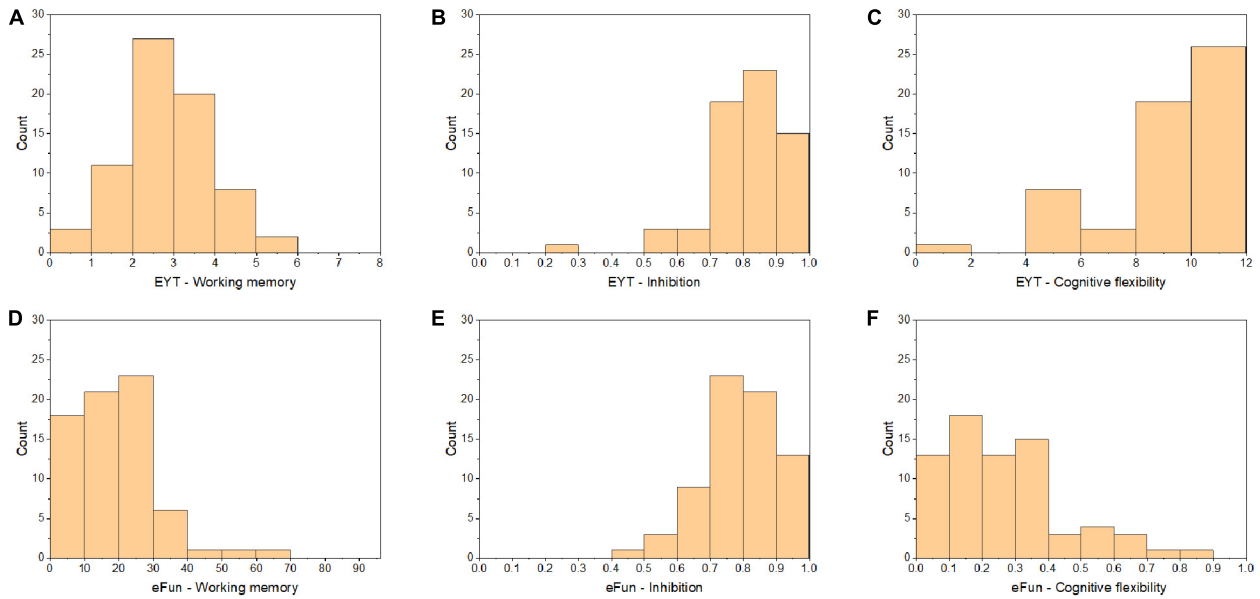


**FIGURE 6 |** Frequency histograms of student performance on **(A)** EYT – Working memory, **(B)** EYT – Inhibition, **(C)** EYT – Cognitive flexibility, **(D)** eFun Working memory, **(E)** eFun Inhibition, and **(F)** eFun Cognitive flexibility.

the different scoring procedures for the tasks. Higher scores (perseveration error rate) on the eFun cognitive flexibility task (Ice Cube Sorting task) indicate worse performance whereas higher scores (number of correct sorts after the preswitch phase) on the EYT cognitive flexibility task (Card Sorting task) indicate better performance.

## DISCUSSION

This study evaluated two newly developed executive function (EF) assessment tools (i.e., EYT and eFun) on an iPad for children in a classroom environment. Building on the knowledge of previous research on child-friendly EF measurement tools (Howard and Okely, 2015; Howard and Melhuish, 2017), this study extends the literature on executive functions assessment in children by targeting (a) children's opinion on the task (b) a new child-friendly design, and (c) a new assessment environment. Results showed that children enjoyed playing both the EYT and eFun tasks in a game-like fashion on an iPad in the classroom. Children's evaluation of the assessment tool has rarely been taken into account when measuring EFs in children in previous research.

**TABLE 1 |** Correlations Between Grades and EF Tasks.

| | English | Math | HASS | Science | Design and tech. | Digital tech. |
|---|---|---|---|---|---|---|
| eFun WM | 0.20 | 0.25* | 0.20 | 0.18 | 0.23 | 0.16 |
| eFun Inhib. | 0.07 | 0.05 | 0.02 | 0.07 | 0.06 | 0.05 |
| eFun CF | −0.02 | −0.04 | −0.12 | −0.14 | −0.08 | 0.01 |
| EYT WM | 0.38* | 0.34* | 0.50* | 0.35* | 0.47* | 0.32* |
| EYT Inhib. | 0.06 | 0.03 | 0.05 | 0.08 | −0.07 | −0.07 |
| EYT CF | 0.06 | 0.02 | 0.18 | 0.18 | 0.13 | 0.12 |

*EYT = Early Years Toolbox, *p < 0.05.*

This article presents an early stage in the validation argument of the eFun and EYT test results. The validation arguments are based on the "Standards" for educational and psychological testing (American Educational Research Association et al., 2014). The "Standards" provide guidelines for assessing the validity of interpretations of test scores. According to the "Standards", one source of validity evidence is evidence based on the test content. The task content for this study was developed in accordance with EF theory and existing EF testing formats. Furthermore, this study tested the task content for appropriateness for the given age group of the sample by asking students for their opinions on the tasks (student questionnaires and focus groups). We were also able to determine from the task performance results that some tasks were too easy and others too hard (ceiling effects). This feedback will be used to further improve the tasks in future iterations. It also became apparent that some children with disabilities such as vision or hearing impairments were not able to undertake the tasks. Thus future research should investigate task modifications to include children with disabilities in the assessment process.

Another source of evidence is based on the response process. This is an aspect of validity that needs to be further investigated in future research. Cognitive processes that are activated during EF testing with EYT and eFun could for example be assessed by neuroimaging (Collette et al., 2006; Cristofori et al., 2019) or "talk aloud protocols". The latter method qualitatively explores how the children are conducting the tasks by recording their thought processes while they are playing the tasks. This approach

would support the assessment of strategy use and help to better understand the thought process of children while playing the tasks (Schrier, 2017; Bratiotis et al., 2019).

Evidence for validity can also be assessed with the analysis of the internal structure of a test. For this study interrelationships between variables were assessed and are in line with theory stating that the core EF constructs are distinct (Miyake et al., 2000). Future research might reassess these relationships with a larger sample and/or within a more diverse target group. Relationships with other external variables also offer a source of evidence for validity. In this study, EF performance was expected to positively associate with academic outcomes, therefore we analyzed correlations between EF scores and grades. We plan on conducting future research to retest these relationship with larger and more diverse samples. Convergent evidence was assessed by investigating the relationship between EFs measured by two different types of assessment (i.e., EYT and eFun). As these types of assessments undergo further refinement, we plan on retesting convergence between these tasks and similar tasks. Lastly, consequences of testing can support the validity argument. A consequence of this study will include teachers' access to the EF test results of their students, which can support the development of targeted learning plans for students. Future research is needed to investigate how well these tasks can be utilized by teachers to better understand and educate their students.

The EF eFun tasks developed for this research aim to measure students' full potential and it has been shown that students that show enjoyment and interest for performance tasks score higher on the performance tasks (Schukajlow and Krug, 2014). Thus, enjoyment with the task should not be neglected when measuring children's task performance. Exploring the relationship between enjoyment and (the related concept of) task-motivation is an important research area for future EF research, however, this is outside the scope of the current study. Executive function tasks and feedback questionnaires could for example be complemented with a task motivation questionnaire (please see Eklöf (2006) for a detailed explanation). Furthermore, the results of this study show that children enjoyed playing the tasks, which can be seen as a positive consequence of the task (Kane, 2013), however, the consequences of the results for students and teachers carry more weight in the validity argument than the enjoyment.

Furthermore, we acknowledge that while established original EF tasks have provided a foundation for the newer tasks that were used in the present study, the tasks in this study differ to the extent that the validity of these tasks cannot be argued based on this basic similarity. The similarity lies in the underlying cognitive mechanism such as remembering something in reversed order (to measure WM) but task response methods and design have been adapted, e.g., from tapping physical blocks to tapping ice floats on an iPad. It will take multiple studies to accumulate an evidence base to provide enough evidence to have full confidence in the more "gamified" tasks that are emerging in the literature (such as EYT and eFun) as alternatives to the more traditional tasks.

## Executive Function Task Ratings

The majority of children enjoyed playing the EYT tasks in addition to over half the children reporting that the EYT tasks are too easy. A possible explanation for the large number of students reporting that the EYT tasks were too easy is that these tasks were designed for children aged 3—6, whereas the age of the students in the current sample was marginally above this age (5–8 years). The EYT cognitive flexibility task (Card Sorting Task) was rated as the easiest and the least exciting and fun task out of all EF tasks in this study. In a focus group that was conducted after the study to inform the design of EF tasks in the future, students explained that the EYT cognitive flexibility task instructions felt too repetitive. The verbal instructions ("now we play the color/shape game") were repeated before every trial. Thus, a future consideration when designing EF tasks for children is to give visual feedback/instructions instead of repetitive verbal instructions.

The majority of children also reported that they enjoyed the eFun tasks and thought they were fun. The highest enjoyment was reported for the eFun inhibition task (Log Chop). It is important to also note that of the eFun tasks collectively, the inhibition task was perceived as the easiest. Yet, at the same time, it was rated as the most enjoyable, fun and exciting task out of all EF tasks applied in this study. This suggests that in the current study children perceive less challenging tasks as more enjoyable. In contrast, literature on challenge and enjoyment typically reports that participants find challenging tasks more enjoyable (e.g., Shernoff et al., 2014). However, when separating voluntary and non-voluntary tasks, research suggests that non-voluntary tasks are enjoyed most when they are of low challenge (Koestner et al., 1987). This further demonstrates one of the challenges of designing enjoyable, yet valid measurement tools for children. If the tasks are too easy, outcome scores result in ceiling effects, despite children enjoying them more. Thus, future research needs to find an appropriate level of challenge that enhances rather than inhibits enjoyment in order to create valid and child-friendly measurement tools.

The eFun inhibition task was enjoyed significantly more than the EYT inhibition task. A possible explanation for this finding is that the design of the eFun task is different to the EYT task in terms of the level and nature of stimuli and interaction. In the eFun task, stimuli (logs and icicles) are moving vertically from the top to the bottom of the screen in varying stimuli locations (i.e., left, middle, and right), whereas the EYT has stimuli (fish and shark) moving horizontally from the left to the right only in the middle of the screen. The response mechanics also differ in that the eFun task requires swiping of the stimuli and the EYT requires tapping the screen. The screen in the EYT task can be tapped anywhere to indicate a response, whereas the eFun task requires the children to "chop" the logs with their fingers. Thus, response locations cannot be anticipated in the eFun task, increasing the difficulty and dynamic nature of the task. Furthermore, the speed of the presentation of stimuli in the eFun increases as the levels get higher, which makes it a highly dynamic game.

## Executive Functions and Grades

As expected, the results showed a link between working memory (measured with the EYT app) and academic outcomes in primary school children. This is in line with prior research showing that working memory is related to school achievement

(van der Sluis et al., 2007; Bull et al., 2008; Sesma et al., 2009; Latzman et al., 2010; Usai et al., 2018). In particular, math (in grade 1-3; Monette et al., 2011; Usai et al., 2018) and English (in 11 and 12 year olds; St Clair-Thompson and Gathercole, 2006) have previously been found to be linked to working memory in children. A recent study investigated associations between the core EFs measured at kindergarten entry and its long term effects on academic achievement in grade 3 (Nguyen and Duncan, 2019). Similar to the results of the current study, Nguyen and Duncan (2019) found that working memory had the strongest associations with math and reading achievement, with math showing the strongest association, whereas inhibition and cognitive flexibility were found to have weaker links with reading and math achievement. Other research supports these findings (Greenfader, 2019), underlying the importance of working memory for academic achievement.

However, it should be noted that the EYT working memory task might measure a slightly different construct that is closer to short-term memory than to working memory. Items need to simply be rehearsed only and not repeated in reverse order in the EYT task; rehearsing items without manipulation has been argued to measure short-term memory rather than working memory (Baddeley, 2012). Higher variability and a normal distribution of the outcome scores for the EYT memory task might explain why this task was found to be related to school grades.

The other EF constructs, inhibition and cognitive flexibility, did not show a relationship with grades in the current study. There are two possible explanations for why no other significant links between EFs and academic achievement were found. Firstly, the participants enjoyed playing the tasks and scored relatively well on the tasks, which is reflected in the data, with relatively low variability across task performances (see **Figure 6** in results section). A lack of variability in the outcome data makes it more difficult to differentiate high performers from low performers, which in turn, limits the potential to find significant results. This is a general measurement issue in research, especially with children (Jacobs and Paris, 1987; Reynolds and Mason, 2009). One potential reason for the lack of variability in the EF tasks is because the EF tasks were based on original EF tasks that were designed to identify executive dysfunction in clinical populations. For example, the Wisconsin Card Sorting Test (WCST) was initially used to identify people with various types of brain dysfunction (e.g., Milner and Petrides, 1984) and to assess neuropsychological dysfunction in school-aged children with developmental psychopathologies (Pennington and Ozonoff, 1996). Thus, in order to assess more subtle levels of varying EF abilities in a typically developing population, we argue that the difficulty levels of the tasks used in the present study could benefit from revision.

A second possible explanation for the low correlations between EF tasks and grades is the age group of the children in this study. In primary school children executive functions are still developing and therefore are more difficult to assess at this younger age. For example, inhibition emerges and undergoes rapid growth in early childhood, particularly between the ages of three to six (Band et al., 2000; Carver et al., 2001; Garon et al., 2008; Wiebe et al., 2012), and continues to mature into early adulthood. Considering the age group of the current sample (5–8 years), inhibitory skills have started to emerge but might not be fully developed yet. Furthermore, children have lower attention spans and therefore are more likely to disengage with the task, which can result in data that is less reliable than data collected with older participants. This is reflected in the research literature, which often reports positive relationships between EF task performance and academic achievement in *older* children (11–16 years; 9–15 years; Sesma et al., 2009; Latzman et al., 2010). This highlights the need for more research with appropriate tools on EFs in relation to academic achievement at the beginning of primary school.

## Correlations Between EYT and eFun

Convergent validity between the two task batteries was found to be low, as low or non-significant correlations were found between matching EF tasks across eFun and EYT. The only significant correlation was between the cognitive flexibility tasks, and it must be noted this was only a small correlation. One reason for the non-significant correlations for working memory and inhibition between the two task batteries could be that the EYT is designed for a younger age group. The EYT tasks are designed for children aged 3–6 (Howard and Melhuish, 2017), and eFun is designed for children aged 5–9 (Berg et al., 2019). Additionally, the eFun test battery has higher difficulty levels than EYT. Another important difference between the tasks is the design. As discussed earlier, the response methods and mechanics differ in the inhibition tasks. Additionally, the constructs that were measured might have differed, especially in the working memory tasks. The eFun WM task exposes a more complex memory construct, since it asks participants to keep stimuli in mind and indicate them in reversed order. On the other hand, the EYT measures a more simple (or short term) memory construct by asking the participants to simply repeat the order of stimuli without having to reverse it. Similarly, the eFun inhibition game is more complex because the stimuli locations vary and the speed increases, whereas the EYT inhibition game does not change stimuli location or the pace by which the stimuli are appearing. In contrast, the cognitive flexibility tasks are both measuring a very complex construct, which might explain the statistically significant low correlation found between these two tasks.

## Task Inter-Correlations

In the current study, there were small or non-significant correlations among EF tasks which is consistent with EF theory by Miyake et al. (2000) that suggests the three core EFs (working memory, inhibition, and cognitive flexibility) are separable constructs. Results showed no significant associations between the EF constructs, with the exception of EYT Inhibition and the EYT cognitive flexibility task, which showed a low significant correlation. Therefore, based on our data, the EF constructs WM and inhibition are distinguishable in the current sample of primary students.

Considering the supporting theories that postulate that executive functions consist of three core constructs (e.g.,

Miyake et al., 2000; Diamond, 2013), using three tasks to measure executive functions is deemed appropriate in this study. However, the findings of the current study do not support research that shows that the three core executive functions are moderately correlated (Miyake et al., 2000; Lehto et al., 2003). The findings need to be interpreted with respect to the population that the theories are based on. Miyake et al. (2000) conducted their research with undergraduate students and Lehto et al. (2003) investigated EFs in typically developing teenagers aged 15 and 16 years. Future research is needed to better understand at what age executive functioning becomes developed to the extent that there are consistent inter-correlations among assessment tasks.

## Limitations

The main limitation of this study was the sample. First, the number of participating children was relatively low. Second, only students in year one and two from one private primary school were included. Additionally, students with severe sensory impairments, such as vision or hearing impairments were not able to take part in the study. Thus the results may not be representative of the typical range of primary students in Australia. The nature of the sample may have also acted to reduce the variability of the data in the present study. Therefore, future research is needed with a larger sample that is more diverse.

An issue in the present study was ceiling effects within our EF performance data. Ceiling effects are a consistent issue within the literature where scholars aim to measure EFs in children, especially when measuring inhibition (Holmboe et al., 2019; Willoughby et al., 2019). For example, Willoughby et al. (2019) found floor and ceiling effects on several inhibition tasks with children (Go/No-Go, Silly Sound Stroop tasks, and Spatial Conflict Arrows). The authors also found low correlations among the executive function tasks in their study and mention that limited task variation in the Go/No-Go task was a problem (for more information please see Willoughby et al., 2019). Similarly, Petersen et al. (2016) and Holmboe et al. (2019) mention variability issues with regards to inhibition tasks. Petersen et al. (2016) explain that ceiling and floor effects are associated with lower variability in the measured construct, which increases Type II error and reduces power to detect associations with other variables. Another reason for floor and ceiling effects can be a small number of trials. Generally, tasks for children are kept short to suit their limited attention span, however, having a small number of trials increases the risk of low variability and ceiling effects in children (Holmboe et al., 2019). Thus, researchers interested in measuring inhibition with children need to be cautious with their selection of trials and outcome variables in order to avoid low variability associated with floor and ceiling effects.

## Implications

The new EF tablet-based measurement tools have useful implications for teachers. The EF apps are an enjoyable activity that can easily be implemented in the classroom schedule, as they do not require any assistance. The outcome data can be uploaded to an online cloud storage, which has potential to give teachers quick access to data about their students' skills. Knowing about students' EF levels can facilitate the development of targeted learning plans. Thus, the purpose of the online EF assessment tool eFun is not to diagnose students but to help teachers to better understand their students' learning profile. For example, during the data collection period of our research one student scored very high on the WM task and passing that information onto the teacher helped the teacher gain a greater understanding of the student's particular cognitive profile. Additionally, pointing out the student's strengths to the student has the potential to increase that student's confidence levels.

Nevertheless, the eFun EF apps are at an early stage of development and still need to be refined to fit primary students' skill levels and to increase variability of the outcome scores. More specifically, this means that the tasks need to be challenging, but enjoyable (with simple and clearly communicated rules) for primary school children. We are planning to adjust the eFun app accordingly and to extend the app by the inclusion of an additional game measuring problem-solving. The problem-solving task is hypothesized to involve all three core executive functions (Zelazo et al., 1997; Diamond, 2013).

# CONCLUSION

The EF tasks that are typically used in the research literature are based on neuropsychological tests that were originally used to diagnose executive dysfunction. The EF tasks used in the present study were modified to be appropriate for use with typically developing children. We examined a newly developed set of EF tasks (eFun; Berg et al., 2019). As expected, children self-reported experiencing all the tasks as fun and enjoyable. Nevertheless, it is an ongoing process to redesign the original tasks to suit participants with a more typical cognitive development profile. This study has contributed to this movement in that we have identified areas that need improvement and have been able to measure children's perspectives on EF tasks for the first time. The challenge of future research is to address the issues identified by refining EF measurement tasks that can maintain valid results, while being easy to use, particularly for primary school-aged children for whom the development of EFs can impact greatly on the early stages of their learning journey.

# DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

# ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Edith Cowan Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

VB is the lead author who conducted the research, developed the measurement tool, and wrote and edited the manuscript. SR edited the manuscript and contributed to the research process and measurement tool development. MM and MG edited the manuscript and contributed to the measurement tool development. DM is the lead developer of the measurement tool and helped with the acquisition of data. All authors contributed to the article and approved the submitted version.

## FUNDING

Funding was received from Edith Cowan University and Cinglevue International as part of an Industry Ph.D. scholarship.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01702/full#supplementary-material

## REFERENCES

American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Aron, A. R. (2011). From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. *Biol. Psychiatry* 69, e55–e68. doi: 10.1016/j.biopsych.2010.07.024

Axelsson, A., Andersson, R., and Gulz, A. (2016). Scaffolding executive function capabilities via play-&-learn software for preschoolers. *J. Educ. Psychol.* 108, 969–981. doi: 10.1037/edu0000099

Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

Band, G. P., van der Molen, M. W., Overtoom, C. C., and Verbaten, M. N. (2000). The ability to activate and inhibit speeded responses: separate developmental trends. *J. Exp. Child Psychol.* 75, 263–290. doi: 10.1006/jecp.1999.2538

Berg, V., McMahon, M., Rogers, S., Garrett, M., and Manley, D. (2019). The eFun App: a new tool to measure executive functions to support learning in a child-friendly, valid and engaging Way. *Paper Presented at the EdMedia+ Innovate Learning*, Waynesville.

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., and Domitrovich, C. E. (2008). Executive functions and school readiness intervention: impact, moderation, and mediation in the Head Start REDI program. *Dev. Psychopathol.* 20, 821–843. doi: 10.1017/S0954579408000394

Bratiotis, C., Steketee, G., Dohn, J., Calderon, C. A., Frost, R. O., and Tolin, D. F. (2019). Should i keep it? Thoughts verbalized during a discarding task. *Cogn. Ther. Res.* 43, 1075–1085. doi: 10.1007/s10608-019-10025-y

Broidy, L. M., Tremblay, R. E., Brame, B., Fergusson, D., Horwood, J. L., Laird, R., et al. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: a six-site, cross-national study. *Dev. Psychol.* 39, 222–245. doi: 10.1037/0012-1649.39.2.222

Bull, R., Espy, K. A., and Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev. Neuropsychol.* 33, 205–228. doi: 10.1080/87565640801982312

Bull, R., and Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Dev. Neuropsychol.* 19, 273–293. doi: 10.1207/s15326942dn1903_3

Carver, A., Livesey, D., and Charles, M. (2001). Age related changes in inhibitory control as measured by stop signal task performance. *Int. J. Neurosci.* 107, 43–61. doi: 10.3109/00207450109149756

Cianchetti, C., Corona, S., Foscoliano, M., Contu, D., and Sannio-Fancello, G. (2007). Modified Wisconsin Card Sorting Test (MCST, MWCST): normative data in children 4–13 years old, according to classical and new types of scoring. *Clin. Neuropsychol.* 21, 456–478. doi: 10.1080/13854040600629766

Collette, F., Hogge, M., Salmon, E., and Van der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience* 139, 209–221. doi: 10.1016/j.neuroscience.2005.05.035

Cristofori, I., Cohen-Zimerman, S., and Grafman, J. (2019). Executive functions. *Handb. Clin. Neurol.* 163, 197–219.

Dawson, P., and Guare, R. (2018). *Executive Skills in Children and Adolescents: A Practical Guide to Assessment and Intervention*. New York, NY: Guilford Publications.

De Greeff, J., Hartman, E., Mullender-Wijnsma, M., Bosker, R., Doolaard, S., and Visscher, C. (2016). Long-term effects of physically active academic lessons on physical fitness and executive functions in primary school children. *Health Educ. Res.* 31, 185–194. doi: 10.1093/her/cyv102

Deci, E., and Ryan, R. (2005). *Intrinsic Motivation Inventory (IMI) [WWW Document]. Self-Determination Theory.* Available online at: http://www.selfdeterminationtheory.org/intrinsic-motivation-inventory (accessed July 6, 2020).

Denson, T. F., Pedersen, W. C., Friese, M., Hahm, A., and Roberts, L. (2011). Understanding impulsive aggression: angry rumination and reduced self-control capacity are mechanisms underlying the provocation-aggression relationship. *Pers. Soc. Psychol. Bull.* 37, 850–862. doi: 10.1177/0146167211401420

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Curr. Dir. Psychol. Sci.* 21, 335–341. doi: 10.1177/0963721412453722

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Diamond, A. (2016). "Why improving and assessing executive functions early in life is critical," in *Executive Function in Preschool-Age Children: Integrating Measurement, Neurodevelopment, and Translational Research*, eds J. A. Griffin, P. McCardle, and L. S. Freund (Washington, DC: American Psychological Association), 11–43. doi: 10.1037/14797-002

Diamond, A., and Wright, A. (2014). An effect of inhibitory load in children while keeping working memory load constant. *Front. Psychol.* 5:213. doi: 10.3389/fpsyg.2014.00213

Dickey, M. D. (2006). Game design narrative for learning: appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. *Educ. Technol. Res. Dev.* 54, 245–263. doi: 10.1007/s11423-006-8806-y

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychol.* 30, 412–431. doi: 10.1016/0001-6918(69)90065-1

Dörrenbächer, S., Müller, P. M., Tröger, J., and Kray, J. (2014). Dissociable effects of game elements on motivation and cognition in a task-switching training in middle childhood. *Front. Psychol.* 5:1275. doi: 10.3389/fpsyg.2014.01275

Eklöf, H. (2006). Development and validation of scores from an instrument measuring student test-taking motivation. *Educ. Psychol. Meas.* 66, 643–656. doi: 10.1177/0013164405278574

Engelmann, J. B., Damaraju, E., Padmala, S., and Pessoa, L. (2009). Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects. *Front. Hum. Neurosci.* 3:4. doi: 10.3389/neuro.09.004.2009

Engelmann, J. B., and Pessoa, L. (2014). Motivation sharpens exogenous spatial attention. *Emotion* 7, 668–674. doi: 10.1037/1528-3542.7.3.668

Falloon, G. (2013). Young students using iPads: app design and content influences on their learning pathways. *Comput. Educ.* 68, 505–521. doi: 10.1016/j.compedu.2013.06.006

Flook, L., Smalley, S. L., Kitil, M. J., Galla, B. M., Kaiser-Greenland, S., Locke, J., et al. (2010). Effects of mindful awareness practices on executive functions in elementary school children. *J. Appl. Sch. Psychol.* 26, 70–95. doi: 10.1080/15377900903379125

Garon, N., Bryson, S. E., and Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychol. Bull.* 134, 31–60. doi: 10.1037/0033-2909.134.1.31

Gioia, G., Isquith, P., Guy, S., and Kenworthy, L. (2001). *Brief Rating Inventory of Executive Function (BRIEF)*. Odessa, FL: Psychological Assessment Resources.

Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., and Weinberger, D. R. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Arch. Gen. Psychiatry* 54, 159–165.

Greenfader, C. M. (2019). What is the role of executive function in the school readiness of Latino students? *Early Child. Res. Q.* 49, 93–108. doi: 10.1016/j.ecresq.2019.02.011

Hackley, S. A., Schäffer, R., and Miller, J. (1990). Preparation for Donders' type B and C reaction tasks. *Acta Psychol.* 74, 15–33. doi: 10.1016/0001-6918(90)90032-B

Holmboe, K., Larkman, C., de Klerk, C., Simpson, A., Christodoulou, C., and Dvergsdal, H. (2019). *The Early Childhood Inhibitory Touchscreen Task*.

Howard, S. J., and Melhuish, E. (2017). An early years toolbox for assessing early executive function, language, self-regulation, and social development: validity, reliability, and preliminary norms. *J. Psychoeduc. Assess.* 35, 255–275. doi: 10.1177/0734282916633009

Howard, S. J., and Okely, A. D. (2015). Catching fish and avoiding sharks: investigating factors that influence developmentally appropriate measurement of preschoolers' inhibitory control. *J. Psychoeduc. Assess.* 33, 585–596. doi: 10.1177/0734282914562933

Hunt, A. R., and Kingstone, A. (2003). Covert and overt voluntary attention: Linked or independent? *Cogn. Brain Res.* 18, 102–105. doi: 10.1016/j.cogbrainres.2003.08.006

Jacobs, J. E., and Paris, S. G. (1987). Children's metacognition about reading: issues in definition, measurement, and instruction. *Educ. Psychol.* 22, 255–278. doi: 10.1207/s15326985ep2203%264_4

Józsa, K., Barrett, K. C., and Morgan, G. A. (2017). Game-like tablet assessment of approaches to learning: assessing mastery motivation and executive functions. *Electron. J. Res. Educ. Psychol.* 15, 665–695. doi: 10.14204/ejrep.43.17026

Kado, Y., Sanada, S., Yanagihara, M., Ogino, T., Ohno, S., Watanabe, K., et al. (2012). Executive function in children with pervasive developmental disorder and attention-deficit/hyperactivity disorder assessed by the Keio version of the Wisconsin card sorting test. *Brain Dev.* 34, 354–359. doi: 10.1016/j.braindev.2011.08.008

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *J. Educ. Meas.* 50, 1–73. doi: 10.1111/jedm.12000

Kessels, R. P. C., van den Berg, E., Ruis, C., and Brands, A. M. A. (2008). The backward span of the Corsi block-tapping task and its association with the WAIS-III digit span. *Assessment* 15, 426–434. doi: 10.1177/1073191108315611

Koestner, R., Zuckerman, M., and Koestner, J. (1987). Praise, involvement, and intrinsic motivation. *J. Pers. Soc. Psychol.* 53, 383–390. doi: 10.1037/0022-3514.53.2.383

Latzman, R. D., Elkovitch, N., Young, J., and Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *J. Clin. Exp. Neuropsychol.* 32, 455–462. doi: 10.1080/13803390903164363

Lehto, J. E., Juujärvi, P., Kooistra, L., and Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *Br. J. Dev. Psychol.* 21, 59–80. doi: 10.1348/026151003321164627

Lim, T., Louchart, S., Suttie, N., Hauge, J. B., Stanescu, I. A., Ortiz, I. M., et al. (2014). *Narrative Serious Game Mechanics (NSGM) – Insights into the Narrative-Pedagogical Mechanism*. Cham.

Lin, H.-L., Lawrence, F. R., and Gorrell, J. (2003). Kindergarten teachers' views of children's readiness for school. *Early Child. Res. Q.* 18, 225–237. doi: 10.1016/S0885-2006(03)00028-0

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203. doi: 10.1037/0033-2909.109.2.163

Mäntylä, T., Rönnlund, M., and Kliegel, M. (2010). Components of executive functioning in metamemory. *Appl. Neuropsychol.* 17, 289–298. doi: 10.1080/09084282.2010.525090

Memisevic, H., Biscevic, I., and Pasalic, A. (2018). Predictors of math achievement in the elementary school students grades 1-3. *Acta Neuropsychol.* 16, 249–258. doi: 10.5604/01.3001.0012.5915

Milner, B., and Petrides, M. (1984). Behavioural effects of frontal-lobe lesions in man. *Trends Neurosci.* 7, 403–407. doi: 10.1016/s0166-2236(84)80143-5

Milner, B., and Taylor, L. (1972). Right-hemisphere superiority in tactile pattern-recognition after cerebral commissurotomy: evidence for nonverbal memory. *Neuropsychologia* 10, 1–15. doi: 10.1016/0028-3932(72)90038-3

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108

Monette, S., Bigras, M., and Guay, M.-C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *J. Exp. Child Psychol.* 109, 158–173. doi: 10.1016/j.jecp.2011.01.008

Morra, S. (1994). Issues in working memory measurement: testing for M capacity. *Int. J. Behav. Dev.* 17, 143–159. doi: 10.1177/016502549401700109

Nguyen, T., and Duncan, G. J. (2019). Kindergarten components of executive function and third grade achievement: a national study. *Early Child. Res. Q.* 46, 49–61. doi: 10.1016/j.ecresq.2018.05.006

Obradović, J., Sulik, M. J., Finch, J. E., and Tirado-Strayer, N. (2018). Assessing students' executive functions in the classroom: validating a scalable group-based procedure. *J. Appl. Dev. Psychol.* 55, 4–13. doi: 10.1016/j.appdev.2017.03.003

O'Donnell, L. A., Deldin, P. J., Pester, B., McInnis, M. G., Langenecker, S. A., and Ryan, K. A. (2017). Cognitive flexibility: a trait of bipolar disorder that worsens with length of illness. *J. Clin. Exp. Neuropsychol.* 39, 979–987. doi: 10.1080/13803395.2017.1296935

Otto, T., Schottler, F., Staubli, U., Eichenbaum, H., and Lynch, G. (1991). Hippocampus and olfactory discrimination learning: effects of entorhinal cortex lesions on olfactory learning and memory in a successive-cue, go–no-go task. *Behav. Neurosci.* 105, 111–119. doi: 10.1037/0735-7044.105.1.111

Pennington, B. F., and Ozonoff, S. (1996). Executive functions and developmental psychopathology. *J. Child Psychol. Psychiatry* 37, 51–87. doi: 10.1111/j.1469-7610.1996.tb01380.x

Pessoa, L., and Engelmann, J. B. (2010). Embedding reward signals into perception and cognition. *Front. Neurosci.* 4:17. doi: 10.3389/fnins.2010.00017

Petersen, I. T., Hoyniak, C. P., McQuillan, M. E., Bates, J. E., and Staples, A. D. (2016). Measuring the development of inhibitory control: the challenge of heterotypic continuity. *Dev. Rev.* 40, 25–71. doi: 10.1016/j.dr.2016.02.001

Posner, M. I., and Cohen, Y. (1984). "Components of visual orienting," in *Attention and performance X: Control of language processes*, eds H. Bouma, and D. G. Bouwhuis (Hillsdale, NJ: Erlbaum), 531–556

Posner, M. I., and Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annu. Rev. Psychol.* 58, 1–23. doi: 10.1146/annurev.psych.58.110405.085516

Reeve, J. (1989). The interest-enjoyment distinction in intrinsic motivation. *Motiv. Emot.* 13, 83–103. doi: 10.1007/bf00992956

Reynolds, C. R., and Mason, B. A. (2009). Measurement and statistical problems in neuropsychological assessment of children," in *Handbook of Clinical Child Neuropsychology*, eds C. R. Reynolds, and E. Fletcher-Janzen (New York, NY: Springer), 203–230. doi: 10.1007/978-0-387-78867-8_9

Rogers, S. L., Barblett, L., and Robinson, K. (2016). Investigating the impact of NAPLAN on student, parent and teacher emotional distress in independent schools. *Aust. Educ. Res.* 43, 327–343. doi: 10.1007/s13384-016-0203-x

Schrier, K. (2017). Designing role-playing video games for ethical thinking. *Educ. Technol. Res. Dev.* 65, 831–868. doi: 10.1007/s11423-016-9489-7

Schukajlow, S., and Krug, A. (2014). "Are interest and enjoyment important for students' performance," in *Proceedings of the Joint Meeting of PME 38 and PME-NA 36*, Vol. 5,eds C. Nicol, S. Oesterle, P. Liljedahl, and D. Allan (Vancouver: PME).

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., and Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychol.* 15, 232–246. doi: 10.1080/09297040802220029

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., and Shernoff, E. S. (2014). "Student engagement in high school classrooms from the perspective of flow theory," in *Applications of Flow in Human Development and Education,* ed. T. B. Gutkin (Dordrecht: Springer), 475–494. doi: 10.1007/978-94-017-9094-9_24

Shute, V. J., Wang, L., Greiff, S., Zhao, W., and Moore, G. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. *Comput. Hum. Behav.* 63, 106–117. doi: 10.1016/j.chb.2016.05.047

Simpson, A., and Riggs, K. J. (2006). Conditions under which children experience inhibitory difficulty with a "button-press" go/no-go task. *J. Exp. Child Psychol.* 94, 18–26. doi: 10.1016/j.jecp.2005.10.003

St Clair-Thompson, H. L., and Gathercole, S. E. (2006). Executive functions and achievements in school: shifting, updating, inhibition, and working memory. *Q. J. Exp. Psychol.* 59, 745–759. doi: 10.1080/17470210500162854

Straker, L., Maslen, B., Burgess-Limerick, R., Johnson, P., and Dennerlein, J. (2010). Evidence-based guidelines for the wise use of computers by children: physical development guidelines. *Ergonomics* 53, 458–477. doi: 10.1080/00140130903556344

Usai, M. C., Viterbori, P., and Traverso, L. (2018). Preschool executive function profiles: implications for math achievement in grades 1 and 3. *J. Res. Child. Educ.* 32, 404–418. doi: 10.1080/02568543.2018.1496957

van der Sluis, S., de Jong, P. F., and van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence* 35, 427–449. doi: 10.1016/j.intell.2006.09.001

Wallisch, A., Little, L. M., Dean, E., and Dunn, W. (2018). Executive function measures for children: a scoping review of ecological validity. *OTJR* 38, 6–14. doi: 10.1177/1539449217727118

Wiebe, S. A., Sheffield, T. D., and Andrews Espy, K. (2012). Separating the fish from the sharks: a longitudinal study of preschool response inhibition. *Child Dev.* 83, 1245–1261. doi: 10.1111/j.1467-8624.2012.01765.x

Willoughby, M. T., Piper, B., Kwayumba, D., and McCune, M. (2019). Measuring executive function skills in young children in Kenya. *Child Neuropsychol.* 25, 425–444. doi: 10.1080/09297049.2018.1486395

Zelazo, P. D. (2015). Executive function: reflection, iterative reprocessing, complexity, and the developing brain. *Dev. Rev.* 38, 55–68. doi: 10.1016/j.dr.2015.07.001

Zelazo, P. D., Blair, C. B., and Willoughby, M. T. (2016). *Executive Function: Implications for Education. NCER 2017-2000.* Washington, DC: U.S. Department of Education.

Zelazo, P. D., Carter, A., Reznick, J. S., and Frye, D. (1997). Early development of executive function: a problem-solving framework. *Rev. Gen. Psychol.* 1, 198–226. doi: 10.1037/1089-2680.1.2.198

**Conflict of Interest:** VB is studying under an industry-based Ph.D. scholarship, which is financially supported jointly by Edith Cowan University and Cinglevue International. Resources that were utilized in the development of the eFun app were provided by Cinglevue International. Cinglevue International intends to incorporate this into their Virtuoso platform and offer it to customers on a commercial basis.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 30 July 2020
doi: 10.3389/fpsyg.2020.01563

# Promoting Students' Well-Being and Inclusion in Schools Through Digital Technologies: Perceptions of Students, Teachers, and School Leaders in Italy Expressed Through SELFIE Piloting Activities

*Sabrina Panesi*, Stefania Bocconi and Lucia Ferlino*

*Institute for Educational Technology, National Research Council of Italy, Genova, Italy*

OPEN ACCESS

***Edited by:***
*Mirta Vernice,*
*University of Milano-Bicocca, Italy*

***Reviewed by:***
*Jesús-Nicasio García-Sánchez,*
*Universidad de León, Spain*
*Joni Tzuchen Tang,*
*National Taiwan University of Science and Technology, Taiwan*

****Correspondence:***
*Sabrina Panesi*
*panesi@itd.cnr.it;*
*panesi.sabrina@gmail.com*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *29 February 2020*
***Accepted:*** *10 June 2020*
***Published:*** *30 July 2020*

***Citation:***
*Panesi S, Bocconi S and Ferlino L (2020) Promoting Students' Well-Being and Inclusion in Schools Through Digital Technologies: Perceptions of Students, Teachers, and School Leaders in Italy Expressed Through SELFIE Piloting Activities. Front. Psychol. 11:1563. doi: 10.3389/fpsyg.2020.01563*

Digital technology in its various forms is a significant component of our working environment and lifestyles. However, there is a broad difference between using digital technologies in everyday life and employing them in formal education. Digital technologies have largely untapped potential for improving education and fostering students' well-being and inclusion at school. To bring this to fruition, systemic and coordinated actions involving the whole school community are called for. To help schools exploit the full range of opportunities digital technologies offer for learning, the European Commission has designed and implemented a self-reflection tool called SELFIE (Self-reflection on Effective Learning by Fostering Innovation through Educational Technology). Based on the DigCompOrg conceptual framework, SELFIE encompasses key aspects for effectively integrating digital technologies in school policies and practices. The present study investigates how SELFIE can also support the school community to self-reflect about students' well-being and inclusion. In Italy, the SELFIE online questionnaire has been completed by 24,715 students, 5,690 teachers, and 1,507 school leaders, for a total of 31,912 users from 201 schools (at primary, lower secondary, and upper secondary levels) located in 10 different regions. The complementary data we have collected regarding student well-being and inclusion highlight significant differences in the perceptions on this issue reported by students, teachers, and school leaders. These findings have important implications for facilitating successful practices within the whole school community in order to promote students' well-being and inclusion using educational technologies, as well as for planning future actions following a systemic approach.

**Keywords: inclusion, well-being, digital education, digital competence, SELFIE, DigCompOrg**

# INTRODUCTION

## Student Well-Being and Inclusion in Schools

Over the past decade, there has been growing interest in students' well-being, not only in relation to how it may impact on their learning, but also at policy level, examining whether and how education systems that prioritize student well-being foster positive and fulfilling life experience (Pollard and Lee, 2003). The Programme for International Student Assessment (PISA) defines well-being as the quality of students' life, focusing on their psychological, cognitive, social, and physical capabilities. It also distinguishes between various dimensions of well-being, including life as a whole, self-related well-being, school-related well-being, and well-being out of school (OECD, 2019). This definition also emphasizes students' well-being as an active inner process to achieve their personal and social goals (Borgonovi and Pál, 2016). Student well-being covers four discrete aspects that are nonetheless strictly corelated. The first is *cognitive well-being*, namely, successful participation in society in a variety of roles—as lifelong learners, as productive workers, as active citizens—thanks in part to their possessing the knowledge and competences required to fulfill those roles effectively. The second aspect is *psychological well-being*, namely, students' opinion and feelings about their own lives, their educational activity, and the personal objectives they have set themselves. The third is *physical well-being*, in other words, their health level and capacity to lead a healthy lifestyle. The last facet is *social well-being*, covering relations with the family, other learners, and educators, as well as perception of the school social environment. In particular, relations with peers and educators often prove to be very strong indicators of other well-being aspects (Suldo et al., 2009; Moore et al., 2017; Littlecott et al., 2018). Investigations that Govorova et al. (2020) conducted in 35 countries within the OECD examined the range and impact on learner well-being of different factors present in learning settings; specifically, they examined measures educational institutions can take to improve learners' perceived sense of well-being. They found that actually such measures had little substantial impact when viewed within the overall sample. Nevertheless, the authors stress the need for schools to adopt a more holistic approach and to ensure that daily educational activities take adequate account of student well-being, especially regarding the social, psychological, and emotional dimensions of the student experience.

A number of investigations such as that by de Róiste et al. (2012) highlight the considerable gains in social and emotional well-being that can be attained through learners' more active involvement in school activities, with benefits in other areas as well, like improved learning outcomes and experiences, higher satisfaction, as well as stronger relationships and engagement levels (Kuurme and Carlsson, 2010; Løhre et al., 2010; Coombes et al., 2013). Analysis that Anderson and Graham (2016) performed of the perceived well-being levels of learners, teachers, school leaders, and support staff in a large-scale study conducted in Australia found that learners consider well-being as multi-faceted; it may comprise such aspects as student voice and regard for their positions, the exercising of rights and commanding respect. In the study, learners and school staff alike highlighted recognition of their voices and their own selves as valuable, respect-worthy members of the community as areas of particular significance for well-being. Research into the student–teacher relationship in its varied forms has revealed that teachers place special importance on discussing a wide range of matters with their students beyond study-related topics as a means of forging strong bonds and thereby contributing to well-being (Maelan et al., 2018). It should be stressed, however, that only teachers were surveyed in this case. Indeed, in much of the extant literature, "teachers" and "school staff" are treated as being synonymous, despite the burgeoning range of different support roles currently being performed by school staff members, which could well have a beneficial impact on learner well-being (van Petegem et al., 2007).

Students' well-being at school is increasingly conceptualized not only at the individual level but also as a collective, school-wide commitment. Schools fostering the individual and collective well-being of students provide essentials for their holistic growth and development, at school and beyond. To this end, schools play a pivotal role in students' individual and collective well-being, incorporating well-being into planning and processes, striving for excellence in teaching and learning, connecting on many levels, and helping to build trusting and respectful relationships for students to succeed. For example, the draft Curriculum for Wales 2022 (Welsh Government, 2020) seeks to embed health and well-being into the core of the new curriculum by making it one of the six "Areas of Learning Experience" for Welsh schools. It is intended to help learners understand and appreciate how the different components of health and well-being are interconnected and recognizes that good health and well-being are important to enable successful learning.

Another framework for investigating school climate and learner well-being is that formulated by the OECD (2019). This covers four areas of school experience. The first is s*afety*, pertaining to maladaptive behaviors like classroom discipline and student bullying, as well as the school's regulations, stances, and strategies on such behavior. The second area is *teaching and learning*, comprising academic support, response and engagement, civic education and socio-emotional abilities, as well as indications of impact deriving from continuing professional development and school leadership, like school vision and teachers' peer collaboration. The third area is *school community*, which comprises the likes of student–teacher relations, student cooperation and teamwork, respect for diversity, involvement of parents, partnerships with the local community, and the sense of involvement and belonging. The fourth area is the *institutional environment*, covering aspects like the school's facilities and resources, availability of learning resources and digital technology, and measures of school organization like the size of the school and its classes.

Similarly, the Australian Well-being Framework for Schools (NSW Department of Education and Communities, 2015) recognizes the strong linkages between student safety, well-being, and learning outcomes, and identifies five key elements for fostering well-being within the complex, multi-dimensional

world of schools. The first element is *leadership*, given that school heads and leaders are pivotal for developing a positive socio-educational environment in which all members of the school community have a sense of being included and connected, feeling both safe and respected. The second area is *inclusion*, whereby the entire school community is proactive in fostering a welcoming socio-educational culture, one in which diversity and sound, respectful relationships are valued. The third area is Student Voice, meaning that students play an active role in their learning and well-being, develop a sense of feel connectedness, and employ their social and emotional abilities to ensure they act respectfully, resiliently, and safely. The fourth area regards collaborative partnerships with families and local communities as partners as part of students' learning, safety, and well-being. The fifth area is *support* among staff, students, and families, who work together to foster a sense of well-being and to promote positive behavior, ultimately leading to more effective teaching and learning. Hence, the Well-being Framework for Schools provides a broader understanding of well-being that encompasses multi-dimensional aspects of schools. It also conceptualizes inclusion to a degree as a different but interrelated factor contributing to well-being in the school context.

Inclusion is defined by Ainscow (2005) as the constant, ongoing quest to optimize responses to diversity, involving the abolition of obstacles to physical presence, active participation, and attainment. These factors are particularly relevant for any students facing the potential risk of being marginalized or of underachieving. In Ainscow's view, the concept of inclusion also extends to the removal of negative responses and/or attitudes to diversity regarding a person's race, ethnicity, gender, sexual orientation, social class, economic status, religion, first language, achievement levels, not to mention disability (Ainscow et al., 2016; Messiou, 2017). A major defining characteristic of inclusive education is response to student diversity through the deployment of learning environments and learning opportunities for all (Slee, 2018). This ensures that all learners have the potential to be an integral part of the school community and to engage actively in all facets of school life (Spratt and Florian, 2015).

In the Italian education context, Benigno et al. (2018) clearly point out that, in recent definitions of school inclusion, students' universal access to education is combined with the pivotal role schools play in fostering a sense of collective belonging to an amicable network of individuals. Indeed, peer interaction is now seen by many as an important component of inclusion, although terms like *integration*, *participation*, and *social inclusion* are rarely defined in clear explicit terms. That notwithstanding, in much of the literature on the subject, aspects related to these fundamental concepts are considered crucial, including constructs like participation in group activities, incidence of peer interactions, perception of acceptance, and friendly relations. The study recently conducted by Sarti et al. (2019) highlights that the perception of support from others and the possibility to ask help when needed is fundamental to build social inclusive schools' environments, especially for students with specific learning disabilities.

Moreover, children with special needs attend regular classes in Italy, from primary to secondary school, hence following the same curriculum as their peers (Zanobini et al., 2017). Despite this, problems have emerged concerning the actual degree of inclusiveness in Italy's schools (Opertti, 2015). Principally, it has become apparent that mere attendance of students with disabilities in mainstream classes does not in itself mean that the curriculum is universally shared (Obiakor et al., 2012). Some research has demonstrated that a degree of exclusion also occurs in what are ostensibly inclusive classes, in which learners with disabilities show signs of feeling low acceptance in class (D'Alessio, 2011; Zanobini, 2013; Benigno et al., 2019).

In order to guide schools through a process of inclusive school development, the Centre for Studies on Inclusive Education has defined the *Index for Inclusion* (Booth and Ainscow, 2011), an effort intended to foster high achievement levels for all staff and students. The view of inclusion that the Index embodies is to minimize barriers to learning and participation within school policies and practices and to emphasize the student diversity as a rich resource for supporting teaching and learning. The CSIE Index is founded on the social model of disability and includes three dimensions: development of inclusive cultures, formulation of inclusive policies, and evolution in inclusive practices. Each of these is associated to a set of indicators and a number of questions. For instance, *Dimension C—Evolving inclusive practices* comprises a group of indicators that instantiate major aspects of inclusive education: all learners are encouraged to participate actively in class; they are actively engaged in their own learning; students collaborate for learning; teachers seek to foster participation and facilitate learning for all students.

## Supporting Student Well-Being and Inclusion in Schools Through Technology

In considering the matter of students' well-being and inclusion at school, it is fundamental to understand the role and impact technology may have on these two dimensions. Evidence suggests that technologies offer opportunities for inclusive education, helping in particular to prepare learners with specific needs (related to disability, immigrant background, and socio-economic disadvantage) acquire skills that enable them to integrate into education and society as well (e.g., Benigno et al., 2019).

The European Digital Strategy (European Commission, 2020) recently announced by the European Commission highlights digital inclusion among the key priorities for the coming years. The Commission's efforts to ensure that everybody can contribute to and benefit from the digital economy and society reflects an inclusion-driven approach through digital technologies that centers around four main pillars: (i) advancing accessible ICT solutions (design for all), (ii) developing assistive technologies enabling people with disabilities to interact, (iii) empowering citizens' skills and digital skills to fight marginalization and social exclusion, and (iv) fostering social inclusion and participation of disadvantaged people in public, social, and economic activities. The use of digital technology provides individuals with opportunities for accessing information, managing their own learning processes, communicating with peers and mentors, and developing,

repurposing, and sharing materials (Bocconi and Ott, 2013, 2014). To effectively design inclusive learning environments, teachers require specific training activities (Caruso and Ferlino, 2018), as well as to further develop a wide range of digital pedagogical competences so as to promote inclusive and personalized learning (Redecker, 2017; Bocconi and Panesi, 2018b; Caena and Redecker, 2019).

Defining the contribution that digital technologies can make in promoting inclusive socio-educational processes, Trentin (2019) proposes the *hybrid inclusive classroom model*. This entails *always-on* education opportunities for homebound students (e.g., those with chemical sensitivity illnesses), who can thereby actively contribute and take part in daily classroom life from a remote location. Hybrid inclusive classrooms unfold within dynamic hybrid spaces formed when participants use their (mobile or fixed) devices to connect online at any time, thereby integrating remote (and/or virtual) spaces and situations within the actuality of a visual/perceptual location/situation. From the learning viewpoint, hybrid classrooms exploit the *liquid nature* of digital interaction, melting the institutional rigidity that typifies schools and thereby opening up spatial–temporal and conceptual crossflows and currents (Trentin, 2017; Benigno et al., 2018). Critically, this approach also allows homebound students to maintain social relations with their peers, something that plays a central role in the development of the mind and of those social, cognitive, and meta-cognitive abilities that empower the individual to grasp and manage their inner world and well-being. Clearly, there are also significant knock-on benefits for the peers (and teachers) of homebound students, whose participation in hybrid classrooms not only impacts positively on their innate sense of inclusion but also broadens and strengthens their sense of how digital technologies can empower and shape educational processes *per se*, with potential benefits for digitally driven well-being.

Acknowledgment that *digital well-being* is intrinsic to digital competence is also a critical consideration for learning institutions (Gui et al., 2017). Authors such as Beetham (2015) argue that *digital well-being* concerns not only the supposed benefits of digital engagement but also some of its possible risks. In this view, the digital well-being of students differs substantially from that of education staff. This manifests, for example, in learners not realizing that certain online behaviors are illegal, or staff members' stress deriving from digital work and health issues connected to digital activities.

In the Italian context, Bicocca University in Milan ran a project called "Digital Well-being – Schools" whose objective was to help educators acquire the skills needed to cooperate with students in the development of a mindful relationship with digital media and for fostering digital well-being in all areas of day-to-day life. The project deployed a training intervention that was subjected to a randomized controlled trial involving 15- to 16-year-old students from 18 high schools located in northern Italy. The outcomes corroborate the belief that fostering more aware employment of digital media can bring benefits of various kinds, one of which is everyday digital well-being. To give some examples, the test cohort began using smartphones in a less obsessive and invasive manner, scaled back their social media use, and experienced less distress related to the internet; in addition, their indices of life satisfaction and happiness increased (Gui et al., 2018).

Approaching digital well-being within education involves helping students to use digital technologies in a safe and effective manner. The need to include safety as a core element of digital education is born out by the risks to students' individual well-being posed by phenomena like technology addiction and cyber bullying. This is supported by Mascia et al. (2020), who shed light on the multifaceted relationships that impact on the well-being of adolescents and confirmed the harm that smartphone addiction exerts on well-being and self-regulation. The European Commission's Digital Competence Framework (DigComp 2.1) (Carretero et al., 2017) positions well-being in relation to competences connected to Safety. The framework describes protecting health and well-being as the individual's ability to safeguard oneself and others from possible dangers and to limit risks while using technologies, including understanding the potential of technologies *for promoting social well-being and inclusion*. Seen thus, DigComp2.1 defines the social *well-being* and *social inclusion* of leaders as complementary objectives that coalesce, especially regarding the affordances and outcomes of employing digital technologies for learning (e.g., Jones and Sandford, 2019).

Similarly, at school level, the European Framework for Digitally Competent Educational Organizations (DigCompOrg) provides a comprehensive background for effectively integrating digital technologies in educational organizations (Kampylis et al., 2015; Mattar et al., 2020). This conceptual framework is the basis for the SELFIE tool[1] that gives schools a holistic view of how students, teachers, and school leaders perceive the digital *status quo* of their policies and practices (Castaño-Muñoz et al., 2018).

## Perceptions of Students, Teachers, and School Leaders on Well-Being and Inclusion Expressed Through SELFIE

SELFIE (Self-assessment tool for digitally capable schools) is one of the 11 initiatives set out in the Digital Education Action Plan adopted by the European Commission (2018) to promote self-assessment of digital and innovative education practices in the school context. Available in the 24 official languages of the European Union, SELFIE gathers—anonymously—the views of students, teachers, and school leaders on how technology is being used in their context. In order to implement SELFIE, an analysis was conducted on several self-assessment tools of digital readiness developed and/or used in Europe (Kampylis et al., 2016), such as Opeka and Ropeka tools in Finland (Tanhua-Piiroinen and Viteli, 2017) and Digital Schools of Distinction in Ireland (O'Leary, 2018) to name a few.

With a pilot involving more than 65,000 schools' actors in 14 countries (e.g., Kampylis et al., 2019), including Italy (Directorate-General for Education, Youth, Sport, and Culture,, 2019; Bocconi et al., in press), SELFIE encompasses elements and descriptors that may be regarded *as intrinsically linked to students'*

[1]https://ec.europa.eu/education/schools-go-digital_en

*well-being and inclusion in schools*, fostering a deep reflection that spans from *organizational responsibilities* (e.g., Leadership, Infrastructure) to *individual responsibilities* (e.g., Teaching and Learning Practices).

While some instruments embody the specific national characteristics, SELFIE is a European level initiative that provides transparency regarding schools' digital competence, thus aiding comparison of education systems across the continent, with benefits both for peer learning and for policymaking. By helping entire learning communities (students, teachers, and school leaders) to engage in a cyclical self-reflection process, SELFIE supports schools to grasp their progress in digitally enhanced teaching and learning. SELFIE helps them plan out an ongoing development path in terms of digital strategies and praxis, and in doing so to address issues of inclusion and learner well-being.

Thus far, studies investigating SELFIE have largely concentrated on describing the tool and associated self-reflection process, and on identifying similarities and differences with similar undertakings. The outcomes of these studies reveal that SELFIE is among the very few instruments designed for comprehensive involvement of learners in the digital self-evaluation that schools conduct (Kampylis et al., 2016, 2019; Castaño-Muñoz et al., 2018). This is a fundamental characteristic of SELFIE and its contribution to promoting Student Voice, especially in identification of any potential misalignments existing between school policy/strategy level and actual teaching/learning activities.

Other research studies (Szûcs, 2019; Bocconi et al., in press) shed light on the way in which SELFIE satisfies the need to address aspects of digital innovation across the entire educational organization. Elsewhere, Jeladze and Pata (2018) demonstrate that the data generated from the SELFIE self-reflection process help schools employing digitally technologies get a better grasp of their progress in this area; this highlights the considerable variation in levels of the digital competence among different schools. Employment of SELFIE outside the school context is investigated by Broek and Buiskool (2020), who thereby shed light on the tool's affordances for application to non-formal and informal learning.

This contribution is unique among the research studies performed thus far on SELFIE in the sense that it reveals how the personalizable questionnaire structure and content SELFIE proposes encompass key aspects in education contexts like learner well-being and inclusion, factors that are often disregarded by similar tools focusing on digital technology use. Specifically, this study is unique among the research performed into SELFIE for the following reasons:

1. it reveals that the detail encompassed within the questionnaire makes it possible to investigate critical domain-independent areas in the school context like how the employment of digital technologies impacts on learner well-being and inclusion. In this case, perceptions on the topics of well-being and inclusion were sampled from a cross-section of the school learning community undergoing SELFIE pilot testing in Italy.
2. it furthers extant research into student well-being and inclusion via digital technology use within the context of schools' innovation vision and planning (Carretero et al., 2017; Govorova et al., 2020). There is currently a lack of published research into these issues, so the work reported here presents fresh insights.
3. the reported findings may help to inform ongoing development of SELFIE so that the tool might help users gain a better understanding of well-being and inclusion, an aspect that characterizes other frameworks (e.g., NSW Department of Education and Communities, 2015) as well as internationally adopted evaluation tools such as PISA (e.g., OECD, 2019).

## Research Question and Hypothesis

The main objective of this study is to investigate students' well-being and inclusion through technologies within school communities, by analyzing the perception that students, teachers, and school leaders express regarding practices deployed inside the school contexts. Specifically, the research question guiding the present study is: *How do students, teachers, and school leaders perceive students' well-being and inclusion through technologies in schools' policies and practices, at different education levels?*

Starting from this research question, our hypotheses are as follows:

> *Hypothesis 1 (H1)*: students, teachers, and school leaders perceive a relationship between well-being and inclusion in the strategies and practices of their school;
> *Hypothesis 2 (H2):* students, teachers, and school leaders across education levels (primary, lower secondary, upper secondary general, upper secondary vocational) perceive students' well-being and inclusion through technologies differently;
> *Hypothesis 3 (H3):* within schools, there is a relationship between the perception of students, teachers, and school leaders on student's well-being and inclusion through technologies at different education levels.

# MATERIALS AND METHODS

## Participants

This research is part of the wider European project to pilot test an online self-assessment questionnaire called SELFIE, which deals with the perceptions that students, teachers, and school leaders have on the use of digital technologies in the school context.

A total of 201 Italian schools in 10 different regions took part in this piloting project in 2017 on a voluntary basis. Respondents to the SELFIE questionnaire comprised 24,715 students, 5,690 teachers, and 1,507 school leaders (including principal, vice-principal, subject coordinators, ICT coordinators, etc.) at different education levels (Primary, Lower Secondary, Upper Secondary general, and Upper Secondary Vocational levels), for a total of 31,912 participants (for details, see **Table 1**).

**TABLE 1 |** Participant type for each education level.

| | Education level | | | | |
|---|---|---|---|---|---|
| **Users type** | **Primary** | **Lower secondary** | **Upper secondary general** | **Upper secondary vocational** | **Total** |
| Students | 3,158 | 5,572 | 14,295 | 1,690 | 24,715 |
| Teachers | 1,651 | 1,212 | 2,424 | 403 | 5,690 |
| School Leaders | 391 | 420 | 572 | 124 | 1,507 |
| Total | 5,200 | 7,204 | 17,291 | 2,217 | 31,912 |

## Procedure

From the end of September to early October 2017, students, teachers, and school leaders participated in the pilot project in Italy by filling in the online SELFIE questionnaires. During the piloting phase, participants were supported by local educational authorities—or by Institutes of Educational Research – under the supervision of the SELFIE national coordinator, namely, the National Research Council of Italy's Institute for Educational Technology (Bocconi et al., in press).

Before the piloting activities commenced, teachers acting as SELFIE coordinators (Bocconi and Panesi, 2018a) in each pilot school were fully trained to inform their local school community (i.e., school leader, teachers, and students) about the initiative and to manage the piloting process in the school.

Furthermore, the local SELFIE coordinators registered their school/s on the SELFIE platform; subsequently, the European Commission sent them an e-mail containing a link granting access to the SELFIE questionnaires. The local SELFIE coordinators then distributed the link to participants so that they could fill in the online self-assessment questionnaires.

## Measures

*The SELFIE tool* (see Castaño-Muñoz et al., 2018, for details on reliability, consistency, and validity data) has been developed on the basis of the DigCompOrg conceptual framework dedicated to the digital competence of educational organizations (Kampylis et al., 2015; see also Kampylis et al., 2016). SELFIE includes three different instantiations, respectively, devoted to the three different target users foreseen for the questionnaire: school leaders, teachers, and students. The information collected from the student questionnaires were on their use of digital technologies in the school context. The teacher questionnaires concentrated on their practices linked to school policies. The information collected from school leader questionnaires regarded policies on the use of digital technologies for learning in the school.

Overall, the three instantiations of the SELFIE questionnaire are composed of originals items (i.e., not derived or linked to other studies) organized in the following categories: (a) Core items: these were constructed according to the DigCompOrg framework and were mandatory for all respondents; (b) Attitude and belief items: these contextualize the information collected from the core items; (c) Optional items: these were constructed according to the DigCompOrg framework but schools could decide whether to include them or not in their self-reflection exercise; (d) School-specific items: these could be composed by the individual schools themselves and added to their questionnaire; (e) Vocational-specific items: these applied exclusively to vocational schools; (f) Background characteristics: questions on the school's demographics, resources, and context (Castaño-Muñoz et al., 2018). All items were evaluated on a five-point Likert scale.

In this study, we identified a set of SELFIE original items focusing on students' well-being and inclusion through technologies, in line with the literature (van Petegem et al., 2007; Booth and Ainscow, 2011; NSW Department of Education and Communities, 2015; Anderson and Graham, 2016; Carretero et al., 2017; OECD, 2019; Trentin, 2019). Specifically, the selected SELFIE original items (i) were common to all three respondent groups and (ii) reflected their perceptions on six components related to students' well-being and inclusion through technologies, namely: *relationships*, *school community*, *safety*, *individual learning needs*, *active learner*, and *collaboration* (for details, see **Table 2**).

## Statistics Analysis

To conduct qualitative investigation of the mean perceptions of students, teachers, and school leaders about items dedicated to students' well-being and inclusion through technologies, descriptive statistics on scores related to those items were calculated. Following the literature (Könings et al., 2014), ANOVAs were not conducted to compare the means from the three groups (students, teachers, school leaders) because, in answering the questionnaire, each group was called upon to focus on different aspects: students on their use of digital technologies in the school context; teachers on their practices linked to school policies; and school leaders on policies related to the use of digital technologies for learning in the school (Castaño-Muñoz et al., 2018). For this reason, the literature stresses that it is not enough to compare only mean perception scores of different stakeholder groups (Könings et al., 2014). The relations among the scores related to the items focused on students' well-being and inclusion for each user group (students, teachers, and school leaders) were investigated with bivariate correlations. A series of confirmatory factor analyses (CFAs), based on covariance matrices, were conducted to verify the latent structure of the perception of students' well-being and inclusion for each user group.

Specifically, we conducted the CFA based on different theoretical models: (1) a single factor model, consistent with DigComp2.1 (Competence area "Safety", skill "4.2 Protecting health and well-being") (Carretero et al., 2017);

**TABLE 2 |** SELFIE Tool: selected original items focused on well-being and inclusion.

| Components of well-being | SELFIE areas Castaño-Muñoz et al., 2018 | SELFIE item ID | Item formulation in SELFIE questionnaire (for students, teachers, and school leaders) |
|---|---|---|---|
| Relationships: Learners' relations with peers and educators within the school context | 1.1 Benefits and challenges are openly discussed (Area 1. Leadership) | SL_1.1<br>T_1.1. A<br>S_1.3 | SCHOOL LEADERS: In our school, we discuss with teachers and students the benefits and challenges of using digital technologies for teaching and learning<br>TEACHERS: In my school, I discuss with school leaders, teachers and students the benefits and challenges of using digital technologies for learning<br>STUDENTS: In my school, our teachers discuss with us the benefits and drawbacks of using digital technologies for learning |
| School community: student–teacher communications, students cooperation and teamwork, and students' feelings about their social life | 1.7 Use of different communication tools (Area 5 Infrastructure) | SL_1.7<br>T_1.4<br>S_1.5 | SCHOOL LEADERS: As part of our digital strategy, we use different communication tools within and beyond the school community according to our different communication purposes and target groups<br>TEACHERS: In my school, I use different communication tools according to the different communication purposes and target groups<br>STUDENTS: In my school, we use digital technologies for communicating with teachers and other students |
| Safety: rules, attitudes and school strategies for minimizing the negative impact on students' mental well-being potentially caused by use of technology | 1.13 Students learn how to behave safely and responsibly (Areas 4 + 6 Digital competence as outcome) | SL_1.13<br>T_1.10<br>S_1.7 | SCHOOL LEADERS: As part of our digital strategy, we have guidelines for students on the safe and responsible use of digital technologies<br>TEACHERS: In my school, I teach my students how to behave in a safe and responsible way, online and offline<br>STUDENTS: In my school, I learn how to behave in a safe and responsible way, online and offline |
| **Components of inclusion** | **SELFIE Areas Castaño-Muñoz et al., 2018** | **SELFIE Item ID** | **Item formulation in SELFIE questionnaire (for students, teachers, and school leaders)** |
| Individual learning needs: personalizing learning and addressing individual learning needs | 2.5 Digital technologies are used to address individual learning needs (Area 2. Pedagogy) | SL_3.9<br>T_4.4<br>S_2.4 | SCHOOL LEADERS: In our school, teachers use digital technologies to address individual learners' needs<br>TEACHERS: When digital technologies are used in school, it is easier to address students' individual needs<br>STUDENTS: In my school, I get to do special digital activities if I need extra help or if I am ahead of the class |
| Actively involve students: all those in the school's learning community take an active part and play a constructive role in positive relationship building | 2.3 Digital technologies are used to actively involve students (Area 2. Pedagogy) | SL_2.3<br>T_2.3<br>S_2.2 | SCHOOL LEADERS: It is part of our digital strategy, to use digital technologies to actively involve students in their learning<br>TEACHERS: As a teacher, I use digital technologies to actively involve students in their learning<br>STUDENTS: In my school, I use digital technologies to become a more active learner |
| Collaboration: Students learn collaboratively | 2.10 Digital technologies are used for student collaboration (Area 2. Pedagogy) | SL_2.10<br>T_2.10<br>S_5.5 | SCHOOL LEADERS: It is part of our digital strategy, that students collaborate using digital technologies<br>TEACHERS: As a teacher, I use digital technologies to help students collaborate with each other<br>STUDENTS: In my school, my classmates and I help each other when we have problems with digital technologies |

(2) a two-factor model, consistent with The Well-being Framework for Schools (NSW Department of Education and Communities, 2015), where inclusions are seen as a different but interrelated factor contributing to well-being in the school context. To conduct CFA, EQS 6.1 software has been used (Bentler, 2006).

Multiple fit indices were considered to compare models: chi-square ($\chi^2$), the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), comparative fit index (CFI), the root mean square error of approximation (RMSEA), the standardized root mean squared residual (SRMR), and the Akaike information criterion (AIC). Chi-square was used to evaluate the appropriateness of the CFA model; non-significant values indicated a minor difference between the covariance matrix generated by the model and the observed matrix and, thus, an acceptable fit. GFI values of $>0.95$ can be considered a good fit and $>0.90$ can be considered an acceptable fit (Schermelleh-Engel et al., 2003). An AGFI index of $>0.90$ is indicative of good fit, while values $>0.85$ may be considered as an acceptable fit (Schermelleh-Engel et al., 2003). CFI values of $>0.97$ can be considered a good fit (Schermelleh-Engel et al., 2003), and $>0.95$ can be considered an acceptable fit (Schermelleh-Engel et al., 2003). RMSEA levels of $<0.05$ indicate a good fit and acceptable if $<0.08$ (Kline, 2005). SRMR $<0.05$ represents a good fit and $<0.10$ is acceptable (Schermelleh-Engel et al., 2003). AIC is used to compare models: the model with the lowest AIC values is to be preferred.

Furthermore, in the two-factor CFA model based on the theoretical well-being framework for Schools (NSW Department of Education and Communities, 2015), for each of the three samples (students, teachers and school leaders) and each of the two factors measured, we calculated McDonald's omega, composite reliability (CR), average variance extracted (AVE),

discriminant validity, and convergent validity with R package (Rosseel, 2012; Jorgensen et al., 2019), also in line with Zhang and Yuan (2016). First, McDonald's omega (McDonald, 1970, 1999) was used to measure internal consistency reliability; values considered acceptable were those above 0.70 (Viladrich et al., 2017). Furthermore, in an effort to gauge construct validity, we employed (a) *convergent validity*, namely, confidence that the adopted indicators had accurately measured a trait, and (b) *discriminant validity*, namely, the extent to which trait measurements are not related (Campbell and Fiske, 1959; Jöreskog, 1969). To determine convergent validity, we referred to the criterion proposed by Fornell and Larcker (1981), a commonly adopted approach for evaluating the level of variance that a model's latent variables may share. According to this criterion, the convergent validity of the measurement model can be assessed by the AVE and CR. AVE measures the level of variance captured by a construct versus the level due to measurement error, and values above 0.70 are considered very good, whereas the values of 0.50 are acceptable. CR is a less biased estimate of reliability than Cronbach's alpha; the acceptable value of CR is 0.70 and above. According to Fornell and Larcker (1981), discriminant validity can be assessed by comparing the amount of the variance capture by the construct and the shared variance with other constructs. Thus, the levels of square root of the AVE for each construct should be greater than the correlation involving the constructs.

Based on the CFA results, composite scores were calculated as the mean of the well-being and inclusion *z* score to represent the two latent dimensions. For each user group (students, teachers, and school leaders), ANOVAs were conducted with the composite well-being and inclusion scores as dependent variables, and education level group membership (primary, lower secondary, upper secondary general, and upper secondary vocational) as the between-subject variable to explore group differences in the well-being and inclusion components. Finally, for each school, the average scores expressed, respectively, by students, teachers, and school leaders on each item were calculated. For each type of participant (students, teachers, and school leaders), composite scores were calculated as the mean of the well-being and inclusion *z* score to represent the two latent dimensions. Bivariate correlations among these scores were conducted separately for each education level.

## RESULTS

Descriptive statistics results for the three groups (students, teachers, and school leaders) concerning the items focused on students' well-being and inclusion through technologies are shown in **Table 3**. Overall, the mean scores of the participants' perspective for each item showed teachers as the group that expressed the most negative reaction, with the exception of two items ("Use of different communication tools" and "Digital technologies are used to address individual learning needs"), where students expressed a more negative perspective.

**Table 4** reports the correlations among SELFIE items focusing on students' well-being and inclusion for the three participant groups: students, teachers, and school leaders. The findings showed reasonable correlations among all items for students ($r$ = 0.194 to $r$ = 0.447; all $p < 0.001$), teachers ($r$ = 0.266 to $r$ = 0.621; all $p < 0.001$), and school leaders ($r$ = 0.392 to $r$ = 0.651; all $p < 0.001$). Particularly strong correlations emerged for the school leader group. Furthermore, for all groups, the three scores related to the answers to items focused on students' well-being showed reasonable correlations with the three scores related to answers to the items focused on inclusion (students: $r$ = 0.212 to $r$ = 0.259; all $p < 0.001$; teachers: $r$ = 0.266 to $r$ = 0.422; all $p < 0.001$; school leaders: $r$ = 0.402 to $r$ = 0.508; all $p < 0.001$).

Based on these last findings, we also performed a set of CFAs to determine the latent structure of students' well-being and inclusion for students, teachers and school leaders. For each group (students, teachers, and school leaders) two alternative models were tested. Model A assumed that all scores loaded on a single factor. In model B, the three well-being scores loaded on one factor and the three inclusion scores loaded on the other factor. **Table 5** summarizes the fit indices for these models. All models showed at least an acceptable fit to the data. For each of the three groups involved, the lowest AIC, RMSEA, and SRMR, and the highest GFI, AGFI, and CFI were found for models B (factor 1: well-being; factor 2: inclusion) Therefore, models B provided the best fit.

The fit of nested models can be compared by subtracting the $\chi^2$ value of the less restricted model. Following previous research (e.g., Miyake et al., 2000; Panesi and Morra, 2020), in this study we also compare models A and B in this direct way. This is because positing one factor is mathematically equivalent to fixing at 1 the value of the parameter phi that represents the correlation between the two factors posited in any of the two-factor models. The fit of model B was significantly better than the fit of model A for students [$\chi^2$diff(1) = 532.348], teachers [$\chi^2$diff(1) = 319.406], and school leaders [$\chi^2$diff(1) = 39.484].

It should also be noted that, in the endorsed models, the estimated correlations between the two latent variables were quite large for students (phi = 0.79), teachers (phi = 0.82), and school leaders (phi = 0.90) (see **Figure 1**), and these findings are in line with the Well-being Framework for Schools (NSW Department of Education and Communities, 2015).

Furthermore, concerning the two-factor model (Model B), **Table 6** shows, for each of the three users (students, teachers, and school leaders), the reliability of the two constructs (factor 1: well-being; factor 2: inclusion) as measured by McDonald's omega, AVE, CR, and the square root of the AVE, to also investigate convergent and discriminant validity.

Concerning *internal reliability*, for the school leaders' group, McDonald's omega met the acceptable level for both well-being and inclusion (factor 1:0.70; factor 2:0.75), while for the teachers' group, McDonald's omega was slightly lower than the acceptable value (0.67) for well-being and above the threshold (0.71) for inclusion. Conversely, for the students, McDonald's omega fell below the acceptable value for both factors (factor 1:0.57; factor 2:0.55). Therefore, reliability indexed by McDonald's omega is sufficient for well-being and inclusion factors, with the exception of students. It should be noted that these findings mimic the characteristics of the SELFIE instrument, where school leaders

**TABLE 3 |** Descriptive statistics.

| | Students | | Teachers | | School leaders | |
|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Benefits and challenges are openly discussed | 3.21 | 1.21 | 2.86 | 1.12 | 3.28 | 0.99 |
| Use of different communication tools | 3.25 | 1.94 | 3.40 | 1.01 | 3.68 | 0.94 |
| Students learn how to behave safely and responsibly | 3.48 | 1.24 | 3.43 | 1.08 | 3.69 | 0.94 |
| Digital technologies are used to address individual learning needs | 2.83 | 1.30 | 3.60 | 0.90 | 3.38 | 0.89 |
| Digital technologies are used to actively involve students | 3.46 | 1.11 | 3.36 | 0.96 | 3.64 | 0.84 |
| Digital technologies are used for student collaboration | 3.71 | 1.13 | 3.12 | 1.08 | 3.44 | 0.97 |

**TABLE 4 |** Zero-order correlations between scores related to the items focused on students' well-being and inclusion (separately for students, teachers, and school leaders).

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **Students** | | | | | | |
| Benefits and challenges are openly discussed | | 0.207*** | 0.447*** | 0.212*** | 0.259*** | 0.257*** |
| Use of different communication tools | | | 0.191*** | 0.212*** | 0.259*** | 0.257*** |
| Students learn how to behave safely and responsibly | | | | 0.233*** | 0.232*** | 0.230*** |
| Digital technologies are used to address individual learning needs | | | | | 0.327*** | 0.263*** |
| Digital technologies are used to actively involve students | | | | | | 0.194*** |
| Digital technologies are used for student collaboration | | | | | | |
| **Teachers** | | | | | | |
| Benefits and challenges are openly discussed | | 0.381*** | 0.402*** | 0.280*** | 0.393*** | 0.425*** |
| Use of different communication tools | | | 0.460*** | 0.302*** | 0.422*** | 0.412*** |
| Students learn how to behave safely and responsibly | | | | 0.266*** | 0.398*** | 0.420*** |
| Digital technologies are used to address individual learning needs | | | | | 0.367*** | 0.324*** |
| Digital technologies are used to actively involve students | | | | | | 0.621*** |
| Digital technologies are used for student collaboration | | | | | | |
| **School leaders** | | | | | | |
| Benefits and challenges are openly discussed | | 0.392*** | 0.452*** | 0.455*** | 0.508*** | 0.462*** |
| Use of different communication tools | | | 0.462*** | 0.428*** | 0.464*** | 0.402*** |
| Students learn how to behave safely and responsibly | | | | 0.462*** | 0.478*** | 0.430*** |
| Digital technologies are used to address individual learning needs | | | | | 0.547*** | 0.497** |
| Digital technologies are used to actively involve students | | | | | | 0.651*** |
| Digital technologies are used for student collaboration | | | | | | |

*$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.*

**TABLE 5 |** Fit indices of the hypothesized models (separately for students, teachers and school leaders).

| | Models | Factors | $\chi^2$ | df | *P* | RMSEA | SRMR | GFI | AGFI | CFI | AIC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Students | A | 1 factor: Well-being/inclusion | 1,769.508 | 9 | 0.000 | 0.089 | 0.042 | 0.975 | 0.942 | 0.910 | 1,751.508 |
| | | 2 factors | 1,237.160 | 8 | 0.000 | 0.079 | 0.041 | 0.983 | 0.957 | 0.937 | 1,221.160 |
| | B | (1) Well-being; | | | | | | | | | |
| | | (2) Inclusion | | | | | | | | | |
| Teachers | A | 1 factor: Well-being/inclusion | 439.839 | 9 | 0.000 | 0.092 | 0.036 | 0.973 | 0.937 | 0.952 | 421.839 |
| | | 2 factors | 120.433 | 8 | 0.000 | 0.050 | 0.021 | 0.993 | 0.982 | 0.987 | 104.433 |
| | B | (1) Well-being; | | | | | | | | | |
| | | (2) Inclusion | | | | | | | | | |
| School Leaders | A | 1 factor: Well-being/inclusion | 90.980 | 9 | 0.000 | 0.078 | 0.029 | 0.979 | 0.950 | 0.974 | 72.978 |
| | | 2 factors | 51.496 | 8 | 0.000 | 0.060 | 0.023 | 0.988 | 0.969 | 0.986 | 35.496 |
| | B | (1) Well-being; | | | | | | | | | |
| | | (2) Inclusion | | | | | | | | | |

**TABLE 7 |** Descriptive statistics, ANOVA, and *post hoc* (Bonferroni).

| | | Students | | | | | | Teachers | | | | | | School leaders | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Education level** | **Descriptive statistics** | | **ANOVA** | | ***Post hoc* (Bonferroni)** | | **Descriptive statistics** | | **ANOVA** | | ***Post hoc* (Bonferroni)** | | **Descriptive statistics** | | **ANOVA** | | ***Post hoc* (Bonferroni)** | |
| | | ***M*** | ***SD*** | ***F*** | ***Sig*** | **Comparisons** | ***Sig.*** | ***M*** | ***SD*** | ***F*** | ***Sig*** | **Comparisons** | ***Sig.*** | ***M*** | ***SD*** | ***F*** | ***Sig*** | **Comparisons** | ***Sig.*** |
| Well-being | Primary | 0.223 | 0.713 | 550.803 | 0.000 | *P* < LS | 0.002 | −0.115 | 0.835 | 15.679 | 0.000 | *P* < LS | 0.000 | 0.018 | 0.792 | 2.474 | 0.060 | *P* = LS | 1.000 |
| | Lower secondary | 0.285 | 0.724 | | | *P* > USG | 0.000 | 0.068 | 0.812 | | | *P* < USG | 0.000 | 0.060 | 0.788 | | | *P* = USG | 1.000 |
| | Upper secondary general | −0.161 | 0.802 | | | *P* > USV | 0.000 | 0.044 | 0.833 | | | *P* < USV | 0.072 | −0.020 | 0.898 | | | *P* = USV | 0.211 |
| | Upper secondary vocational | 0.005 | 0.742 | | | LS > USG | 0.000 | 0.000 | 0.792 | | | LS = USG | 1.000 | −0.164 | 0.879 | | | LS = USG | 0.815 |
| | | | | | | LS > USV | 0.000 | | | | | LS = USV | 0.936 | | | | | LS = USV | 0.054 |
| | | | | | | USG < USV | 0.000 | | | | | USG = USV | 1.000 | | | | | USG = USV | 0.503 |
| Inclusion | Primary | 0.168 | 0.743 | 132.916 | 0.000 | *P* > LS | 0.001 | −0.223 | 0.888 | 59.365 | 0.000 | *P* < LS | 0.000 | −0.035 | 0.820 | 1.637 | 0.179 | *P* = LS | 1.000 |
| | Lower secondary | 0.102 | 0.783 | | | *P* > USG | 0.000 | 0.009 | 0.870 | | | *P* < USG | 0.000 | −0.021 | 0.882 | | | *P* = USG | 0.644 |
| | Upper secondary general | −0.079 | 0.769 | | | *P* > USV | 0.000 | 0.149 | 0.873 | | | *P* < USV | 0.000 | 0.061 | 0.976 | | | *P* = USV | 1.000 |
| | Upper secondary vocational | 0.016 | 0.771 | | | LS > USG | 0.000 | −0.008 | 0.854 | | | LS < USG | 0.000 | −0.099 | 0.917 | | | LS = USG | 0.926 |
| | | | | | | LS > USV | 0.000 | | | | | LS = USV | 1.000 | | | | | LS = USV | 1.000 |
| | | | | | | USG < USV | 0.000 | | | | | USG > USV | 0.005 | | | | | USG = USV | 0.443 |

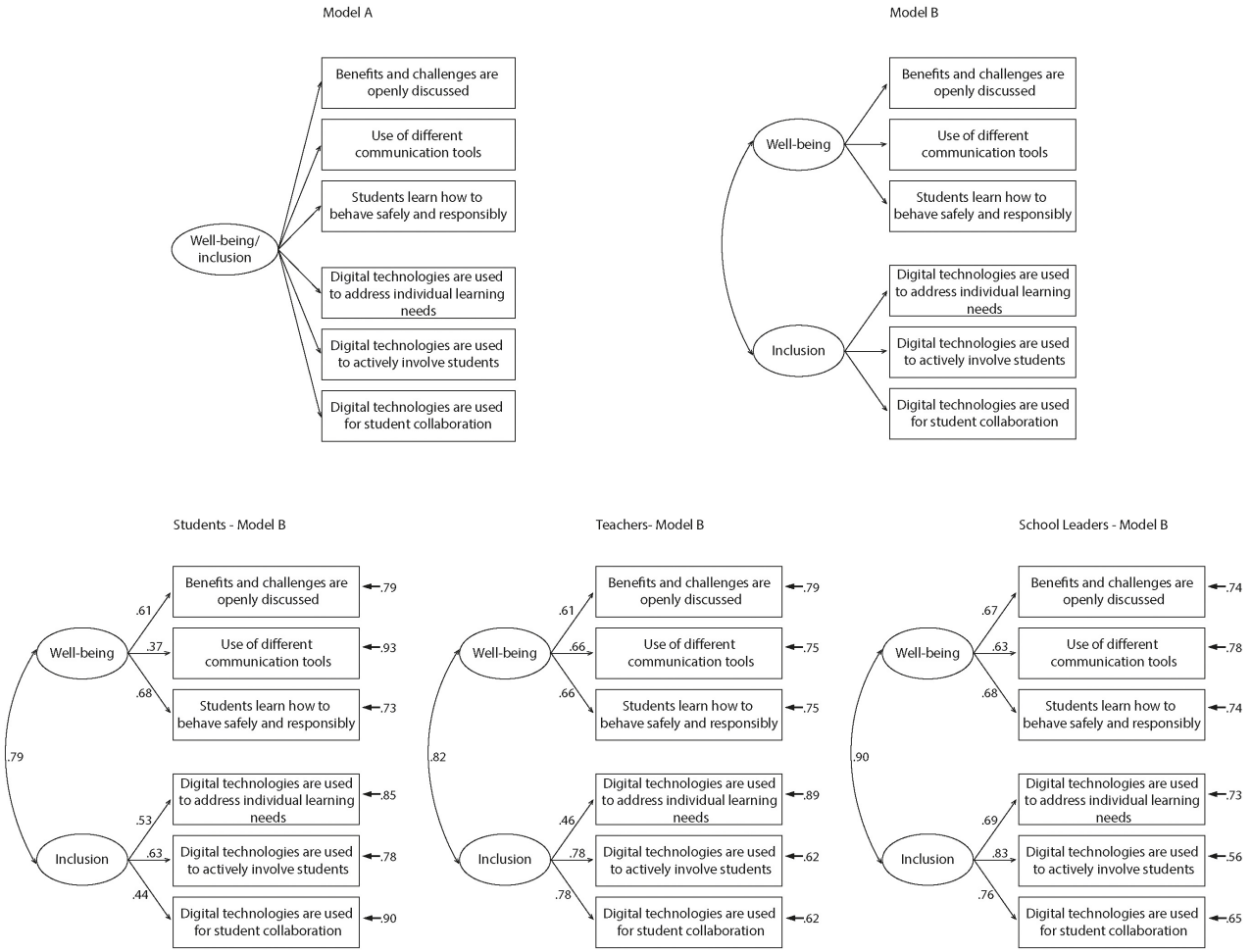


**FIGURE 1 |** Well-being and inclusion models. Models B are the endorsed models, respectively, for students, teachers, and school leaders (standardized parameters are reported).

and teachers provide, respectively, the perspective on the two factors (well-being and inclusion) in schools' policy and practice, while students' perspective falls in between the two layers, hence reasonably having a more limited perception of both.

**TABLE 6 |** The two-factor CFA model, for each of the three samples' McDonald's omega, AVE, CR, and the square root of the AVE.

| | McDonald's omega | AVE | CR | √AVE | *phi* |
|---|---|---|---|---|---|
| **Students** | | | | | |
| Well-being | 0.57 | 0.32 | 0.71 | 0.56 | 0.79 |
| Inclusion | 0.55 | 0.29 | 0.71 | 0.54 | |
| **Teachers** | | | | | |
| Well-being | 0.67 | 0.40 | 0.82 | 0.63 | 0.82 |
| Inclusion | 0.71 | 0.47 | 0.95 | 0.68 | |
| **School leaders** | | | | | |
| Well-Being | 0.70 | 0.44 | 0.84 | 0.66 | 0.90 |
| Inclusion | 0.75 | 0.50 | 0.91 | 0.71 | |

*Internal validity (McDonald's omega, >0.70); convergent validity: AVE (>0.50) and CR (>0.70); discriminant validity (comparison of AVE squared root > phi, correlation score of pair of latent variables).*

Regarding the *convergent validity*, the CR for both well-being and inclusion surpassed the 0.70 threshold for all the three users. The AVE reached the recommended level of 0.50 in factor 2—inclusion for the school leaders, except for teachers and students. AVE values also scored below the threshold in factor 1—well-being for all the three actors. Although the AVE acceptable minimum cutoff point is 0.50, convergent validity may still be considered adequate because all latent factors had CR values above 0.70 (Malhotra and Dash, 2011). According to Fornell and Larcker (1981), convergent validity can be established via CR alone because AVE is a more conservative measure and relatively strict compared to CR. Likewise, other authors (Angel et al., 2019; Ibrahim et al., 2019; Shaakumeni and Csapoì, 2019) also insisted that latent variables' AVE can fall below 0.50 if its CR is satisfactory.

Finally, concerning *discriminant validity*, the square root of the AVE values for both factor 1—well-being and factor 2—inclusion did not exceed the correlation between the two

constructs for all the three groups as required by Fornell and Larcker (1981). It is important to note, however, that discriminant validity is not exclusively an empirical means to validate a model (Farrell, 2010). Theoretical foundations and arguments should provide reasons for constructs correlating or not (Bollen and Lennox, 1991), as concepts are partly defined by their relationships with other concepts in a conceptual network (Henseler et al., 2015). Therefore, failure to establish discriminant validity between the two constructs (well-being and inclusion) does not necessarily imply that the underlying concepts are identical, especially when research studies continued support for conceptualizing these two constructs as separated but interrelated (NSW Department of Education and Communities, 2015; OECD, 2019).

Based on the CFA results, for each group (students, teachers, and school leaders), two composite scores representing well-being and inclusion were calculated as the mean of the *z* scores, as follows: (A) for well-being: the *z* score average of the answers to the items (1) Benefits and challenges are openly discussed; (2) Use of different communication tools; (3) Students learn how to behave safely and responsibly; (B) for inclusion: the *z* score average of answers to the items (1) Digital technologies are used to address individual learning needs; (2) Digital technologies are used to actively involve students; (3) Digital technologies are used for student collaboration.

The results of ANOVAs conducted with the two composite well-being and inclusion measures as dependent variables and education level group membership (primary, lower secondary, upper secondary general, and upper secondary vocational) as the between-subject variable showed that for students and teachers, differences between education level groups emerged in both well-being and inclusion. In particular, differences were found among students of different education level groups. For the well-being component, the lower secondary group returned a more positive perception than all the others, followed by the primary group and the upper secondary vocational group. For the inclusion component, the primary group returned a more positive perception than all the others, followed by the lower secondary group and the upper secondary vocational group (for details, see **Table 7**). Concerning teachers, the primary group returned lower scores than other groups in both well-being and inclusion. For inclusion only, differences between the lower secondary group and the upper secondary general group and between the upper secondary general group and the upper secondary vocational group emerged. No differences were found in either the well-being or inclusion components among school leaders of any education level group.

Finally, average school scores were calculated for students, teachers, and school leaders on each item. Subsequently, composite scores were calculated as the mean of the well-being and inclusion *z* score to represent the two latent dimensions, respectively, for students, teachers, and school leaders. **Table 8** reports the correlations among the two composite well-being and inclusion measures that reflect students', teachers', and school leaders' perception at different education levels. Overall, the findings showed strong correlations among teachers' and school leaders' perception on both well-being and inclusion at all education levels ($r = 0.476$ to $r = 0.622$, all $p < 0.001$), except for upper secondary vocational level.

Concerning students' perception on well-being and inclusion, some particular results may be observed at school level. In primary schools, the highest correlations with teachers' perception on well-being and inclusion ($r = 0.357$ to $r = 0.495$, all $p < 0.01$) emerged, but also reasonable correlations with school leaders' perception on inclusion and an acceptable correlation between students' perception on inclusion and school leaders' perception on well-being emerged (all $p < 0.05$). In lower secondary schools, all scores correlated significantly (all $p < 0.01$), and in particular, students' perception on well-being correlated highest with teachers' perception on well-being ($r = 0.432$, $p < 0.001$) and subsequently with school leaders' perception on well-being ($r = 0.396$, $p < 0.01$) and students' perception on inclusion correlated highest with teachers' perception on inclusion ($r = 0.652$, $p < 0.001$) and subsequently with school leaders' perception on inclusion ($r = 0.525$, $p < 0.001$). In upper secondary general schools, high correlations with teachers' perception on well-being and inclusion emerged (all $p < 0.001$), followed by significant correlations with school leaders' perception on well-being and inclusion (all $p < 0.01$). In

**TABLE 8 |** Zero-order correlations within schools on well-being and inclusion (separately for education levels).

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **Primary (School *n* = 59)** | | | | | | |
| 1. Students' well-being | | 0.392** | 0.407** | 0.357** | 0.237 | 0.309* |
| 2. Students' inclusion | | | 0.418** | 0.495*** | 0.279* | 0.423** |
| 3.Teachers' well-being | | | | 0.806*** | 0.540*** | 0.542*** |
| 4.Teachers' inclusion | | | | | 0.519*** | 0.476*** |
| 5.School leaders' well-being | | | | | | 0.665*** |
| 6. School leaders' inclusion | | | | | | |
| **Lower secondary (School *n* = 65)** | | | | | | |
| 1.Students' well-being | | 0.516*** | 0.432*** | 0.380** | 0.396** | 0.272* |
| 2.Students' inclusion | | | 0.494*** | 0.652*** | 0.480*** | 0.525*** |
| 3.Teachers' well-being | | | | 0.796*** | 0.519*** | 0.491*** |
| 4.Teachers' inclusion | | | | | 0.523*** | 0.511*** |
| 5.School leaders' well-being | | | | | | 0.791*** |
| 6. School leaders' inclusion | | | | | | |
| **Upper secondary general (School *n* = 62)** | | | | | | |
| 1.Students' well-being | | 0.679*** | 0.650*** | 0.494*** | 0.461*** | 0.351** |
| 2.Students' inclusion | | | 0.531*** | 0.464*** | 0.412** | 0.453*** |
| 3.Teachers' well-being | | | | 0.901*** | 0.621*** | 0.555*** |
| 4.Teachers' inclusion | | | | | 0.622*** | 0.587*** |
| 5.School leaders' well-being | | | | | | 0.849*** |
| 6. School leaders' inclusion | | | | | | |
| **Upper secondary vocational (School *n* = 14)** | | | | | | |
| 1.Students' well-being | | 0.434 | 0.635* | 0.246 | 0.224 | 0.348 |
| 2.Students' inclusion | | | 0.284 | 0.390 | 0.049 | 0.143 |
| 3.Teachers' well-being | | | | 0.757** | 0.338 | 0.521 |
| 4.Teachers' inclusion | | | | | 0.114 | 0.252 |
| 5.School leaders' well-being | | | | | | 0.348 |
| 6. School leaders' inclusion | | | | | | |

*$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$.

upper secondary vocational, only a correlation between students' perception on well-being and teachers' perception on well-being ($r = 635$, $p < 0.05$) emerged.

## DISCUSSION AND CONCLUSION

The findings of this study highlight the strong relationship between students' well-being and inclusion through technologies perceived by students, teachers, and school leaders in schools' policies and practices. In addition, depending on the education level, the perceptions of the three actors emerge as different and correlated.

A number of studies in the literature highlight a strong relationship between well-being and inclusion in schools (e.g., NSW Department of Education and Communities, 2015), including through technologies (Carretero et al., 2017; OECD, 2019). To the best of our knowledge, however, few works explore the relationship among the perceptions of students, teachers, and school leaders on students' well-being and inclusion through technologies specifically in schools' policies and practices in order to effectively meet the needs of students.

This study attempted to identify how students, teachers, and school leaders perceive students' well-being and inclusion through technologies in schools' policies and practices and has highlighted some interesting insights in relation to the formulated hypothesis.

In particular, concerning the perception of students, teachers, and school leaders about the relationship between well-being and inclusion in the policies and practices of their school (*H1*), the findings indicate that (i) the different components of well-being (*relationships*, *school community*, *safety*) and of inclusion (*individual learning needs*, *active learning*, and *collaboration*) are perceived by all the three actors as associated; in particular, the highest correlations emerged in school leaders' perceptions. (ii) Well-being and inclusion are perceived by all the three actors as two separated but strongly correlated factors, mostly by school leaders. These findings confirm Hypothesis 1 and are in line with consolidated international frameworks where inclusion is seen as a different but interrelated factor contributing to well-being in the school context (NSW Department of Education and Communities, 2015; OECD, 2019). The fact that the strongest correlation between components of well-being and inclusion emerged in school leaders' perceptions could be explained either by the systemic and holistic point of view that characterize school leaders' and by their commitment to ensure compliance with mandated education policy (Tosh and Doss, 2020).

Depending on the education level, students, teachers, and school leaders perceive students' well-being and inclusion through technologies differently (*H2*). Specifically, the most positive perception on well-being was expressed by students in lower secondary, followed by students in primary, in upper secondary vocational, and, lastly, in upper secondary general (e.g., licei and technical schools). On the other hand, the most positive perception on inclusion was expressed by students in primary school, followed by students in lower secondary and upper secondary vocational. Once again, the most negative perception was expressed by students in upper secondary general. These findings are in line with recent PISA results (OECD, 2019), indicating that around one-third of 15-year-old students were not satisfied with their lives (both in and out of school). Concerning teachers, few differences emerged at various school levels. The most negative perception regarding both inclusion and well-being emerge from primary school teachers. This finding is line with the literature, which shows primary teachers' need for collaboration, shared responsibilities, common planning time, and professional development to foster inclusion and well-being through technologies (Sam et al., 2015; Peacock, 2016; Manrique et al., 2019). Regarding school leaders, no significant differences emerge at different education levels. These findings partially confirm Hypothesis 2. As argued by Castaño-Muñoz et al. (2018), school leaders' perceptions are unlikely to vary significantly among education levels as they reflect the policy and strategy dimension of the school context (Tuytens and Devos, 2010). By contrast, students' and teachers' perceptions reflect classroom practices, which are most likely affected by factors related to the education level.

Concerning our third hypothesis, within schools, a relationship between the perception of students, teachers, and school leaders on student's well-being and inclusion through technologies emerged at different education levels (*H3*). The findings highlight that (i) teachers and school leaders' perceptions are strongly correlated at all education levels, except for iVET; this could result from the limited size of the sample of iVET schools involved in the study. (ii) Students' perceptions most closely correlate with teachers' perceptions with respect to school leaders' perceptions at all education levels, except for iVET. These findings confirm Hypothesis 3 and are in line with the literature that highlights relationships between students, teachers, and school leaders' perception (e.g., Castaño-Muñoz et al., 2018) on students' well-being and inclusion through technologies (Walker and Logan, 2009; Soutter et al., 2014; Ovbiagbonhia et al., 2019). In particular, the strong relationships that emerged between students and teachers, and between teachers and school leaders highlight the importance of considering the perceptions of all the three main actors in the school community to promote well-being and inclusion (Kampylis et al., 2015; OECD, 2019). This is in line with the evidence that school climate is as an important factor to be considered to improve engagement in school activities, but it is effective only when its influence can modify the well-being experience of the students (Lombardi et al., 2019).

A major limitation of the present research study regards the low degree of reliability and validity of the SELFIE tool and process for the specific purpose of assessing the school community's (students', teachers', and school leaders') perceptions of how digital technology use affects students' well-being and inclusion. Indeed, while an adequate level of convergent validity was returned for all three participant groups, the internal reliability of both well-being and inclusion factors was insufficient for the student group, as was the well-being factor related to the teachers' group. Likewise, the discriminant validity level was inadequate across the board, i.e., for both factors in relation to all the user groups. These results can be attributed to two main reasons: (i) as a tool expressly designed to support schools' self-evaluation of the use of digital technologies for teaching and learning *per se*, SELFIE's focus on well-being and

inclusion through technology is quite limited in scope; (ii) this aligns epistemologically with the theoretical frameworks (NSW Department of Education and Communities, 2015; OECD, 2019) underpinning the two-factor CFA models, which in fact treat well-being and inclusion as distinct but tightly intertwined factors.

Those endeavoring to repeat the research study reported here are advised to include external variables as a means of validating our theoretical assumption, mentioned above, that well-being and inclusion are distinct factors that in fact are tightly intertwined. In addition, they could look in greater detail on SELFIE's limited scope for gauging school-wide perceptions on the relations between digital technology use and learners' well-being and inclusion. One possible step in this direction would be to consider employing other tools as well, such as PISA (OECD, 2019) for well-being and the Index for Inclusion (Booth and Ainscow, 2011) for inclusion. Another constructive step could be to engage other stakeholder groups—especially parents—in the SELFIE process, especially to increase the tool's scope to address digital technology use relations with well-being and inclusion. Clearly, systematic field studies are required to gain a firmer grasp of how digital technologies can be employed for positive, school-level practices that support well-being and inclusion.

In conclusion, the findings presented here highlight significant factors regarding the relations between digital technology use and students' well-being and inclusion, factors that ought to be addressed so as to satisfy the needs of students in an effective manner.

Practical implication to improve schools' digital policies and practices can be drawn from our findings. In particular, the strong relationship between students and teachers' perceptions can guide teachers' cooperation, by exchanging ideas and sharing best practices on how to provide extra help or giving students opportunities to express their ideas in relation to well-being and inclusion through technologies. In addition, the relationship of school leaders' perceptions with both teachers and students can contribute to design consistent schools' policies, building trusting relationships with teachers and students, and to offer enriching activities for effectively responding to students' needs. This is in line with recent empirical studies on well-being (Lombardi et al., 2019) that emphasize the importance of integrating socio-emotional development in schools' daily practice (Burns and Gottschalk, 2019). Such type of school-level interventions, which have emerged to be more successful in some European countries including Italy (Govorova et al., 2020), encompasses key components of well-being and inclusion through technologies such as *relationships*, *school community*, *safety*, *individual learning needs*, *active learning*, and *collaboration*, explored in the SELFIE self-reflection process (Kampylis et al., 2016, 2019). Similarly, research currently being performed in the EC ERASMUS+SHERPA project (ErasmusPlus, 2020) includes the design and testing of a pedagogical innovation toolkit that complements and operationalizes schools' SELFIE self-reflection on digital technology use. The kit will scaffold individual school's efforts to formulate sound policies and practices in this area, and in doing so will provide adequate scope for addressing aspects related to students' well-being and inclusion. The focus on the school-wide environment will help in identifying variables in the school climate that impact on students' well-being and inclusion, thereby also enriching the SELFIE tool and process itself.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

Ethical review and approval was not required for this study involving human participants in accordance with the local legislation and institutional requirements. All answers provided through SELFIE were anonymous. Individual students, teachers, school leaders or other staff members replying to the questions and statements were not identified personally.

## AUTHOR CONTRIBUTIONS

SP contributed to the organization and data collection, performed statistical analyses, participated in the interpretation of results, and contributed to writing the manuscript, in particular the sections Materials and Methods, Results, and Discussion and Conclusion. SB has organized and supervised data collection, conducted the literature review, participated in interpretation of the results, and contributed to writing the manuscript, in particular the sections Introduction and Discussion and Conclusion. LF has contributed to the literature review and provided feedback on the manuscript. All authors read and approved the submitted version.

## FUNDING

Financial support for carrying out the SELFIE pilot in Italy was obtained through grant No. 930158-2016 A08-IT of 05/08/2016 (EC - JRC B.4). The present work was supported by grant No. 612867-EPP-1-2019-1-EL-EPPKA3-PI-FORWARD from the EC EACEA Agency (SHERPA).

## ACKNOWLEDGMENTS

The authors thank all the school leaders, teachers, and students who participated in the SELFIE pilot study in Italy. Special thanks go to Dr. Panagiotis Kampylis and Dr. Yves Punie of the European Commission's Joint Research Centre for their extraordinary commitment and support during this study. We are grateful to Prof. Sergio Morra for his valuable support and to our CNR-ITD colleagues Jeffrey Earp and Marcello Passarelli for their feedback. Finally, the authors also would like to express their gratitude to the referees, who offered extremely valuable suggestions for improvements in the earlier version of this study.

## REFERENCES


Ainscow, M. (2005). Understanding the development of inclusive education system. *Rev. Electron. Investig. Psicoeduc. Psigopedag.* 3, 15–20.

Ainscow, M., Dyson, A., Hopwood, L., and Thomson, S. (2016). *Primary Schools Responding to Diversity: Barriers And Possibilities*. Abingdon: Cambridge Primary Review Trust.

Anderson, D. H., and Graham, A. P. (2016). Improving student well-being: having a say at school. *Sch. Eff. Sch. Improv.* 27, 348–366. doi: 10.1080/09243453.2015.1084336

Angel, C. M., Woldetsadik, M. A., McDaniel, J. T., Armstrong, N. J., Young, B. B., Linsner, R. K., et al. (2019). Confirmatory factor analysis of the enriched life scale among US military veterans. *Front. Psychol.* 10:2181. doi: 10.3389/fpsyg.2019.02181

Beetham, H. (2015). *Thriving in a Connected Age: Digital Capability And Digital Well-Being*. Bristol: JISC.

Benigno, V., Caruso, G., Fante, C., Ravicchio, F., and Trentin, G. (2018). *Classi Ibride E Inclusione Socio-Educativa: Il Progetto TRIS*. Milan: Franco Angeli.

Benigno, V., Ferlino, L., and Trentin, G. (2019). *Tecnologie abilitanti. Un Nuovo Paradigma per l'inclusione. I quaderni Pearson Academy Serie su Inclusione a 360°. Equitaĭ e Valorizzazione Dei Talenti*. London: Pearson.

Bentler, P. M. (2006). *EQS 6 Structural Equations Program Manual*. Temple City, CA: Multivariate Software Inc.

Bocconi, S., and Ott, M. (2013). "ICT and universal access to education: towards a culture of accessibility," in *Information Systems E-Learning and Knowledge Management Research. WSKS 2011. Communications in Computer and Information Science*, eds M. D. Lytras, D. Ruan, R. D. Tennyson, and P. Ordonez, (Cham: Springer), 330–337. doi: 10.1007/978-3-642-35879-1_39

Bocconi, S., and Ott, M. (2014). "Bridging the concepts of educational software and assistive technology," in *Educational Technology Use and Design for Improved Learning Opportunities*, ed. M. Khosrow-Pour, (Pennsylvania: IGI Global), 185–202. doi: 10.4018/978-1-4666-6102-8.ch010

Bocconi, S., and Panesi, S. (2018a). *SELFIE Pilot Country Report*. Brussels: European Commission.

Bocconi, S., and Panesi, S. (2018b). "Teachers' professional learning and competence in the digital era: the DigCompEdu framework," in *Teacher Education; Training on ICT Between Europe and Latin America*, eds M. Ranieri, L. Menichetti, and M. K. Borges, (Rome: Aracne), 39–48.

Bocconi, S., Panesi, S., and Kampylis, P. (in press). Fostering the digital competence of schools: piloting SELFIE in the Italian education context. *IEEE-RITA*.

Bollen, K. A., and Lennox, R. (1991). Conventional wisdom on measurement: a structural equation perspective. *Psychol. Bull.* 110, 305–314. doi: 10.1037/0033-2909.110.2.305

Booth, T., and Ainscow, M. (2011). *Index for Inclusion. Developing Learning and Participation in Schools*. Bristol: Centre for Studies on Inclusive Education.

Borgonovi, F., and Pál, J. (2016). *A Framework for the Analysis of Student Well-Being in the PISA 2015 Study: Being 15 In 2015*. OECD Education Working Papers, No. 140. Paris: OECD Publishing.

Broek, S., and Buiskool, B. J. (2020). *Adapting the SELFIE tool for work-based learning systems in Vocational Education and Training*, (Report No. EUR 30079 EN), Luxembourg: Publications Office of the European Union.

Burns, T., and Gottschalk, F. (eds) (2019). *Educating 21st Century Children: Emotional Well-being in the Digital Age*. Paris: OECD.

Caena, F., and Redecker, C. (2019). Aligning teacher competence frameworks to 21st century challenges: the case for the European Digital Competence Framework for Educators (Digcompedu). *Eur. J. Educ.* 54, 356–369. doi: 10.1111/ejed.12345

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56, 81–105. doi: 10.1037/h0046016

Carretero, S., Vuorikari, R., and Punie, Y. (2017). *DigComp 2.1: The Digital Competence Framework for Citizens with Eight Proficiency Levels And Examples Of Use*. Report No. EUR 28558 EN. Luxembourg: Publications Office of the European Union.

Caruso, G., and Ferlino, L. (2018). "The proximity of distance education," in *INNODOCT/18. International Conference on Innovation Documentation and Education*, eds S. Estelles Miguel, J. O. Montesa Andrés, J. V. Oltra Gutiérrez, and M. J. Verdecho Sáez, (València: Editorial Universitat Politècnica de València), 45–67.

Castaño-Muñoz, J., Costa, P., Hippe, R., and Kampylis, P. (2018). "Within-school differences in the views on the use of digital technologies in Europe: evidence from the SELFIE tool," in *EDULEARN18 Proceedings*, eds L. Gómez Chova, A. López Martínez, and I. Candel Torres, (Valencia: IATED Academy), 10417–10426.

Coombes, L., Appleton, J., Allen, D., and Yerrell, P. (2013). Emotional Health and Well-being in Schools: involving young people. *Child. Soc.* 27, 220–232. doi: 10.1111/j.1099-0860.2011.00401.x

D'Alessio, S. (2011). *Inclusive Education in Italy. A Critical Analysis of the Policy of Integrazione Scolastica*. Boston, MA: Sense Publishers.

de Róiste, A., Kelly, C., Molcho, M., Gavin, A., and Gabhainn, S. (2012). Is school participation good for children? Associations with health and well-being. *Health Educ.* 112, 88–104. doi: 10.1108/09654281211203394

Directorate-General for Education, Youth, Sport, and Culture, (2019). *Assessment of Tools And Deliverables Under The Framework For European Cooperation In Education And Training (ET2020)*. Final report. Luxembourg: Publications Office of the European Union.

ErasmusPlus, (2020). *SELFIE HElpeR & Pedagogical innovation Assistant (SHERPA)*. Brussels: European Commission. Available online at: https://ec.europa.eu/programmes/erasmus-plus/projects/eplus-project-details/#project/612867-EPP-1-2019-1-EL-EPPKA3-PI-FORWARD (accessed July 3, 2020).

European Commission, (2018). *Communication from the Commission to the European Parliament the Council the European Economic and Social Committee and the Committee of the Regions on the Digital Education Action Plan* (COM(2018) 22 final). Luxembourg: Publications Office of the European Union.

European Commission, (2020). *Shaping Europe's Digital Future*. Luxembourg: Publications Office of the European Union.

Farrell, A. M. (2010). Insufficient discriminant validity: a comment on Bove, Pervan, Beatty, and Shiu (2009). *J. Mark. Res.* 63, 324–327. doi: 10.1016/j.jbusres.2009.05.003

Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 18, 39–50. doi: 10.1177/002224378101800104

Govorova, E., Benitez, I., and Muniz, J. (2020). How schools affect student well-being: a cross-cultural approach in 35 OECD countries. *Front. Psychol.* 11:431. doi: 10.3389/fpsyg.2020.00431

Gui, M., Fasoli, M., and Carradore, R. (2017). Digital well-being. Developing a new theoretical tool for media literacy research. *IJSE* 9, 155–173. doi: 10.14658/pupj-ijse-2017-1-8

Gui, M., Gerosa, T., Garavaglia, A., Petti, L., and Fasoli, M. (2018). *Digital Well-being. Validation of a Digital Media Education Proposal in High Schools. "Benessere Digitale" Research Centre*. Available at: http://www.benesseredigitale.eu/wp-content/uploads/2019/02/Digital_wellbeing_schools_Research_Report.pdf (accessed July 3, 2020).

Henseler, J., Ringle, C. M., and Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* 43, 115–135. doi: 10.1007/s11747-014-0403-8

Ibrahim, S. N. H., Suan, C. L., and Karatepe, O. M. (2019). The effects of supervisor support and self-efficacy on call center employees' work engagement and quitting intensions. *Int. J. Manpw.* 40, 688–703. doi: 10.1108/ijm-12-2017-0320

Jeladze, E., and Pata, K. (2018). The model of self-organization in digitally enhanced schools. *IxD A* 39, 61–77.

Jones, A., and Sandford, R. (2019). "Safety. Protecting health and well-being," in *Design Thinking for Digital Well-being: Theory and Practice for Educators*, eds F. C. Chambers, A. Jones, O. Murphy, and R. Sandford, (Abingdon: Routledge), 170–186.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34, 183–202. doi: 10.1007/BF02289343

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., and Rosseel, Y. (2019). *Semtools: Useful Tools For Structural Equation Modeling. R Package Version 0.5-2*. Available at: https://CRAN.R-project.org/package=semTools (accessed July 3, 2020).

Kampylis, P., Devine, J., Punie, Y., and Newman, T. (2016). "Supporting schools to go digital: from conceptual model towards the design of a self-assessment tool

for digital-age learning," in *ICERI Proceedings*, eds L. G. Chova, A. L. Martínez, and I. Candel Torres, (València: IATED Digital Library), 816–825.

Kampylis, P., Hodson, D., Petkova, S., Hippe, R., Cachia, R., Sala, A., et al. (2019). *SELFIE Forum - Teaching and Learning in the Digital Age*. Report No. JRC117482. Luxembourg: Publications Office of the European Union.

Kampylis, P., Punie, P., and Devine, J. (2015). *Promoting Effective Digital-Age Learning* (EUR 27599). Luxembourg: Publications Office of the European Union.

Kline, R. B. (2005). *Methodology in the Social Sciences. Principles and Practice of Structural Equation Modeling, 2nd edn*. Guilford Press.

Könings, K. D., Seidel, T., Brand-Gruwel, S., and van Merriënboer, J. J. (2014). Differences between students' and teachers' perceptions of education: profiles to describe congruence and friction. *Instr. Sci.* 42, 11–30. doi: 10.1007/s11251-013-9294-1

Kuurme, T., and Carlsson, A. (2010). The Factors of Well-Being in Schools as a Living Environment According to Students' Evaluation. *J. Teach. Educ. Sustain.* 12, 70–88. doi: 10.2478/v10099-009-0055-9

Littlecott, H. J., Moore, G. F., and Murphy, S. M. (2018). Student health and well-being in secondary schools: the role of school support staff alongside teaching staff. *Pastor Care Educ.* 36, 297–312. doi: 10.1080/02643944.2018.1528624

Løhre, A., Lydersen, S., and Vatten, L. (2010). School well-being among children in grades 1 - 10. *BMC Public Health* 10:526. doi: 10.1186/1471-2458-10-526

Lombardi, E., Traficante, D., Bettoni, R., Offredi, I., Giorgetti, M., and Vernice, M. (2019). The impact of school climate on well-being experience and school engagement: a study with high-school students. *Front. Psychol.* 10:2482. doi: 10.3389/fpsyg.2019.02482

Maelan, E., Tjomsland, H., Baklien, B., Samdal, O., and Thurston, M. (2018). Supporting pupils' mental health through everyday practices: a qualitative study of teachers and head teachers. *Pastor Care Educ.* 36, 1–13. doi: 10.1080/02643944.2017.1422005

Malhotra, N. K., and Dash, S. (2011). *Marketing Research an Applied Orientation*. London: Pearson Publishing.

Manrique, A. N., Dirani, E., Frere, A. F., Moreira, G. E., and Arezes, P. M. (2019). Teachers' perceptions on inclusion in basic school. *Int. J. Manag. Rev.* 33, 409–419. doi: 10.1108/IJEM-02-2018-0058

Mascia, M. L., Agus, M., and Penna, M. P. (2020). Emotional intelligence, self-regulation, smartphone addiction: which relationship with student well-being and quality of life? *Front. Psychol.* 11:375. doi: 10.3389/fpsyg.2020.00375

Mattar, J., Piovezan, M., Souza, S., Santos, C., and Santos, A. (2020). Critical presentation of the european digital competence framework (DigComp) and related frameworks. *Res. Soc. Dev.* 9:e172943062. doi: 10.33448/rsd-v9i4.3062

McDonald, R. P. (1970). The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *Br. J. Math. Stat. Psychol.* 23, 1–21. doi: 10.1111/j.2044-8317.1970.tb00432.x

McDonald, R. P. (1999). *Test Theory: A Unified Treatment*. New Jersey: Lawrence Erlbaum Associates.

Messiou, K. (2017). Research in the field of inclusive education: time for a rethink? *Int. J. Incl. Educ.* 21, 146–159. doi: 10.1080/13603116.2016.1223184

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moore, G. F., Littlecott, H. J., Evans, R., Murphy, S., Hewitt, G., and Fletcher, A. (2017). School composition, school culture and socioeconomic inequalities in young people's health: multi-level analysis of the health behaviour in school-aged children (HBSC) survey in wales. *Br. Educ. Res. J.* 43, 310–329. doi: 10.1002/berj.3265

NSW Department of Education and Communities, (2015). *The Well-Being Framework for Schools*. Sydney: NSW Governament.

Obiakor, F., Harris, M., Mutua, K., Rotatori, A., and Algozzine, B. (2012). Making inclusion work in general education classrooms. *Educ. Treat. Child.* 35, 477–490. doi: 10.1353/etc.2012.0020

OECD, (2019). *PISA 2018 Results (Volume III): What School Life Means for Students' Lives*. Paris: OECD Publishing.

O'Leary, R. (2018). *Digital Schools of distinction Feasibility Report. Digital Schools of Europe*. Available at: http://www.digitalschoolseurope.eu/wp-content/uploads/2018/10/Feasibility-Report.pdf (accessed July 3, 2020).

Opertti, R. (2015). Inclusive education in Italy: a response to Anastasiou, Kauffman and Di Nuovo. *Eur. J. Spec. Needs Educ.* 30, 452–453. doi: 10.1080/08856257.2015.1060077

Ovbiagbonhia, A. R., Kolloffel, B., and den Brok, P. (2019). Educating for innovation: students' perceptions of the learning environment and of their own innovation competence. *Learn. Environ. Res.* 22, 387–407. doi: 10.1007/s10984-019-09280-3

Panesi, S., and Morra, S. (2020). Executive functions and mental attentional capacity in preschoolers. *J. Cogn. Dev.* 21, 72–91. doi: 10.1080/15248372.2019.1685525

Peacock, D. (2016). *Teacher Perceptions of the Effectiveness of Inclusion in Elementary Classroom Settings*, Doctoral thesis, Walden University, Minneapolis, MI.

Pollard, E. L., and Lee, P. D. (2003). Child well-being: a systematic review of the literature. *Soc. Indic. Res.* 61, 59–78. doi: 10.1023/A:1021284215801

Redecker, C. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu*. Report No. EUR 28775 EN. Ispra: JRC.

Rosseel, Y. (2012). lavaan: an r package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.18637/jss.v048.i02

Sam, K. L. S., Ho, F., and Lam, C. S. C. (2015). Perception of teachers towards inclusive education (PTIE): an investigation on scale construct and item calibration. *Int. J. Cult. Stud.* 1, 54–62. doi: 10.18178/ijch.2015.1.1.010

Sarti, D., Bettoni, R., Offredi, I., Tironi, M., Lombardi, E., Traficante, D., et al. (2019). Tell me a story: socio-emotional functioning, well-being and problematic smartphone use in adolescents with specific Learning Disabilities. *Front. Psychol.* 10:2369. doi: 10.3389/fpsyg.2019.02369

Schermelleh-Engel, K., Moosbrugger, H., and Muller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *MPR Online* 8, 23–74.

Shaakumeni, S. N., and Csapoì, B. (2019). Exploring the factorial validity of the beliefs about nature of science questionnaire. *Sci. Educ. Intern.* 30, 38–44. doi: 10.33828/sei.v30.i1.5

Slee, R. (2018). *Defining the Scope Of Inclusive Education: Think Piece Prepared For The 2020 Global Education Monitoring Report Inclusion and Education*. Report No. ED/GEMR/MRT/2018/T1/1. Paris: UNESCO. Available online at: https://unesdoc.unesco.org/ark:/48223/pf0000265773 (accessed July 3, 2020).

Soutter, A. K., O'Steen, B., and Gilmore, A. (2014). The student well-being model: a conceptual framework for the development of student well- being indicators. *J. Youth Adolesc.* 19, 496–520. doi: 10.1080/02673843.2012.754362

Spratt, J., and Florian, L. (2015). Inclusive pedagogy: from learning to action. Supporting each individual in the context of 'everybody'. *Teach. Educ.* 49, 89–96. doi: 10.1016/j.tate.2015.03.006

Suldo, M. S., Friedrich, A. A., White, T., Farmer, J., Minch, D., and Michalowski, J. (2009). Teacher support and adolescents' subjective well-being: a mixed-methods investigation. *Sch. Psychol. Rev.* 38, 67–85. doi: 10.1080/02796015.2009.12087850

Szûcs, A. (2019). Digital education: from hype and disappointment to change. *Opus Educ.* 5, 386–396.

Tanhua-Piiroinen, E., and Viteli, J. (2017). "Opeka and Ropeka, the self-assessing services for teachers and principals," in *Lecture Notes in Computer Science: Vol. 10474. Data Driven Approaches in Digital Education*, eds E. Lavoué, H. Drachsler, K. Verbert, J. Broisin, and M. P. érez-Sanagustín, (Cham: Springer), 602–605. doi: 10.1007/978-3-319-66610-5_72

Tosh, K., and Doss, C. D. (2020). *Perceptions of School Leadership: Implications for Principal Effectiveness*. Report No. RR-2575/5-1-BMGF. Santa Monica, CA: RAND Corporation.

Trentin, G. (2017). Connettività, Spazi Ibridi e Always-on Education. *Riv. AEIT* 5, 14–22.

Trentin, G. (2019). "Tecnologie e inclusione: come far di necessità virtù," in *Tecnologie Per L'educazione*, eds P. C. Rivoltella, and P. G. Rossi, (London: Pearson), 57–68.

Tuytens, M., and Devos, G. (2010). The influence of school leadership on teachers' perception of teacher evaluation policy. *Educ. Stud.* 36, 521–536. doi: 10.1080/03055691003729054

van Petegem, K., Aelterman, A., Rosseel, Y., and Creemers, B. (2007). Student perception as moderator for student well-being. *Soc. Indic. Res.* 83, 447–463. doi: 10.1007/s11205-006-9055-5

Viladrich, C., Angulo-Brunet, A., and Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Ann. Psicol.* 33, 755–782.

Walker, L., and Logan, A. (2009). *Using Digital Technologies To Promote Inclusive Practices In Education*. Kuala Lumpur: FutureLab.

Welsh Government, (2020). *Health and Well-being. Guidance to Help Schools And Settings Develop Their Own Curriculum Enabling Learners To Develop Towards The Four Purposes*. Cardiff: Welsh Governament.

Zanobini, M. (2013). Some considerations about inclusion, disability, and special educational needs: a reply to Giangreco, Doyle and Suter (2012). *Life Span Disabil.* 16, 83–94.

Zanobini, M., Viterbori, P., Garello, V., and Camba, P. (2017). Parental satisfaction with disabled children's school inclusion in Italy. *Eur. J. Spec. Needs Educ.* 33, 597–614. doi: 10.1080/08856257.2017.1386318

Zhang, Z., and Yuan, K. (2016). Robust coefficients alpha and omega and confidence intervals with outlying observations and missing data: methods and Software. *Educ. Psychol. Measur.* 76, 387–411. doi: 10.1177/0013164415594658

**Disclaimer:** The views expressed in this article are purely those of the authors and should not be regarded as the official position of the European Commission.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 15 September 2020
doi: 10.3389/fpsyg.2020.549653

# Teacher Learning in Difficult Times: Examining Foreign Language Teachers' Cognitions About Online Teaching to Tide Over COVID-19

*Lori Xingzhen Gao*[1†] *and Lawrence Jun Zhang*[1,2*†]

[1] College of Foreign Languages, Taiyuan University of Technology, Taiyuan, China, [2] Faculty of Education and Social Work, University of Auckland, Auckland, New Zealand

The sudden global outbreak of COVID-19 in late 2019 has led to thriving online teaching, including the teaching of languages, across the world. As the online teaching of English-as-a-foreign-language (EFL) in Chinese universities is facing new challenges, EFL teachers have been positively exploring new solutions. To understand how EFL teachers were coping with the challenges, we set up this research as part of a larger study to examine EFL teachers' cognitions about online teaching in response to the disruption of normal teaching plans. We did so by taking a qualitative approach through analyzing in-depth interviews with three EFL teachers from a Chinese university. Through thematic analysis we found that teachers had clear cognitions about features, advantages, and constraints of online EFL teaching and that they acquired information and communication technology (ICT) literacy through understanding students' learning needs, online teaching practice, and the necessity of integrating traditional classroom teaching methods into online delivery. We conclude this study with a discussion on its pedagogical implications for similar contexts or colleagues facing similar challenges in other parts of the world.

**Keywords: COVID-19, language teacher cognition, online EFL teaching, information technology literacy (ITC), foreign language teaching and learning, China**

OPEN ACCESS

***Edited by:***
*Jie Chi Yang,*
*National Central University, Taiwan*

***Reviewed by:***
*Ricky Lam,*
*Hong Kong Baptist University, Hong Kong*
*Tiefu Shaun Zhang,*
*University of Electronic Science and Technology of China, China*

****Correspondence:***
*Lawrence J. Zhang*
*lj.zhang@auckland.ac.nz*

***†ORCID:***
*Lori X. Gao*
*orcid.org/0000-0001-9328-7343*
*Lawrence J. Zhang*
*orcid.org/0000-0003-1025-1746*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *07 April 2020*
***Accepted:*** *24 August 2020*
***Published:*** *15 September 2020*

***Citation:***
*Gao LX and Zhang LJ (2020) Teacher Learning in Difficult Times: Examining Foreign Language Teachers' Cognitions About Online Teaching to Tide Over COVID-19. Front. Psychol. 11:549653. doi: 10.3389/fpsyg.2020.549653*

## INTRODUCTION

The real-time coronavirus map updated in the early hours of March 28 on Google[1] showed that the total number of confirmed cases worldwide reached 587,958, of which the United States topped the list with 104,011 cases and that European countries such as Italy (86,498), Spain (65,719), Germany (50,871), France (32,964), and the United Kingdom (14,590) etc. were severely afflicted. In Asia, China (81,394) and South Korea (9,332) ranked high in the disease-hit countries. Statistics of confirmed cases in the Middle East, Australasia, and Africa were not optimistic, either. Since late 2019, COVID-19 has swept over the whole globe as a worldwide concern and brought havoc to people's health and life. Facing the high-risk challenge, to minimize the adverse effect of the epidemic, many countries have activated emergency response and implemented a wide range of measures, of which home-based quarantine and physical distancing are basic ones.

[1]https://google.org/crisisresponse/covid19-map

The epidemic situation of COVID-19 in China started in late December 2019 and worsened quickly in the first 3 months of 2020. The number of confirmed cases of COVID-19 increased rapidly. To control the sources of infection, cut off the channels of transmission, and make every possible effort to curb the spread of the disease, China's central government enforced the policy of strict home-based quarantine and physical distancing among her people across towns and cities, while life-saving battles in hospitals and academic research into drug and vaccine development were going on at the same time. To protect the young generation from being affected by the coronavirus, the Ministry of Education of China called for online teaching and learning among teachers and students at all levels immediately afterward. Educational institutions were temporarily closed down. Online teaching therefore substituted the traditional way of classroom teaching and became the mainstream mode of delivering teaching.

For language teachers in Chinese universities, their teaching plans were disrupted and their knowledge and skills of ICT literacy were challenged. Due to the life-threatening global pandemic, language teachers, as well as teachers of other disciplines, had to move instruction online. In this process, the change in teachers' cognitions about education and language teaching must have changed substantially. How to organize efficient activities via online teaching? How would students respond to online delivery, especially when the subject matter was about learning a foreign language? None of the teachers was sure about the effectiveness of such large-scale online language instruction. Worries and stress lingered on the teachers' minds. How did EFL teachers in Chinese universities perceive and respond to their disrupted teaching when online teaching became the main mode of delivery? It is a question that needs urgent investigation as the teachers' perceptions and responses to online teaching over COVID-19 greatly influence the quality of language education in Chinese universities. As part of a larger study, this paper presents a qualitative study, through in-depth interviews, on three EFL teachers' cognitions about online teaching and their responses to, and strategies for, coping with their disrupted teaching. Unfortunately, little has been reported on how EFL teachers have responded to such a drastically challenging and evolving teaching and learning environment (cf. Fu and Zhou, 2020). This study was set up to fill the research gap. Our findings are expected to help language teaching professionals to understand how they themselves as frontline teachers can benefit from learning about how their colleagues were learning to cope with the new challenges by trying to understand online teaching platforms, online language teaching methods, online class management, among many other things, through scrutiny of three teachers' cognitions and personal experiences.

# LITERATURE REVIEW

## Language Teacher Cognition

Research on teacher cognition began in the 1970s, thrived in the mid-1990s, and became a major area of enquiry in the field of foreign language education (Borg, 2006; Gao, 2019). The last 30 years has witnessed prolific research on language teacher cognition from a relatively new and undeveloped area into an important field of inquiry (Borg, 2019; Sun and Zhang, 2019; Li, 2020). The foci of language teacher cognition research include conceptual understandings of teacher cognition, factors affecting the development of teacher cognition, and the relationships between teachers' cognition and their actual classroom instructional practice, among other things (Zhang and Rahimi, 2014; Gao, 2019).

According to Borg (2015, 2019), the concept of teacher cognition is complex, with about 60 distinctive terms used in language teacher cognition research to refer to it. Although the terms diverge in various contexts, the features can be classified into the following categories: Personal nature, mental lives, life or learning experiences, and the interactive influence of cognitive processes and instructional practice (Borg, 2003). Teacher cognition may involve teachers' beliefs, knowledge, theories, attitudes, images, assumptions, metaphors, and conceptions about teaching, teachers, learning, students, subject matters, curricula, materials, instructional activities, among other things (Borg, 2003). In fact, teachers hold a wide range of practical theories which inform how they behave and teach in classrooms. Moreover, teachers' practical teaching experiences contribute to the development of their cognitions (Kagan, 1992; Gao and Ma, 2011). The existing body of research shows that teacher cognition is defined, partly, by personal factors based on teachers' own understanding of the practical classroom instructional activities (Clandinin and Connelly, 1987; Sun and Zhang, 2019). The body of research has a further, widely accepted implication: To better understand the process of teaching, both actions and cognitions underlying every decision that teachers make in their pedagogical practice need to be described and taken into consideration.

The generally accepted conclusion about factors contributing to language teacher cognition is that it derives from teachers' own learning experiences as language learners within formal classrooms (Pajares, 1992; Ellis, 2006; Nunan and Richards, 2015), and through "apprenticeship of observation" (Borg, 2003) in early teaching experiences and teacher training courses (Popko, 2005), which significantly influence the way in which they view and approach teaching (Sun, 2017; Sun and Zhang, 2019). As Richards and Lockhart (1996, p. 30) claimed, other sources may include "teachers' personality factors, educational principles and research-based evidence," different stages the teachers are in Berliner (1994), teachers' emotions (Zembylas, 2005), and language policy (Farrell and Kun, 2007). In short, many sources contribute to the formation of teacher cognition. Borg's (2015) model intends to capture the potential factors affecting language teacher cognition, which is presented below (see **Figure 1**).

As shown in **Figure 1**, the three categories that influence teacher cognition are schooling, professional coursework, and contextual factors (classroom practice). For schooling, teachers' personal history and specific experience in classrooms help define their early cognitions and shape their understanding of teachers and teaching. Teachers' professional coursework also may have an effect on their existing cognitions, although its impact may

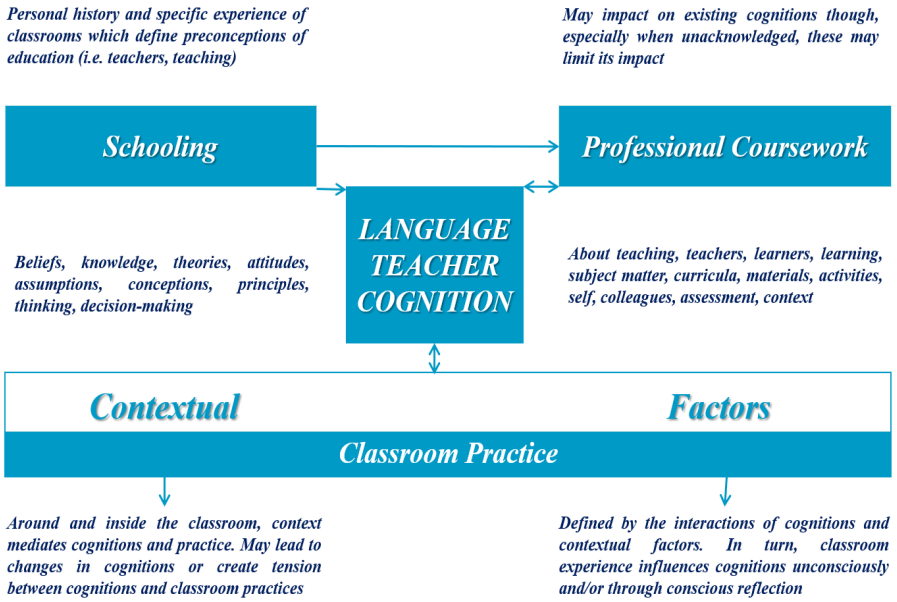


**FIGURE 1 |** Elements and processes in language teacher cognition (adapted from Borg, 2015, p. 333).

be limited because teachers are unaware of the relationship of their coursework to their practice. Contextual factors are related to, and, mediate cognition-practice relationship inside the classroom context and can cause teachers' cognitions to change or form a tension between their cognitions and instructional practice. Classroom practice is bounded by the interactions of cognitions and contextual factors, and in turn, teachers' classroom practice influences their cognitions in an unconscious and/or conscious way.

How teachers' cognitions are related to their actual classroom instructional practice is increasingly a research focus in the field of teacher education, including the cognition-practice congruence, that is, the extent to which teachers' instructional practice is consistent with their cognitions and any inconsistency between the two (Borg, 2019; Sun and Zhang, 2019). Teachers possess theoretical beliefs about teaching, and their beliefs and cognitions provide a basis for their teaching behavior (Borg, 2011; Basturkmen, 2012; Zhang and Ben Said, 2014; Farrell and Ives, 2015; Kaymakamoglu, 2018), and their beliefs guide their thought and behavior. Inconsistency between teachers' stated cognitions and their observed classroom practice has been reported in case studies of experienced and novice language teachers. Such studies were carried out with teachers in primary schools (Farrell and Lim, 2005), secondary schools (Ng and Farrell, 2003), high schools (Mitchell, 2005), and tertiary settings (Sun, 2017), and they have examined topics ranging from beliefs about, for example, teachers' roles (Anstrom, 2003), communicative language teaching (Sugiyama, 2003), teaching methods (Tucker, 2001), and students' needs (Gilliland, 2015). The inconsistency between teacher cognition and practice is also identified in the research in a number of areas (Richards, 1998) in mainstream education (Fang, 1996), literacy education (Cummins et al., 2004), second/foreign language teaching (Xiang and Borg, 2014; Sun, 2017), and the teaching of specific skills in English as a foreign language, including grammar (Graus and Coppen, 2016), writing (Zhang, 2016a,b), reading (Vaish, 2012), vocabulary (Gerami and Noordin, 2013), corrective feedback on oral communication (Zhang and Rahimi, 2014; Rahimi and Zhang, 2015), speaking (Baleghizadeh and Nasrollahi Shahri, 2014), pronunciation (Buss, 2016), and corrective feedback on pronunciation (Cooper, 2019). The degree of inconsistency, from partial congruence to clear divergence, with a low positive relationship or limited correspondence, is also reported in the research.

## Language Teachers' ICT Literacy

The benefit of technology for language teaching and learning has been highlighted in many studies; its role is not only to engage learners in the learning process, but also to promote learners' motivation and learner-centered instruction (Chapelle, 2005). The benefit has been further enhanced and expanded by virtue of technological advancements which have introduced mobile devices such as smart-phones and tablet computers to language learning contexts. Teaching and learning of languages have been enabled to transcend time and space limitations and made "more fun and interactive" (Demouy et al., 2016, p. 19). The convenience, mobility, and effectiveness of mobile learning have become evident (e.g., Evans, 2008; Liu et al., 2008). The benefits of mobile learning should be extendable to EFL contexts as well. The integration of mobile technology into EFL teaching and learning is thus feasible in ways that facilitate students' enhancing their language competence. Although mobile learning seems to be more central to learners than teachers, guidance and advice from teachers are always required by students

(Kukulska-Hulme, 2009). However, teaching with technology can be a complicated and difficult task for some teachers under the influence of social and contextual factors (Koehler et al., 2011). Therefore, before applying mobile learning, teachers need to learn and grasp adequate technological and pedagogical knowledge (Tai et al., 2015). The technological pedagogical content knowledge (TPACK) model, as adopted in our study is a framework that can help us understand how teachers think and take actions accordingly.

Based on the original work of Shulman (1986), where he proposed his conceptualization of teacher knowledge through his pedagogical content knowledge (PCK) model, the TPACK model was developed by Koehler and Mishra (2005) as a way to describe the relationships and interactions between teachers' knowledge of technology, pedagogy, and the subject matter, as shown in **Figure 2**.

As is clear in **Figure 2**, teachers' technological knowledge (TK, i.e., knowledge of how technologies should be used in the content domain), pedagogical knowledge (PK, i.e., knowledge of how contents are learned and taught), and content knowledge (CK, i.e., knowledge of the subject matter to be taught and learned) are connected to one another and lead to three new knowledge forms: PCK, technological content knowledge (TCK), and technological pedagogical knowledge (TPK). The interrelationships of these three forms of knowledge give rise to TPACK, the core concept central to the TPACK model. TPACK, as the term denotes, is characterized by the integration of technological knowledge and pedagogical knowledge into the teaching and learning of content knowledge. While the concept of TPACK gathered the attention of many education researchers, any elaboration on its relevance to specific subject domains such as language teaching and learning remains limited. Chinese university EFL teachers' information and communication technology literacy, as a form of TPACK, in online EFL teaching during COVID-19 has been scarcely explored in the field of second or foreign language education.

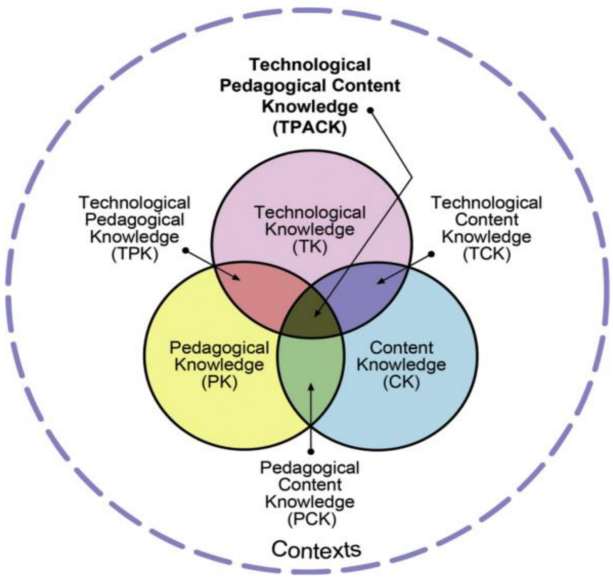


**FIGURE 2 |** TPACK model (Source: tpack.org).

## Online Teaching Over COVID-19 in China

Teachers' ICT literacy is not a new topic; instead it is an important component in the Chinese Ministry of Education's effort for modernizing education in China, as emphasized several times in various official documents. In 2016, the Ministry of Education of China issued *The 13th Five-Year Plan of Education Informationalization,* which called for the integration and development of technology and education, creation of a good environment of informationalized education, and upgrading of teaching concepts, modes and content to train talents for the information age. *The Guideline for College English Teaching* (Ministry of Education of China, 2017) pointed out that as teachers' quality and competence are key factors affecting teaching quality, university English teachers must actively upgrade their knowledge and skills in modern education technology and adapt themselves to meet the needs of university English teaching in the technology-enhanced learning and teaching environment. In the *China Education Modernization 2035*, it is proposed that education reform be carried out through the cultivation of a high-quality, professional and creative teaching contingent to quicken the mode of talent training with modern technology in the information age.

The abrupt outbreak of COVID-19 in China in late December 2019 severely disrupted people's regular schedule of life, education and work. Universities were not exempted at all. Lessons at all levels scheduled for the Spring Semester 2020 were being delivered online instead of face-to-face in the classroom. Shouldering a historical mission of education amid the COVID-19 crisis, scholars and teachers have researched online teaching recently in the field of education in recent weeks in other parts of the world. However, due to the scarcity of time and lack of facility to carry out empirical research, our literature search resulted in only three publications on Chinese EFL teachers related to online teaching over COVID-19 and in the section below we review them for the purpose of helping us to understand the challenges EFL teachers faced and their coping strategies in having to deliver online teaching.

The first paper reports a study by Fu and Zhou (2020) that investigated the opportunities and challenges of online teaching. The authors reported that the hardware facilities and wi-fi conditions are uneven across schools and areas; teachers' information technology skills and resilience cannot meet the needs of online teaching; the online resources and platforms are insufficient for online teaching; students, parents, and schools have different expectations of online teaching; the need for quality individualized education cannot be satisfied by online teaching. In view of these challenges of online teaching, the authors suggested optimizing wi-fi and hardware facilities and integrating quality online resources and platforms be seriously considered.

The second publication by Jiang et al. (2020) discussed the principles and mode of online teaching over COVID-19. The principles of online teaching over COVID-19 the authors reported include simplicity in technology, immediate feedback, interactive communication, and precision management. Their suggestions are consistent with the countermeasures and suggestions for enhancing the quality of online teaching proposed by Wang et al. (2020), which is the third paper we review below.

Wang et al. (2020) explored the advantages and rapid development of online teaching in China during COVID-19 and the benefit of the assistance from 5G technology in improving teaching and learning. They proposed a range of suggestions for improving online teaching quality and efficiency. The suggestions include creating a smart environment for home-based teaching and learning, choosing suitable online teaching modes, and enhancing interaction in online teaching. These are all useful suggestions, but unfortunately, they are not based on empirical evidence or supported by any empirical data.

The literature review shows that though language teacher cognition and teachers' information technology literacy have been profoundly and systematically researched in terms of conceptual understanding, relationships among components within each model, and application on various education levels and in various subject areas, no exploration has been carried out into language teacher cognition about online EFL teaching in China, especially language teachers' cognition about their information technology literary for online EFL teaching over COVID-19. The available research into online teaching over COVID-19, in spite of its paucity, provides insights into further research. Our study is an attempt to investigate three language teachers' cognitions about online EFL teaching and their information technology literacy in response to disrupted teaching plans in a Chinese university. To be more exact, the study addresses the following research questions:

(1) What are the language teachers' cognitions about online teaching over COVID-19?
(2) How did teachers acquire ICT literary in the initial stage of COVID-19?

# MATERIALS AND METHODS

As part of a larger study, the study reported here is a qualitative inquiry into what the language teachers think about, how they carry out and reflect on online EFL teaching over COVID-19, trying to construct meaning from their experience, feelings, and thoughts in this trying time. We carried out this study by framing it within the theory of constructivism. According to constructivism, reality is constructed by individuals interrelating with their cultural and social world: Human beings, and by necessity, including teachers, seek understanding and perceptions of the real world where they live and work, and develop subjective understanding from their own experience. As Merriam (1998) advises, qualitative researchers are interested in understanding the meanings people construct, and this was the purpose of our study as well.

## Participants

### Sampling Strategy

The participating teachers in the present research were chosen through convenience sampling due to the sudden outbreak of COVID-19 and teachers' subsequent festinate preparation of online EFL teaching. Convenience sampling is "selecting a sample based on time, money, location, availability of sites or respondents, and so on" (Merriam and Tisdell, 2016, p. 98). Through convenience sampling we contacted some EFL teachers from a university in a Northern Chinese city. They expressed their concerns, worries, and anxieties over the influence of epidemic on EFL teaching and their follow-up online instruction. After initial talks, three participants agreed to take part in our study. Their backgrounds are described in the following section.

### Participant Backgrounds

The participants in this research are three teachers of English from a university in a Northern Chinese city. They were chosen mainly because the courses they taught online represented three different types: Comprehensive English, English Listening, and Linguistic Pragmatics. At the same time, the three teachers' educational backgrounds and teaching experiences varied substantially. To protect personal information and for the convenience of data presentation, the participants are referred to as Jane, Mary, and Shirley. **Table 1** provides an overview of these teachers' demographic information:

As shown in **Table 1**, Jane, Mary and Shirley bear differences in their ages, educational backgrounds, years of teaching, courses taught, students' levels, and the online platforms they used. In terms of students' level, the numbers in **Table 1** represent their grades; namely, students in level 1 refer to freshmen, whose English proficiency is lower than level 2 (sophomores) and 3 (juniors). These differences among the participants present a diverse coverage of online EFL teaching. Their online teaching had been conducted for 6 weeks when the data were collected, during which their online teaching underwent a process of design, trial, and stabilization. The three participants are typical of EFL teachers in the university.

## Interviewing as a Research Tool

Methodologically interviews are used frequently in educational research because they are an instrument that can be used for in-depth investigation into the issue at hand (Miles et al., 2014). According to Creswell (2014), interviews can be carried in a face-to-face and one-on-one way, by telephone, in focus groups, or through the Internet (emails, for example). Interviews involve generally open-ended questions, which are few in number and designed to elicit views and opinions from the participants. Interviews are a powerful data collection instrument for qualitative research as they are useful when participants cannot be directly observed; through interviews participants can provide historical information; interviews allow the researcher to control over the line of questioning. Interviews are thus a natural and socially acceptable way of collecting qualitative data

**TABLE 1 |** Overview of participants' demographics.

| Name | Gender | Age | Educational degree | Years of teaching | Course | Students' university year level | Online platforms |
|---|---|---|---|---|---|---|---|
| Jane | Female | 39 | Master of Arts | 18 | Comprehensive English (IV) | 2 | Chaoxing, QQ |
| Mary | Female | 49 | Ph.D. in Arts | 27 | English Listening (I) | 1 | Ding Talk |
| Shirley | Female | 28 | Ph.D. in Arts | 1 | Linguistic Pragmatics | 3 | Chaoxing, QQ, Baidu Netdisk |

in various situations while focusing on diverse topics for in-depth information (e.g., Zhang, 2010). The interviewer can adopt flexible approaches and probe into newly emerging issues, and the interview protocol helps to keep the interaction on the right track, which is a systematic coverage of the domain (Rubin and Rubin, 2011). However, the disadvantages of interviews cannot be denied and neglected. Interviews provide indirect information filtered through the views of interviewees; interviews provide information in a designated place rather than a natural setting; the researcher's presence may bias responses; not all people are equally articulate and perceptive.

According to the degree of structure, the one-to-one interviews can be divided into structured interviews, unstructured interviews, and semi-structured interviews. Structured interviews and unstructured interviews are two extreme types of interviews where the interviewer either closely follows a prepared and elaborate interview schedule with little room for variation in the responses or follows the interviewee in unpredictable directions with interruptions kept to a minimum. Semi-structured interviews offer a compromise between the two extremes; with a set of pre-prepared guiding questions and prompts, the interviewer provides direction and guidance and at the same time allows the interviewee to elaborate on some issues (Dörnyei, 2007).

Interviews are suitable and appropriate for the present research as its intention is to gather information about the teachers' cognitions about online EFL teaching and their acquisition of ICT literacy in the initial stage of online EFL teaching. The semi-structured interview was adopted as the main tool to elicit verbal data for this study. The first author worked as the interviewer and asked two broad open-ended questions to direct the interview, and the participant/interviewees were encouraged to give detailed responses to the following two broad questions: (1) What do you think of online EFL teaching over COVID-19? (2) How did you acquire your ICT literacy in the initial stage of COVID-19? The semi-structured interview was used in the present research because it avoided the disadvantages of both structured interviews and unstructured interviews.

## Data Collection Procedures

The data for the present study were collected from March 25 to 28 through WeChat interviews with Jane on a one-to-one basis. Jane was interviewed three times and each interview lasted 45 min or so. Mary's and Shirley's data were in the form of written summaries and reflections on their online EFL teaching based on their responses to the two interview questions. **Table 2** below is a record of the interviews with the three participants.

As **Table 2** shows, Jane was interviewed three times via WeChat. In the first interview, Jane shared her cognitions about online EFL teaching. The data were in the form of texts and audio clips. In the second interview, Jane introduced her preparations for online EFL teaching and showed the pictures/screenshots and video clips she made prior to her online teaching. In the third interview, Jane updated the interviewer with her reflections on online EFL teaching with data in the form of texts and audio clips, etc. Mary's and Shirley's summaries and reflections on online EFL teaching were mainly texts, pictures, and screenshots. The collected data include texts, pictures/screenshots, audio clips, and video clips, which helped to ensure the trustworthiness of the present qualitative study. Besides the interview data, the author took field notes as well during and after the interviews. Having different forms of data also made possible the triangulation of data. The data collection ceased when there were sufficient data for the exploration of the participants' cognitions about online teaching and their teaching practice during the period of COVID-19.

## Data Analysis

The data analysis procedures began immediately after the first interview by both authors. Subsequent data in audio clips collected from the interviews were transcribed and processed. In addition, the notes taken during each interview were typed, read, and summarized. They were then processed through thematic analysis. Thematic analysis is a method for identifying, analyzing and reporting patterns (themes) emerging from data (Braun and Clarke, 2006; Miles et al., 2014), during which process the researcher scrutinizes the data for typical themes and concepts (Miles and Huberman, 1994; Rubin and Rubin, 2011).

During the interview transcriptions, no translation was done of the non-English data because the original language from the participant could better convey their meanings in their first language, mandarin Chinese, which is also the first language of both authors, which means that they are fully understood for meaningful analysis. This decision was made because as stated by Wang et al. (2020), "the relation between subjective experience and language is a two-way process; language is used to express meaning, but the other way round, language influences how meaning is constructed" (pp. 313–314). Much information will be lost in the course of translation because of the lack of equivalent vocabulary, syntax, idioms, and concepts between the source language and the target language (Sechrest et al., 1972). In doing the analysis we were constantly reminded of these possible shortcomings and therefore we mainly followed the original Chinese interviews or texts. For the purposes of reporting

**TABLE 2 |** Statistical record of data collection.

| Name | Time | Date | Duration | Mode | Main topic | Form of data |
|---|---|---|---|---|---|---|
| Jane | 1 | March 25 | 48 min | WeChat | Cognitions about online EFL teaching | Texts, audio clips |
| | 2 | March 26 | 47 min | WeChat | Preparations for online EFL teaching | Texts, pictures, video clips |
| | 3 | March 28 | 48 min | WeChat | Reflections on online EFL teaching | Texts, audio clips |
| Mary | Written summaries and reflections on online EFL teaching | | | | | Texts, pictures, screenshots, etc. |
| Shirley | Written summaries and reflections on online EFL teaching | | | | | Texts, pictures, screenshots, etc. |

the findings that can be understood by a larger international audience, all the interviews were presented in English.

Our data analysis was also guided by the Interactive Model (see **Figure 3**). As is shown, the condensation, display, and conclusion drawing/verifying are closely interwoven or connected to data collection, despite the whole process being called "analysis" in the general sense. Inspired by the Interactive Model, we created a five-step data analysis procedure (see **Table 3**).

# FINDINGS AND DISCUSSION

## Teachers' Cognitions About Online EFL Teaching Over COVID-19

The sudden outbreak of COVID-19 across the whole globe posed a crisis to traditional face-to-face ways of EFL teaching in Chinese universities. In the meantime, it also provided an opportunity for all parties involved to update their cognitions about EFL teaching and upgrade their knowledge and skills of information technology literacy to meet the needs of online teaching. We cannot reset 2020, but it can be argued that we can reset ourselves with updated cognitions and upgraded knowledge and skills of information technology literacy.

### Teachers' Perceptions of Online EFL Teaching Over COVID-19

The participating teachers had diverse perceptions of online EFL teaching over COVID-19 as they compared it with traditional classroom language teaching to explore the features of online EFL teaching. Based on their online teaching experiences, they looked at online EFL teaching from different perspectives. Jane expressed her view of teaching and learning as limited by teachers' mastery of information technology in online EFL teaching:

> Jane: *Online teaching is diametrically different from traditional classroom teaching. In the traditional teaching mode, teachers prepare the lessons and deliver them in the classroom, and interact with students face-to-face, relying little on information technology and the network infrastructure. They can fulfill their teaching tasks within the specified time. However, online teaching is conducted in a virtual space, in which face-to-face interaction can hardly be achieved with the same effect as in a physical classroom. Limited by the online teaching conditions, it is difficult to guarantee the full participation of students; the learning outcomes are closely related to self-management and metacognitive ability on the students' part.*

As reflected in Jane's observation, online EFL teaching was totally different from traditional classroom teaching because of its heavy reliance on information and communication technology, which subsequently leads to uncertain learning outcomes on the students' part. Jane's negative view of online EFL teaching was a result of her worry and anxiety about the new and unfamiliar form of teaching in the initial stage. If a chronological case study can be conducted with Jane, findings concerning the change in her ideas of online EFL teaching will be expected. Different form Jane's view, online EFL teaching is "new and efficient extension of traditional classroom teaching" in Mary's and Shirley's eyes. Mary in details illustrated her favorite idea about online EFL teaching in comparison with traditional classroom teaching:

> Mary: *Online EFL teaching expands traditional classroom teaching in three main aspects. First, such online learning platforms as MOOC provide quality resources for teaching and learning, from which the teacher can choose some key content for the students. Second, in terms of supervision and assessment of students' learning process and outcome, more objective records can be retrieved from online teaching. Third, the teacher–student interaction in online teaching can take more forms. Students are more active because they don't need to show their image as they did in traditional classroom teaching which effectively emboldens their courage and confidence in answering questions. Furthermore, the students' adoption of various forms of participation (such as words, pictures, animations, Emojis, and voice messages, etc.) in online teaching livens the atmosphere.*

Online EFL teaching, in Mary's eye, is more positive than negative. Mary's perception about the advantages of online EFL teaching was echoed and supplemented in Shirley's written reflection:

> Shirley: *Online teaching of EFL courses has its unique merits. Students can repeatedly watch teaching videos to meet their learning needs. Teachers can keep track of students' progress and provide targeted support and advice.*

In spite of their optimistic views about online EFL teaching, Mary and Shirley didn't ignore its drawbacks. Mary explained her structured understanding of the disadvantages of online EFL teaching related to teacher–student interaction, reflection of students' learning progress, and inequality in education caused by imbalance among network devices across regions.

> Mary: *Online EFL teaching has its insuperable weaknesses in terms of timely teacher–student interaction, reflection of students' learning development, and requirement of network devices on students' part. First, none of the online teaching platforms can guarantee instantaneous interaction between teachers and students, which subsequently brings negative effect on the teaching efficiency. Due to the time lost in the lengthening of teacher–student interaction, online teaching capacity is less than that of traditional*

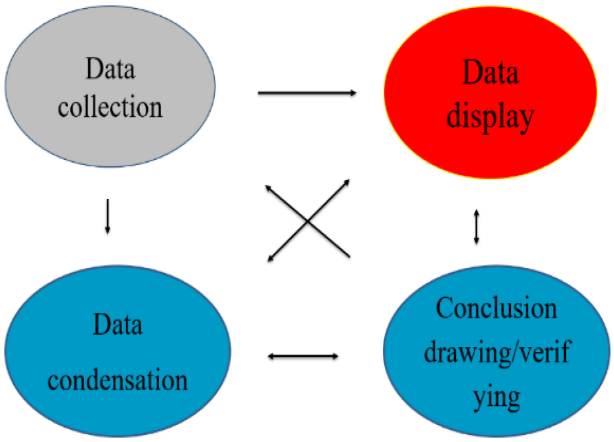


**FIGURE 3 |** Components of data analysis: Interactive model (adapted from Miles and Huberman, 1994; Miles et al., 2014).

*classroom teaching and a lot of work needs to be assigned for the students to do after class. Second, as the teacher cannot see every student in online teaching, it is more difficult to monitor their learning, not to mention their after-class assignments. Third, the imbalance of network devices across regions inevitably leads to inequality in education among students.*

Mary's practice-driven insight into online EFL teaching reflects her thorough understanding of teaching and learning process of EFL courses, which is in turn central to her online teaching practice. Several weeks of online EFL teaching experience widens her understanding of language education and enriches her cognition of online teaching in particular. Mary's thought about the drawbacks of online EFL teaching was overall and general, which was fortunately remedied by Shirley's reflection based on her online classes.

Shirley: *In my online classes, two things are found to hinder teaching efficiency. Prepared teaching materials turn out to be too much for the online class due to my inexperience and unpredictable network signal. Online class activities such as open discussion are inappropriate.*

In addition to the features of online EFL teaching, the change in the role of teachers is also one of the major themes emerging from the data. Jane stated that the role of the teacher changed from the traditional knowledge imparter and the classroom activity organizer to the resource integrator and the supervisor for students' autonomous learning in online teaching mode. She also mentioned that in her online teaching practice she carefully chose online resources provided by leading publishers in foreign language teaching, such as Foreign Language Teaching Press (Beijing) and Foreign Language Education Press (Shanghai), as effective supplementary materials to the textbooks adopted for the coursed she taught.

## Challenges Facing EFL Teachers Teaching Online Over COVID-19

Jane, Mary, and Shirley all expressed their confusion and anxiety during the initial stage of online teaching. In her interview, Jane talked about her psychological pressure during her preparation for online EFL teaching, and mentioned that her worries were mainly about her lack of proper information technology literacy for online teaching, the insufficient conditions for online teaching and learning on both teachers' and students' part, and invalid class management during online teaching.

Jane: *What platform should I use for online teaching? Will it collapse during online teaching? What if unexpected breakdown happens and I am not able to solve it? I am so ill-prepared and worried.*

The first challenge Jane met is typical of EFL teachers and central to online teaching. Teachers were familiar with teaching methods in face-to-face delivery in classrooms before the abrupt breakout of COVID-19 and their information technology literacy was limited to the integration of digital equipment into classroom teaching, with little knowledge and skills for online teaching. This put EFL teachers in a poor position, which restricts them from conducting online teaching effectively especially when they are supposed to do so after a very short period of training. Their uncertainty over the platform, channel, and specific skills, etc. for online teaching poses a challenge before them.

The second challenge is insufficient network conditions for online teaching and learning. Online teaching and learning requires a large amount of data transmission and thus relies heavily on the wi-fi infrastructure or monthly data plans. This is costly for students, too. If a flaw occurs in one segment, the whole online class will be forced to stop. And Jane also worried about the materials for online teaching and learning. As COVID-19 broke out in late December 2019, the students were having winter vacation with their family. After she contacted the class representative, usually called "class monitor" in China, she found that few of the students took their textbooks along when they

**TABLE 3 |** Five-step data analysis model used in the present research.

| Steps | Description of the process |
|---|---|
| 1 Cleaning the original data | Reading through the original data, find and correct the errors in the original data files, check the consistency of the data, etc. |
| 2 Coding the data | Repeatedly reading the original data and generating open codes. |
| 3 Generating themes | Comparing the original codes, checking the relevance between codes, and naming the themes. |
| 4 Categorizing the themes | Putting the themes into different categories according to certain standards and principles. |
| 5 Producing the report | Producing a scholarly report of the analyzed data with vivid, compelling extract examples, relating back to the research question and literature. |

left the university. She was concerned that such situations might hinder her successful online teaching and learning.

The third challenge Jane met was the possible invalid class management during online teaching. Since the teacher and students were not in one actual room during online teaching, as is usually the case, class management became more challenging. The teacher were not able to observe the students and give timely feedback through non-verbal means such as eye contact. The inefficiency in class management may lead to students' idling away and unsatisfactory learning outcome.

### EFL Teacher's Readiness to Deliver Online Teaching to Tide Over COVID-19

In view of the three challenges, Jane, Mary, and Shirley did a lot of autonomous learning and exploration of relevant elements and technological skills. As a result, they found the solutions and became ready for the online teaching.

To solve the first challenge, they learned how to use various Internet platforms such as Chaoxing, Ding Talk, MOOC, etc. for online teaching, and some chat groups in social networking apps such as WeChat and QQ. These platforms and apps have been developing fast since the outbreak of COVID-19. Through learning and trials, they got familiar with their functions including live streaming, uploading, and downloading files, and roll calling, etc. They were also aware of the advantages and disadvantages of each platform and app. The ability to make a choice of suitable online platform(s) prepared them for online EFL teaching and enhanced their confidence in information technology for online teaching.

Facing the second challenge, the three participants all contacted the students for accurate information and took corresponding measures. In Jane's case, she contacted the class monitors 2 weeks before online teaching began and updated her knowledge about the situation of her students and their readiness for online learning. What the monitors told her made her come to a comfortable realization that her preparedness for online teaching met her students' expectations.

> Jane: *The monitor told me that most of the students in her class left their textbooks (Contemporary College English Book Four) in their dormitory when they left for home. Instead they took Exercise books for TEM-4 with them and planned finish the exercises in winter vacation. As for Internet access, every student has wi-fi at home and smart devices for online learning.*

After she knew the students' lack of textbook, Jane asked the monitors to search for the PDF version of the textbook. One day later, this was solved and every student in the class was equipped with the basic learning material – the textbook.

As for class management, Jane sorted the students into learning groups of four or five, assigned learning tasks related to the textbook or the *Test for English Majors – Band 4 (TEM-4)* grammar and vocabulary for each group and shared their learning outcome in the online platform. This worked well in the past 6 weeks, as it kept every student engaged in learning activities. The sharing and feedback encouraged the students to continuously perform for better.

## Teacher's Acquisition of ICT Literacy in the Initial Stage of COVID-19

Three themes related to teachers' acquisition of ICT literacy emerged. In Jane's case, her ICT literacy was acquired as a result of a clear understanding of her students' learning needs; it was also acquired through online teaching practice; her integration of traditional classroom teaching into online teaching with information technology increased her confidence. These themes were also embodied in Mary's and Shirley's cases.

### ICT Literacy Acquired Through Clear Understanding of Students' Learning Needs

In Jane's case, before starting the online teaching, Jane was fully aware of her students' learning needs in the Spring Semester. She mentioned her cognition when she talked about her preparations for online teaching.

> Jane: *My main teaching task this semester is Comprehensive English (IV) for English 1801/1802. The second semester of the sophomore year is one of the most critical semesters for English majors in universities. Along with completing the heavy learning tasks of various courses, they are also required to take part in Test for English Majors Band Four (TEM-4) which is supposed to check the learning outcome in their first 2 years of university.*

As is the rule in Chinese universities, English majors' performance in TEM-4 is very important for the students, failing which they will be deprived of their degree. In the arrangement of tutoring students to prepare for the exam, the course *Comprehensive English* is customarily responsible for two blocks: *Grammar and Vocabulary*, and *Cloze*. These two blocks in TEM-4 are difficult for the students but take a high percentage in the total marks. Students' learning needs put forward higher requirements and challenges, and therefore bring great psychological pressure to Jane and the students. With a clear understanding of the students' learning needs, Jane made a neat choice of teaching materials and teaching methods as well as the most suitable platform for online teaching: Chaoxing. Jane's choice of Chaoxing Platform for online teaching was a result of guidance from her university and college.

> Jane: *Under the guidance and help of the university and college, I quickly got familiar with the operation process of Chaoxing Platform for online learning, and decided to choose it as the main platform for my online teaching platform.*

The Chaoxing online platform has multiple functions and evident advantages as it is specially designed for online teaching and learning: Required textbook chapters and other editable teaching materials make the platform user-friendly, integrated, and systematic; Chaoxing's cloud drive has a large storage capacity, and the data stored in its cloud drive can be synchronized to its mobile phone app for convenient operation; rich online classroom activities in Chaoxing, such as roll calling, voting, topic discussion, question posing, surveys, etc., basically cover activities in offline physical classroom teaching, and effectively promote classroom activities and improve online teaching efficiency; the notification function is

very practical in that it cannot only quickly send notifications to targeted students, but also report the number of students who have read the notification in time to the teacher and be operated to remind those who have not read of reading in time; the homework function provides set an array of question forms, such as multiple choice, gap filling, short answer tasks, etc., in line with the actual needs of teaching, and this function can also grade some homework, which effectively reduces the workload of teachers. The multiple functions of Chaoxing help Jane get familiar with online teaching process, master online teaching methods and skills as soon as possible, and ensure the teaching efficiency of her online teaching.

Shirley's case was similar to Jane's. She chose Chaoxing as her main platform for online teaching and QQ chatgroups for class discussion and Baidu Netdisk for transmitting learning materials.

However, in Mary's case, as she recognized in the *English Listening* course that students needed to do a substantial amount of practice to acquire the language skill, she chose Ding Talk as the online teaching platform for live teaching of the course. Such platforms as MOOC and Chaoxing were conceived by her as improper platforms for online learning of *English Listening* because "they emphasize knowledge transference instead of student-centered practice." In her Ding Talk live class, she directly engaged the students in the listening practice like in a traditional classroom. She started a lesson with providing the students with knowledge about new words and cultural background of the listening item. Then, she played the audio clips four times to the students and asked them to do follow-up exercises during the intervals and provided them with necessary support such as the answers to the exercises and the audio scripts. She asked the students to upload their notes of listening to Ding Talk for a further look at their mistakes and gave them corresponding guidance.

## ICT Literacy Acquired Through Online Teaching Practice

Jane encountered many difficulties and challenges in preparing for her online teaching, such as making files containing both PowerPoint slides and voicing over her PowerPoint slides (PPT + voice), uploading teaching materials to the Chaoxing cloud drive. The first challenge Jane encountered during her preparation for her online teaching was recording the files with EV Capture. **Figure 4** below shows the "PPT + voice" files made by Jane.

> Jane: *When I made the first "PPT + voice" files, I was unfamiliar with the software and often made operational mistakes. In the meantime, without students' response as in the traditional classroom, I was always unsatisfied with the sound quality and atmosphere of recording: it was not lively as supposed to be. Then I need to repeatedly abandon the unfinished files and rework on them from the very outset. As a result, a presentation of a few minutes can take an hour, which upsets me at the end of the day.*

The difficulty in video recording had Jane stumped for quite some time, and then she solved the problem by adding classroom

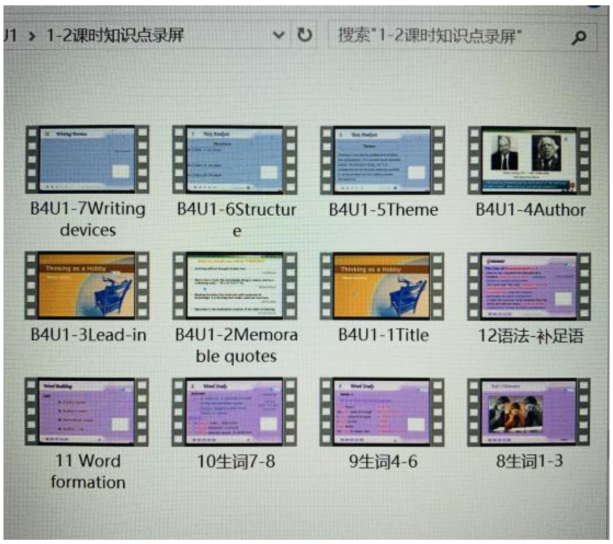


**FIGURE 4 |** The "PPT + Voice" files made by Jane.

teaching skills such as raising questions after knowledge explanation, waiting some seconds for students to digest, and providing feedback to imaginary answers from students.

> Jane: *With imagined teacher–student interaction added, the "PPT + voice" files has become more vivid and lively.*

The second challenge Jane encountered in her preparation for online teaching was uploading the teaching materials onto the Chaoxing cloud drive. She mentioned the reason why it was a challenge for her.

> Jane: *As Chaoxing only operates on its app of mobile phones instead of a software program on computers, how can I transfer teaching materials from my computer to the app? I was totally clueless and frustrated.*

Later, Jane consulted her colleagues and the staff of Chaoxing and learned how to upload the materials to the Chaoxing cloud drive through "Erya," an online teaching channel run by the university. The materials then can be inserted into the catalog of chapters, synchronized with Chaoxing app, ready for the students to access and download from their smart phones. In the meantime, Jane also learned to develop her courseware on the Chaoxing website and its operations, becoming increasingly familiar with Chaoxing for online teaching.

Similar processes happened to Shirley. She found the difficulties in creating "PPT + voice" files lied in her lively tone as lecturing before her 69 students and the prediction of students' obstacles in learning.

## Integrating Traditional Classroom Teaching Into Online EFL Teaching With ICT

In traditional classrooms, before getting into the lessons, Jane would call the roll to make sure all the students were present and ready to learn. She integrated roll calling into her online class as well. The first activity in her online teaching before

actual teaching began was roll calling. The roll calling function in Chaoxing often collapsed in the first few weeks as millions of students across China were logging onto the platform for online learning.

> Jane: *My targeted countermeasures include extending the roll calling time by starting it half an hour earlier, transferring the roll calling to social networking platforms such as QQ or WeChat groups, or replacing it with random mass calling during lessons.*

Jane found that these countermeasures worked well. They helped ensure the on-time attendance and full participation of her students in her online lessons.

The most important content of teaching in either traditional classrooms or online teaching is to impart knowledge to students and inspire students to think. Some hours before online class time, Jane uploaded the recorded "PPT + voice" video files and PDF teaching materials through the Erya channel into her Chaoxing cloud drive and course folders. In this way the students could access these materials through Chaoxing app on their smart phones. In the online class, Jane asked students to watch the video files and learn the texts with the help of the teaching materials. After watching the videos and self-study, students, started the discussion during which Jane raised questions to provoke students' critical thinking and analysis, summarize important grammar items in the text, and reflect on their learning processes.

Accurate and effective teaching of EFL is very important in classroom teaching. So Jane integrated this belief into her online teaching by formatting the two 50-min lessons into three sections. The first section started with daily greetings and accurate descriptions about the learning objectives and corresponding learning tasks. Her students were asked to type 1 if they were clear about the tasks. Then the students were asked to spend 20 min or so watching the video files and learning autonomously. They typed 2 when they completed the task. In the second section, Jane raised questions about the knowledge students learned from the videos and the PDF files as described above. The third section began 10 min before the class was over, during which Jane highlighted the key and difficult points and assigned after-class homework for review of what was learned and preview of the next lesson content.

One of the key teaching methods in the traditional classroom teaching process is raising open-ended questions and providing feedback. Open-ended questions can better trigger students' divergent thinking. Jane applied this method to her own online teaching. She used Text A in Unit One, *Thinking as a Hobby,* as an example to illustrate how she carried this out in her online teaching.

> Jane: *In this text, the author explained his own thoughts on thinking and the dynamic development process by describing views and expressions of many characters on thinking in his process of growing up.*

To help students grasp how the author formed and developed his thinking, Jane assigned homework of character analysis through Chaoxing, as shown in **Figure 5**. The students first listed what the characters said and did as narrated in the text, and then summarized their characters and personalities. In the discussion section of online teaching, Jane led the students to further analyze the features of the characters (such as Mr. Houghton and Ruth) and their thinking. She also asked students find out the role they played in the development of the author's thinking. In the text, three grading scales for measuring thinking are used (Grade 3, Grade 2, and Grade 1). Grade-one thinking is ranked the highest and indicates the best level of thinking; the characteristics of each grade of thinkers are also illustrated in the text. Through comparison and analysis, the students learned that they should be neither grade-three thinkers full of ignorance and hypocrisy, nor grade-two thinkers who destroy without the power to create; instead they should be trying to be grade-one thinkers who are independent thinkers and set out to seek truth based on logical reasoning.

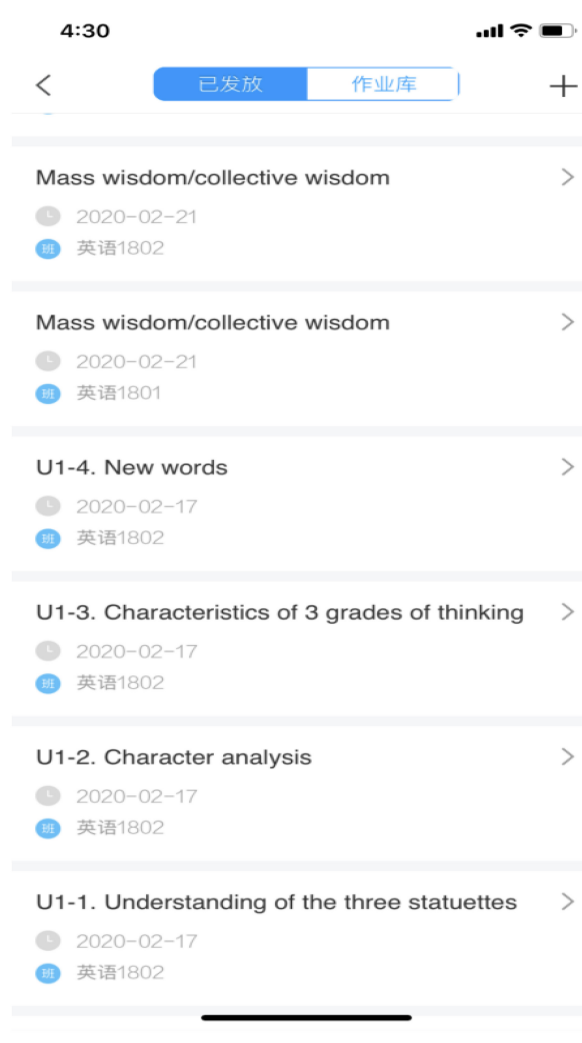


**FIGURE 5 |** The homework designed by Jane in Chaoxing.

In view of the research questions stated above, we can see that teachers have identifiable cognitions about online EFL

teaching over COVID-19. The interview data and the reflections showed that the participating teachers realized that online EFL teaching, different from traditional classroom teaching, is limited by teachers' mastery of information technology; online EFL teaching has its apparent advantages such as quality online resources, retrievable online records of students' learning process and outcome, and lively teacher–student interaction, etc.; meanwhile, online EFL teaching has insuperable deficiencies in its failure to offer instantaneous teacher–student interaction, to monitor their in-class performance, and to ensure education equality caused by imbalance in network devices across regions. This is consistent with Borg's (2015) model of language teacher cognition: Classroom practice as a contextual factor affects teachers' cognitions unconsciously and through conscious reflection.

Teachers' cognitions about online EFL teaching are congruent with their online instructional practice. They chose appropriate platforms and adopted various teaching methods based on their cognitions about students' learning needs in their courses. This finding is in keeping with the findings of many studies on cognition-practice relationship of language teachers (e.g., Borg, 2011; Basturkmen, 2012; Farrell and Ives, 2015; Li, 2020).

Concerning our second research question, teachers acquired their ICT literacy through their clear understanding of students' learning needs and was facilitated by online teaching practice and integrating traditional classroom teaching. This finding aligns with the TPACK model developed by Koehler and Mishra (2005) in that TPK is the knowledge domain that has overlaps with TK and PK. Teachers' integration of their teaching methods in traditional classrooms into online EFL teaching is a new and meaningful finding that adds new empirical evidence to the existing research on teacher cognitions and TPACK.

## CONCLUSION

This qualitative inquiry, as part of a larger study, was set up to examine EFL teachers' cognitions about online teaching in response to their disrupted teaching plans, and how they acquired their information technology literacy in the initial stage of the COVID-19 outbreak. Such research on teacher cognition about online teaching during COVID-19 and the findings have theoretical implications for research on online EFL teaching and on teacher cognition. It has modestly expanded the knowledge about online EFL teaching by focusing on what teachers think and believe about online EFL teaching. The qualitative research adopted in the present study has counterbalanced the dominant trend in foreign language research that is predominantly quantitative in methodology. This study has also provided new empirical evidence for research on teacher cognition about online EFL teaching because the findings related to teachers' cognitions about the features, advantages, and weaknesses of online EFL teaching are new to the field of language teacher cognition research. The findings from the present study might also have pedagogical implications for EFL teachers. These are contextually relevant and meaningful findings in this trying time. These findings further illustrate that research on language teacher cognition needs to be further investigated with reference to specific contexts. Hopefully, our findings can help EFL teachers in the Chinese context or any other similar context in other parts of the world to understand that language teaching is an endeavor full of complexities and unexpected events. Language teachers, as colleagues in other disciplines, need to be flexible, resilient, and ready to learn new skills for tiding over unexpected challenges such as COVID-19.

Despite its significance, our study has one obvious limitation: The lack of data from students. COVID-19 imposed challenges to both teachers and students. Data from students' perspectives would allow for a more holistic study. Future research might consider including data from students' perspectives for a more generalizable and holistic study. Teachers' personal experiences such as birthplaces, schooling, and professional coursework may lead to differences in their cognitions about online EFL teaching. Meanwhile, students' cognitions about online EFL teaching and learning collected through interviews and/or other data collection instruments would provide feedback to teachers and facilitate their cognitions about EFL teaching, especially online teaching of EFL to students in similar contexts.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the participants for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

LG and LZ conceptualized the study. LG collected the data, analyzed them, and wrote the first draft. LZ contributed to the rewriting and revising of the first draft and subsequent revisions. Both authors contributed to the article and approved the submitted version.

## FUNDING

The research reported in this manuscript was supported by a grant from the Science and Technology Department (SATD) of Shanxi Province for the project (No. 2019041009-4) titled

"A Study on English Teachers' Competence and Professionalism in Information Technology-Enhanced Teaching and Learning Environments" and a grant of the Teaching Reform and Innovation Initiative from Taiyuan University of Technology, titled "Chinese University Teachers' Cognition and Practice in Teaching English-as-a-foreign-language (EFL) Listening."

## ACKNOWLEDGMENTS

We are grateful to the participants for their willingness to be involved in this research study. We are also thankful to the reviewers and editors for their very constructive feedback that has helped us improve the clarity of our manuscript.

## REFERENCES

Anstrom, K. (2003). *Team Teaching Between English as a Second Language and Content Specialists at the Secondary Level: A Case Study of Teacher Beliefs and Practices*. doctoral dissertation, George Washington University, Washington DC.

Baleghizadeh, S., and Nasrollahi Shahri, M. N. (2014). EFL teachers' conceptions of speaking competence in English. *Teach. Teach.* 20, 738–754. doi: 10.1080/13540602.2014.885702

Basturkmen, H. (2012). Review of research into the correspondence between language teachers' stated beliefs and practices. *System* 40, 282–295. doi: 10.1016/j.system.2012.05.001

Berliner, D. C. (1994). "Expertise: the wonder of exemplary performances," in *Creating Powerful Thinking in Teachers and Students*, eds J. N. Mangieri and C. C. Block (Fort Worth, TX: Harcourt Brace), 161–186.

Borg, S. (2003). Teacher cognition in language teaching: a review of research on what language teachers think, know, believe, and do. *Lang. Teach.* 36, 81–109. doi: 10.1017/s0261444803001903

Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Lang. Teach. Res.* 10, 3–31. doi: 10.1191/1362168806lr182oa

Borg, S. (2011). The impact of in-service teacher education on language teachers' beliefs. *System* 39, 370–380. doi: 10.1016/j.system.2011.07.009

Borg, S. (2015). *Teacher Cognition and Language Education: Research and Practice*. London, England: Bloomsbury.

Borg, S. (2019). "Language teacher cognition: perspectives and debates," in *Second Handbook of English Language Teaching*, ed. X. Gao (Cham: Springer Nature).

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Q. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Buss, L. (2016). Beliefs and practices of Brazilian EFL teachers regarding pronunciation. *Lang. Teach. Res.* 20, 619–637. doi: 10.1177/1362168815574145

Chapelle, C. (2005). "Computer-assisted language learning," in *Handbook of Research in Second Language Teaching and Learning*, ed. E. Hinkel (Mahwah, NJ: Lawrence Erlbaum Associates), 743–755.

Clandinin, D. J., and Connelly, F. M. (1987). Teachers' personal knowledge: what counts as 'personal' in studies of the personal. *J. Curric. Stud.* 19, 487–500.

Cooper, G. (2019). Teachers' cognitions of corrective feedback on pronunciation: their beliefs, perceptions and practices. *System* 84, 41–52. doi: 10.1016/j.system.2019.04.003

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th Edn. Thousand Oaks, CA: Sage.

Cummins, C., Cheek, E., and Lindsey, J. (2004). The relationship between teachers' literacy beliefs and their instructional practices: a brief review of the literature for teacher educators. *E J. Teach. Learn. Diverse Sett.* 1, 175–188.

Demouy, V., Jones, A., Kan, Q., Kukulska-Hulme, A., and Eardley, A. (2016). Why and how do distance learners use mobile devices for language learning? *EuroCALL Rev.* 24, 10–24. doi: 10.4995/eurocall.2016.5663

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press.

Ellis, E. M. (2006). Language learning experience as a contributor to ESOL teacher cognition. *TESL EJ* 10, 1–20.

Evans, C. (2008). The effectiveness of m-learning in the form of podcast revision lectures in higher education. *Comput. Educ.* 50, 491–498. doi: 10.1016/j.compedu.2007.09.016

Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educ. Res.* 38, 47–65. doi: 10.1080/0013188960380104

Farrell, T. S., and Ives, J. (2015). Exploring teacher beliefs and classroom practices through reflective practice: a case study. *Lang. Teach. Res.* 19, 594–610. doi: 10.1177/1362168814541722

Farrell, T. S., and Kun, S. T. K. (2007). Language policy, language teachers' beliefs, and classroom practices. *Appl. Linguist.* 29, 381–403. doi: 10.1093/applin/amm050

Farrell, T. S., and Lim, P. C. P. (2005). Conceptions of grammar teaching: a case study of teachers' beliefs and classroom practices. *TESL EJ* 9, 1–13.

Fu, W., and Zhou, H. (2020). Challenges brought by 2019-nCoV epidemic to online education in China and coping strategies. *J. Hebei Normal Univ. (Educ. Sci.)*, 22, 14–18.

Gao, L. X. (2019). *Teacher Cognition and Practice in Classroom Listening Instruction: Case Studies of Five English-as-a-Foreign-Language (EFL) Teachers in a Chinese University*. PhD thesis, The University of Auckland, Auckland.

Gao, X., and Ma, Q. (2011). Vocabulary learning and teaching beliefs of pre-service and in-service teachers in Hong Kong and mainland China. *Lang. Awareness* 20, 327–342. doi: 10.1080/09658416.2011.579977

Gerami, M. R., and Noordin, N. B. (2013). Teacher cognition in foreign language vocabulary teaching: a study of Iranian high school EFL teachers. *Theory Pract. Lang. Stud.* 3, 1531–1545.

Gilliland, B. (2015). High school teacher perspectives and practices: second language writing and language development. *Lang. Educ.* 29, 287–301. doi: 10.1080/09500782.2014.1001398

Graus, J., and Coppen, P. (2016). Student teacher beliefs on grammar instruction. *Lang. Teach. Res.* 20, 571–599. doi: 10.1177/1362168815603237

Jiang, H., Jiang, L., and Yu, R. (2020). The construction of online teaching modes during COVID-19. *China Educ. Technol.*, 399, 40–41.

Kagan, D. M. (1992). Implication of research on teacher belief. *Educ. Psychol.* 27, 65–90. doi: 10.1207/s15326985ep2701_6

Kaymakamoglu, S. E. (2018). Teachers' beliefs, perceived practice and actual classroom practice in relation to traditional (teacher-centered) and constructivist (learner-centered) teaching (note 1). *J. Educ. Learn.* 7, 29–37. doi: 10.5539/jel.v7n1p29

Koehler, M., and Mishra, P. (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *J. Educ. Comput.Res.* 32, 131–152. doi: 10.2190/0ew7-01wb-bkhl-qdyv

Koehler, M. J., Shin, T. S., and Mishra, P. (2011). "How do we measure TPACK? Let me count the ways," in *Educational Technology, Teacher Knowledge, and Classroom Impact: A Research Handbook on Frameworks and Approaches*, eds R. N. Ronau, C. R. Takes, and M. L. Ness (Hershey, PA: IGI Global), 16–31. doi: 10.4018/978-1-60960-750-0.ch002

Kukulska-Hulme, A. (2009). Will mobile learning change language learning? *ReCALL* 21, 157–165. doi: 10.1017/s0958344009000202

Li, L. (2020). *Language Teacher Cognition: A Sociocultural Perspective*. London: Palgrave Macmillan.

Liu, J., Yu, S., and Ran, M. (2008). "Research on the communicative mobile English learning model," in *Paper Presented at the 5th IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education*, Wuhan.

Merriam, S. B. (1998). *Case Study Research in Education: A Qualitative Approach*. San Francisco, CA: Jossey-Bass.

Merriam, S. B., and Tisdell, E. (2016). *Qualitative Research: A Guide to Design and Implementation* 4th Edn. San Francisco, CA: Jossey-Bass.

Miles, M., and Huberman, M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* 2nd Edn. Thousand Oaks, CA: Sage.

Miles, M. B., Huberman, A. M., and Saldana, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* 3rd Edn. Thousand Oaks, CA: Sage.

Ministry of Education of China. (2017). The guideline for college English teaching. http://www.360doc.com/content/17/0203/14/413468_626210661.shtml.

Mitchell, E. W. (2005). *The Influence of Beliefs on the Teaching Practices of High School Foreign Language Teachers*. doctoral dissertation, University of Massachusetts Amherst, Massachusetts.

Ng, E., and Farrell, T. (2003). "Do teachers' beliefs of grammar teaching match their classroom practices? A Singapore case study," in *English in Singapore: Research on grammar teaching*, eds D. Deterding, A. Q. Brown, and E. L. Low (Singapore: McGraw Hill), 128–137.

Nunan, D., and Richards, J. C. (2015). *Language Learning Beyond the Classroom*. New York, NY: Routledge.

Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Rev. Educ. Res.* 62, 307–332. doi: 10.3102/00346543062003307

Popko, J. (2005). "How MA-TESOL students use knowledge about language in teaching ESL classes," in *Applied Linguistics and Language Teacher Education*, ed. N. Bartels (New York, NY: Springer), 387–403. doi: 10.1007/1-4020-2954-3_22

Rahimi, M., and Zhang, L. J. (2015). Exploring non-native English-speaking teachers' cognitions about corrective feedback in teaching English oral communication. *System* 55, 111–122. doi: 10.1016/j.system.2015.09.006

Richards, J. C. (1998). *Beyond Training: Perspectives on Language Teacher Education*. Cambridge: Cambridge University Press.

Richards, J. C., and Lockhart, C. (1996). *Reflective Teaching in Second Language Classrooms*. Cambridge: Cambridge University Press.

Rubin, H. J., and Rubin, I. S. (2011). *Qualitative Interviewing: The Art of Hearing Data*. Thousand Oaks, CA: Sage.

Sechrest, L., Fay, T. L., and Zaidi, S. H. (1972). Problems of translation in cross-cultural research. *J. Cross-Cultural Psychol.*, 3, 41–56. doi: 10.1177/002202217200300103

Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educ. Res.* 15, 4–14. doi: 10.3102/0013189x015002004

Sugiyama, A. (2003). *Beliefs and Reality: How Educational Experiences in the United States Affect Teaching Practices of Japanese EFL Teachers*. doctoral dissertation, State University of New York at Buffalo, New York, NY.

Sun, Q. (2017). *An Exploratory Study of Chinese University English-As-a-Foreign-Language (EFL) Teachers' Cognitions and Practices About English Grammar Teaching*. doctoral dissertation, The University of Auckland, Auckland.

Sun, Q., and Zhang, L. J. (2019). Ontological, epistemological, and methodological perspectives on research into language teacher cognition. *Foreign Lang. Res.* 38, 73–81.

Tai, H., Pan, M., and Lee, B. (2015). Applying technological pedagogical and content knowledge (TPACK) model to develop an online English writing course for nursing students. *Nurse Educ. Today* 35, 782–788. doi: 10.1016/j.nedt.2015.02.016

Tucker, R. (2001). *Beliefs and Practices of Teachers of English as a Foreign Language in Guiyang China Middle Schools*. doctoral dissertation, Wayne State University, Michigan.

Vaish, V. (2012). Teacher beliefs regarding bilingualism in an English medium reading program. *Int. J. Biling. Educ. Bilingu.* 15, 53–69. doi: 10.1080/13670050.2011.594496

Van Nes, F., Abma, T., Jonsson, H., and Deeg, D. (2010). Language differences in qualitative research: is meaning lost in translation? *European J. Ageing*, 7, 313–316. doi: 10.1007/s10433-010-0168-y

Wang, Y., Wang, Y., Li, Y., and Li, D. (2020). Countermeasures and suggestions for improving the quality of online education during epidemic prevention and control. *China Med. Educ. Technol.*, 34, 119–124, 128.

Xiang, M., and Borg, S. (2014). College english teachers' beliefs about effective language teaching. *Asian J. Engl. Lang. Teach.* 24, 97–120.

Zembylas, M. (2005). Beyond teacher cognition and teacher beliefs: the value of the ethnography of emotions in teaching. *Int. J. Qual. Stud. Educ.* 18, 465–487. doi: 10.1080/09518390500137642

Zhang, L. J. (2010). A dynamic metacognitive systems account of Chinese university students' knowledge about EFL reading. *TESOL Q.* 44, 320–353. doi: 10.5054/tq.2010.223352

Zhang, L. J. (2016a). "A dynamic metacognitive systems perspective on language learner autonomy," in *Language Learner Autonomy: Teachers' Beliefs and Practices in Asian Contexts*, eds R. Barnard and J. Li (Phnom Penh: IDP Education), 150–166.

Zhang, L. J. (2016b). Reflections on the pedagogical imports of western practices for professionalizing ESL/EFL writing and writing-teacher education. *Austr. Rev. App. Linguist.* 39, 203–232. doi: 10.1075/aral.39.3.01zha

Zhang, L. J., and Ben Said, S. (2014). "Toward a global understanding of local initiatives in language teaching and teacher education: global rules, local roles," in *Language Teachers and Teaching: Global Perspectives, Local Initiatives (pp. xviii-xxx)*, eds S. Ben Said and L. J. Zhang (New York, NY: Routledge).

Zhang, L. J., and Rahimi, M. (2014). EFL learners' anxiety level and their beliefs about corrective feedback in oral communication classes. *System* 42, 429–439. doi: 10.1016/j.system.2014.01.012

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 05 January 2021
doi: 10.3389/fpsyg.2020.558466

# Technology-Assisted Self-Regulated English Language Learning: Associations With English Language Self-Efficacy, English Enjoyment, and Learning Outcomes

*Zhujun An[1], Chuang Wang[1], Siying Li[2], Zhengdong Gan[1]* and Hong Li[1]*

[1] University of Macau, Taipa, China, [2] Chongqing University of Education, Chongqing, China

This study investigated Chinese university students' technology-assisted self-regulated learning (SRL) strategies and whether the technology-based SRL strategies mediated the associations between English language self-efficacy, English enjoyment, and learning outcomes. Data were collected from 525 undergraduate students in mainland China through three self-report questionnaires and the performance on an English language proficiency test. While students reported an overall moderate level of SRL strategies, they reported a high level of technology-based vocabulary learning strategies. A statistically significant positive relationship was noted between the use of technology-based SRL strategies and students' English learning outcomes. English language self-efficacy and English language enjoyment were both related to technology-based SRL strategies. Furthermore, SRL strategies fully mediated the relationship between English enjoyment and English learning outcomes, but the association between English enjoyment and SRL strategies was only partially mediated by English language self-efficacy. Pedagogically, findings of this study suggest that training and instruction aimed at promotion of modern educational technology among students need to give attention to developing their strategic awareness of motivation regulation in optimizing effectiveness of their technology use in learning the target language.

**Keywords: technology-assisted language learning, self-regulated learning, English language self-efficacy, English enjoyment, learning outcomes**

OPEN ACCESS

***Edited by:***
*Ana Lucia Pereira, Universidade Estadual de Ponta Grossa, Brazil*

***Reviewed by:***
*Jun Wei, Tsinghua University, China*
*Hassan Mohebbi, European Knowledge Development Institute (EUROKD), Turkey*

****Correspondence:***
*Zhengdong Gan*
*zhengdonggan@um.edu.mo*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *02 May 2020*
***Accepted:*** *04 December 2020*
***Published:*** *05 January 2021*

***Citation:***
*An Z, Wang C, Li S, Gan Z and Li H (2021) Technology-Assisted Self-Regulated English Language Learning: Associations With English Language Self-Efficacy, English Enjoyment, and Learning Outcomes. Front. Psychol. 11:558466. doi: 10.3389/fpsyg.2020.558466*

## INTRODUCTION

The use of technology has received increasing recognition as a means capable of bridging formal and informal settings in the target language learning (Jung, 2015; Botero et al., 2018) and enabling students to actively and effectively use technology both inside and outside the classroom. There has been an increasingly large body of research on students' use of technology for second or foreign language learning (e.g., Lee et al., 2016; Godwin-Jones, 2018; Lai et al., 2018; Su et al., 2018). These research studies have generally concentrated on students' perceptions and evaluations of the suitability of technological devices for language learning, adoption of these technological devices in the classroom settings, and the factors that affect the effectiveness of language learning

in classroom technology-using conditions (Winke and Goertler, 2008; Peters et al., 2009; Steel and Levy, 2013). For example, Winke and Goertler (2008) found that songs and movies were the most frequently used technologies and the ease of access was the strongest predictor of the frequency of technology use. Peters et al. (2009) also reported that listening to music and viewing video files were two of the most highly preferred activities. Recent research on technology-facilitated language learning, however, has been mostly laboratory and classroom experiments of technology applications in the formal educational contexts (e.g., Burston, 2015; Demouy et al., 2016; Kukulska-Hulme, 2016; Tseng and Yeh, 2019). Consequently, our knowledge and understanding of students' self-regulated use of technology for target language learning is still limited. Aspects of technology-assisted language learning such as goal setting, motivation-regulation, and cognitive strategy use particularly in an English as a Foreign Language (EFL) context remains in need of further empirical inquiry. After all, in the course of learning a second or foreign language, learners are at the center of learning and play an instrumental role in shaping outcomes of their learning experiences. Key to this view of learner-centredness is self-regulation and learners taking the responsibility for their own learning (Holec, 1981). Nevertheless, what is lacking in recent research on technology-assisted language learning is a systematic examination of SRL strategies in technology-using conditions particularly in an EFL context. This study investigated Chinese university students' technology-assisted self-regulated learning (SRL) strategies and whether the technology-based SRL strategies mediated the associations between English language self-efficacy, English enjoyment, and learning outcomes.

## Self-Regulated Learning

Self-regulated learning has been widely acknowledged to be learners' systematic effort to manage and regulate their learning process in order to achieve particular learning goals (Pintrich, 2004; Zimmerman and Schunk, 2011). While different theoretical models provide different definitions of SRL, there is a consensus that SRL is a multifaceted construct containing cognitive, metacognitive, behavioral, and self-motivational aspects (Zimmerman and Schunk, 2001; Schunk and Zimmerman, 2012). According to Pintrich et al. (1991), cognitive strategies refer to the skills that learners use to process the information and knowledge when completing a task. They help students to construct, transform, and apply second language (L2) knowledge (Oxford, 2013). Metacognitive strategies refer to the skills students use to control and regulate the cognition and cognitive resources, which helps in goal setting, planning, monitoring, and evaluating their learning outcomes (Winne, 2011). Social-behavioral strategies, as a key aspect of self-regulation, involve learners' control over their learning behavior under the influence of contextual aspects (Zimmerman, 2011). Finally, motivational strategies refer to the procedure or thoughts students applied intentionally to sustain or increase their interest to engage in a task (Wolters, 1999). Note that in various models of SRL in the literature (e.g., Butler and Winne, 1995; Nicol and Macfarlane-Dick, 2006), self-regulated learners are depicted as being capable of controlling over the cognitive, emotional, motivational, and behavioral aspects of learning (Zimmerman and Schunk, 2011). Research also shows that those more effective at self-regulation use a broader repertoire of learning strategies and persist longer in the face of adversity compared to their less self-regulated counterparts (Pajares, 2009).

## Self-Regulated Language Learning in Technology-Using Conditions

Benson (2001) described two important categories of learning resources: traditional learning resources (e.g., reference and course books) and resources provided by modern educational technology (e.g., information communication technology applications). According to Benson, self-regulation is manifested not only in the active regulation of learning strategies but also in the management of different kinds of learning resources. Self-regulated language learning in technology-using conditions has thus often been described as being characterized by learners developing learning strategies such as planning and resource management, and reflecting on as well as evaluating their learning behavior and outcomes (Carneiro et al., 2007). As such, technology-based self-regulated English learning (SRL) strategies refer to specific actions taken by the learners to learn English or to enhance their English learning in technology-using conditions. A large quantity of technology-assisted SRL strategies were identified in previous studies conducted in a variety of research contexts, such as consulting online dictionaries, using translation software, reading texts on the computer, searching the web for information, listening to the radio, exploring cultural knowledge on YouTube and so on (Steel and Levy, 2013; Lai et al., 2018; An et al., 2020; Wang and Chen, 2020).

A number of studies also provide evidence that technology-assisted SRL strategies impact the improvement of learning performance. Chang and Chang (2014) explored Taiwanese college students' listening comprehension strategies on the platform of YouTube, with data collected by Oxford's (1990) Strategy of Inventory for Language Learning (SILL). They found that students performed significantly better on listening comprehension tests after the metacognitive instructional process. Bekleyen and Hayta (2015) investigated the role of mobile phone technology in the employment of language learning strategies among Turkish undergraduate students. Their study employed a self-designed questionnaire to collect data on students' language learning strategies, which was also based on Oxford's (1990) classification of the language learning strategies. Their results show that different types of mobile phone-assisted language learning strategies are helpful in improving students' English proficiency. Nevertheless, constrained by the adoption of Oxford's (1990) classification of the language learning strategies, these studies largely focused on students' use of cognitive and metacognitive strategies. Furthermore, self-regulation is context- and situation-specific, which means that measurement of technology-based self-regulated language learning should be domain-specific (Wang and Zhan, 2020).

Somewhat unlike the studies reviewed above, Lai and Gu (2011) examined how language learners relied on their metacognitive knowledge to regulate different aspects of their

language learning experience, and further identified various factors that affected the participants' selective use of technology for language learning. Encouraging as Lai and her associates' findings were, their study did not include and report cognitive strategies learners used in technology-assisted second language learning contexts. Given the potential technology opens up for language learning, we believe that knowledge of what technology-based cognitive learning strategies students prefer and what strategies were possibly omitted in the previous research will be useful in designing student training programs. We therefore suggest that technology-assisted strategic language learning as a process be examined from a multidimensional perspective that includes an understanding of cognitive, metacognitive, social, and motivational components, and an understanding of how they interact with individual factors and learning achievement.

## Self-Efficacy

Self-efficacy refers to individuals' personal evaluations of their capability of accomplishing a particular task (Schunk et al., 2008). According to Bandura (2006), efficacy beliefs influence the courses of action people choose to pursue, the challenges and goals they set for themselves and their commitment to them, how much effort they put forth in given endeavors, and the outcomes they expect their efforts to produce. Social cognitive theorists have emphasized the role of self-efficacy beliefs in SRL as predictors of academic performance as self-efficacy beliefs can be modified in school practice to promote better academic performance (Zuffianò et al., 2013). The basic presumption as to why self-efficacy has an effect on students' SRL is that when students experience feelings of worth and a perception of improved capability, students are likely to perform better and therefore to experience successful performance (Panadero et al., 2017). For example, in Bassi et al.'s (2007) study, high self-efficacy students were found to report higher academic aspirations and pursuits than low self-efficacy students, and spend more time in homework and primarily associate learning activities with optimal experience. Self-efficacy beliefs have therefore been acknowledged to fulfill a significant role in understanding the academic lives of students as it influences their motivation, affect, and behaviors (Bandura, 2006). While there is prolific research on self-efficacy in the general education field, it is only within the past two decades that self-efficacy has been attracting researchers' attention in the field of L2 acquisition. Similarly, in a study of graduate pre-service teachers' language learning strategies and language self-efficacy in Malaysia, Wong (2005) reported that high self-efficacy pre-service teachers reported more frequent use of more language learning strategies than did low self-efficacy pre-service teachers. A study of the influence of self-efficacy and other motivational self-beliefs on the achievement among college intermediate French students also revealed that self-efficacy for self-regulation was the most significant predictor of intermediate French language achievement, and that students who perceived themselves as capable of using effective metacognitive strategies to monitor their academic work time effectively were more apt to experience academic success in intermediate French (Mills et al., 2007).

Note that more recently, a number of L2 studies tended to develop new self-report self-efficacy measurements to investigate the role of self-efficacy in the L2 learning process. For example, to address the need for valid and reliable tools to assess ESL learners' self-efficacy, Wang et al. (2013) developed the English Self-Efficacy Questionnaire to measure English self-efficacy in the following four areas: (1) English listening; (2) English speaking); (3) reading; and (4) writing. Subsequent Confirmatory factory analysis (CFA) with data from Chinese university students confirmed a second-order common factor with these four first-order latent constructs: English listening, English speaking, English reading, and English writing. To date, studies that adopted the English Self-Efficacy Questionnaire showed that students' English language self-efficacy influenced their use SRL (e.g., Wang and Bai, 2017). English language self-efficacy was also found to positively influence students' feedback preferences and behavior in academic English course settings (Bai and Wang, 2020; Gan et al., 2020). In the existing studies on learners' use of information and communication technologies, however, self-efficacy tends to be operationalized as students' confidence in their ability to select appropriate technological solutions and utilize the chosen technologies effectively to meet learning needs (Lai and Gu, 2011). To the best of our knowledge, the role of the learners' English language self-efficacy in determining their selection and utilization of technology for English learning and enhancement has not been researched.

## Foreign Language Enjoyment

Positive emotions, such as enjoyment, pride, and flow, have been regarded as being efficient in facilitating learning (MacIntyre and Gregersen, 2012; Lake, 2013). Among the positive emotions, enjoyment has been recognized as a most typical positive emotion experienced by foreign language learners and has received increasing attention from researchers in the field of educational psychology (e.g., Csikszentmihalyi, 1990; Dewaele and MacIntyre, 2016, Dewaele et al., 2017; Li et al., 2018). Enjoyment was a sense of satisfaction and reward that generated from activities or the achievement of activities (Ainley and Hidi, 2014). In the literature of educational psychology, enjoyment is often defined as a positive psychological state coming from the efforts by the person who stretches beyond himself to accomplish something challenging or difficult (Csikszentmihalyi, 1990). In the area of English language education, English enjoyment refers to students' liking for learning English as a foreign language. Specifically, in the foreign language learning context, experiencing enjoyment involves concentration, clearing goals and immediate feedback that can help learners build resources (Li et al., 2018). Dewaele and MacIntyre (2016) observed that individual learners who were more proficient than their classmates and who eventually reached a higher level proficiency in the target language demonstrated a significantly higher level of enjoyment than their peers. Foreign language enjoyment was also positively associated with academic achievement by promoting psychological resiliency, relieving negative arousal, and broadening learners' instant thought-action repertoires (Ryan et al., 1990; Lai et al., 2018; Piniel and Albert, 2018). In the literature on students' use of technology for language learning,

while there were various accounts of language learners actively engaging in self-initiated learning activities, it was also frequently reported that there was little and limited use of technologies for language improvement particularly outside the classroom, despite the frequent use of a wide repertoire of technologies for entertainment or infotainment (Lai and Gu, 2011). Such little and limited use of technologies for real language learning purposes is often attributed to a lack of intrinsic interest or enjoyment toward the target language the students are learning. Nevertheless, this assumption has not been empirically well-tested yet.

Clearly, it can be concluded from the above review that while there has been an attempt to integrate learning strategies with elements of SRL and metacognition in the context of technology supported language learning, the literature on the application of self-regulation in technology-supported second language learning is still fairly limited. Although the importance of the role of the strategic and motivational factors in first and second language contexts has been well documented, how these factors function in relation to students' learning achievement in the context of technology use for self-regulated language learning has been under-researched.

## Associations Among Technology-Based SRL Strategies, English Language Self-Efficacy, English Enjoyment, and Learning Outcomes

According to the Expectancy-Value Theory (EVT) (Wigfield and Eccles, 2000; Eccles, 2009), students' expectancies and values play significant roles in their academic learning. Specifically, individual learners' expectancies for success and their subjective values (i.e., attainment, intrinsic value, extrinsic utility, and cost) are assumed to directly influence their educational and behavioral choices. Expectancies and values also directly influence learners' performance, effort, and achievement (Wigfield and Eccles, 2000). Expectancies for success are usually similar to Bandura's efficacy expectation in his discussion of self-efficacy, referring to individuals' beliefs about how well they will do on upcoming tasks, either in the immediate or longer term future. Subjective task values include dimension such as intrinsic interest (e.g., enjoyment) and utility (e.g., usefulness). From a social cognitive theory perspective, self-efficacy beliefs are shaped through a blend of personal experiences, vicarious experiences, social persuasion, and interpretations of physical and emotional condition (Bandura, 1994; Walker and Greene, 2009), and emotional arousal is an essential source of the self-efficacy formation (Bandura, 1997). As such, the sense of enjoyment experienced in relation to a given task situation should thus help people feel confident about their ability to successfully organize and execute a given course of action to solve a problem or accomplish a task (Bandura, 1994). Consequently, EVT and Bandura's self-efficacy theory serve as the major conceptual framework supporting the existence of the chain mediation process, i.e., enjoyment→self-efficacy→SRL→learning outcomes.

Empirical studies have also consistently shown that learning enjoyment and self-efficacy are critical determinants of learners' SRL strategy use and learning achievement. For example, Kim et al. (2015) study revealed that students with high levels of self-efficacy are more likely to devote efforts and use various SRL strategies during the process of learning. Bai and Guo (2019) also found that learning interest, self-efficacy, and growth mindset can significantly promote learners' SRL strategy use and impact their learning achievement. In Bai and Wang's (2020) study, English self-efficacy and intrinsic value were found to be positive predictors of students' self-regulated language learning strategies which in turn contribute to the improvement of English learning outcomes. Similarly, in their study of the role of the motivation variables of self-efficacy, enjoyment, and learning goal orientation in predicting the use of Web-based information systems, Yi and Hwang (2003) found that enjoyment significantly predicted students' self-efficacy. More recently, Hong et al.'s (2017) study on the interrelatedness between intrinsic motivation, online learning self-efficacy, flow experience and students' learning progress also showed that increasing intrinsic motivation (e.g., enjoyment) for language learning increases online learning self-efficacy and flow experience. These studies thus provide empirical evidence that foreshadows the mediation process, i.e., enjoyment→self-efficacy→SRL→learning outcomes. It needs to be pointed out, however, that previous studies that have investigated the associations among enjoyment, self-efficacy, SRL strategies, and performance have rarely included these variables in one model in relation to students' learning achievement in the language context.

## Present Research

Guided by the EVT (Wigfield and Eccles, 2000; Eccles, 2009), this study focused on examining in Chinese university EFL undergraduate students the interplay between English language self-efficacy, English enjoyment, technology-based SRL strategies, and their English learning outcomes. The students' English learning outcomes in this study refer to their score they obtained in the National College English Test-Band 4 (CET-4). In line with the expectancy and value theory perspective and empirical studies discussed above, the research model in **Figure 1** proposes that both English language self-efficacy and English enjoyment influence technology-based SRL strategies which in turn influence students' English learning outcomes, and that English enjoyment is hypothesized to predict students' English language self-efficacy.

Five research questions were thus addressed in this study:

1. What are the reported types and frequencies of Chinese university EFL students' technology-based SRL strategies?
2. How do different types of technology-based SRL strategies contribute to EFL students' English learning outcomes?
3. How does students' English language self-efficacy associate with technology-based SRL strategies?
4. To what extent does English enjoyment correlate with technology-based SRL strategies?
5. What are the relationships among students' English language self-efficacy, English enjoyment, technology-based SRL strategies, and English learning outcomes?

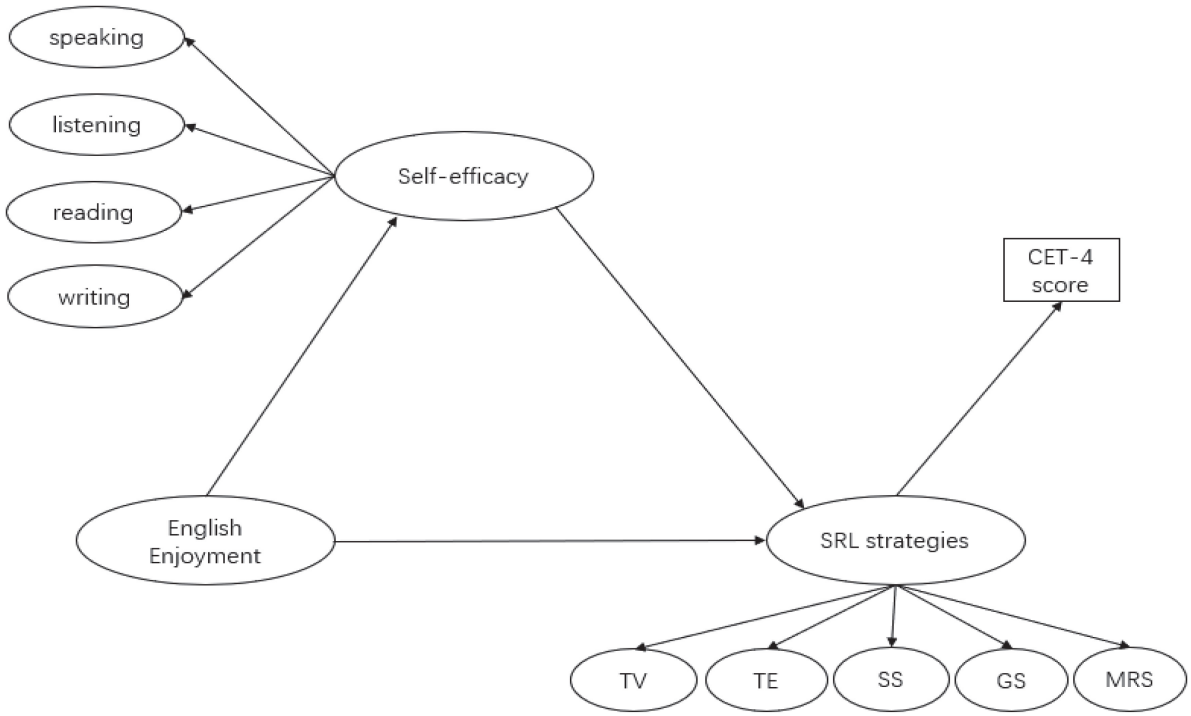


**FIGURE 1 |** Hypothesized structural model of the relationships among technology-based SRL strategies, English language self-efficacy, English enjoyment, and English learning outcomes (i.e., CET-4 score). SRL strategies, technology-based self-regulated English learning strategies; MRS, motivational regulation strategies; GS, goal setting and planning; SS, social strategies; TE, technology-based English song and movie learning; TV, technology-based vocabulary learning.

In order to answer the research questions, and based on the analysis of the literature, the following hypotheses are established accordingly:

> H1: Different types of technology-based SRL strategies display differential associations with English learning outcomes.
> H2: English language self-efficacy positively predicts technology-based SRL strategies.
> H3: English enjoyment positively predicts technology-based SRL strategies.
> H4: English enjoyment positively predicts students' English language self-efficacy.
> H5: There is a significant positive relationship between the use of technology-based SRL

strategies and English learning outcomes.

# MATERIALS AND METHODS

## Participants

A total of 525 undergraduate students forming a volunteer multidisciplinary sample from a university in Northern China were recruited. All participants were first-year and second-year undergraduate students, whose age ranged from 17 to 25 ($M$ = 20.50, $SD$ = 7.97). Most participants were majoring in accounting, finance, auditing, and management. Thus, female students ($n$ = 377) outnumbered male students ($n$ = 148). At the time of data collection, all the students had received formal English education for over 6 years and experienced College Entrance Examination in mainland China. English is a compulsory course for first-year and second-year students in the university where English teachers and students meet in class for 3–4.5 h per week in classrooms. All students in the university were required to pass CET4 before graduation.

## Procedure for Data Collection

The first author contacted the College English course instructors at the university about their willingness to include their students in this survey. Ethical approval was obtained from the university before the study was carried out. Those instructors who were willing to include their students in the survey were then asked to help distribute the online survey weblink to their students by any means that they normally used to contact their students. All participants were informed that their participation was voluntary and that they could withdraw from the study at any time. The participants were also informed that their responses to our survey would be kept confidential and that all data collected would be securely stored in the research center, and would only be used for research purposes. They completed the online survey outside the classroom on the platform of Wenjuanxing with no time limit. The average time the participants used in completing the online survey was 230.01 s.

## Instruments

### Technology-Based Self-Regulated English Learning Strategies Questionnaire

The technology-based SRL strategies questionnaire (TSRLSQ) was designed to assess participants' SRL strategy use in technology-using language learning conditions. Items in the questionnaire originated from three major sources: (a) constructs related to self-regulation or SRL strategies (e.g., cognitive, metacognitive, behavioral, and motivational regulation strategies) outlined by Pintrich and De Groot (1990); Zimmerman (2011), and Oxford (2013); (b) focus group interviews with 15 undergraduate students about the strategies they used when learning English in technology-using conditions; (c) existing instruments assessing students' technology use in language learning or SRL strategies (Barnard et al., 2009; Toffoli and Sockett, 2010; Lai et al., 2012; Lee et al., 2016). These processes resulted in a generation of 36 initial items. Two professional academics who were familiar with SRL strategies with technology were invited to assess the face and content validity of each item. An item was retained if both the two academics agreed that it is appropriate to be used to evaluate students' technology-based SRL in the Chinese EFL context. As a result of this validation process, 30 items were retained and slight modifications were made in the wordings of a few items in light of the two academics' comments.

The 30-item questionnaire in which items were scored on a 7-point Likert scale, ranging from 1 (not at all true of me) to 7 (very true of me), was then piloted to a sample of 155 university EFL students in China. Exploratory factor analyses (EFA) was conducted to examine the underlying factor structure of the TSRLSQ. The Bartlett's spherical test provided a significant chi-square value of 3,069.86 ($p < 0.001$), and the Kaiser-Meyer-Olkin (KMO) statistic was 0.92 exceeding the minimum adequacy value of 0.50 (Tabachnick and Fidell, 2007), meaning that the data were suitable for structure detection (Kaiser, 1958). The Laiser's eigenvalues-greater-than-1.00 criterion (Kaiser, 1960) and the scree plot (Raubenheimer, 2004) extracted a five-factor structure with 26 items, accounting for 69.62% of the total variance: (1) motivational regulation strategies (9 items); (2) goal setting and learning evaluation (5 items); (3) social strategies (4 items); (4) technology-based English song and movie learning (5 items); and (5) technology-based vocabulary learning (3 items).

The above 26-item five-factor structure of Chinese EFL students' TSRLSQ was cross-validated in the present study with a sample of 525 undergraduate students, suggesting an overall acceptable model fit with $X^2 = 1018.14$ (df = 281, $p < 0.001$); CFI = 0.91; TLI = 0.90; SRMR = 0.048; RMSEA = 0.07). Standardized factor loadings for CFA of the TSRLSQ ranged from 0.33 to 0.87. A scale analysis of the total 26 items in TSRLSQ revealed high internal consistency (Cronbach's alpha = 0.95). The Cronbach's alpha coefficients for the five factors were: 0.91 for motivational regulation strategies, 0.85 for goal setting and learning evaluation, 0.87 for social strategies, 0.86 for technology-based English song and movie learning, and 0.68 for technology-based vocabulary learning.

### English Language Self-Efficacy Questionnaire

The 16 items in the English language self-efficacy Questionnaire in this study were adapted from Author (2013) that measures students' English language self-efficacy in terms of four different domains: speaking, listening, reading, and writing, using a scale from 1 (not at all true of me) to 7 (very true of me). In the current study, CFA was also performed to confirm the four factor domains, and satisfying model fit indices were found with $X^2 = 436.49$ (df = 96, $p < 0.001$), CFI = 0.96, TLI = 0.95, RMSEA = 0.08, SRMR = 0.03. Standardized factor loadings for CFA ranged from 0.75 to 0.93. A Cronbach's alpha coefficient of 0.97 was found for the total items in the questionnaire. Furthermore, the Cronbach's alpha for the four domains of English language self-efficacy were:0.92 for speaking, 0.90 for listening, 0.93 for reading, and 0.93 for writing.

### English Enjoyment Questionnaire

The English Enjoyment Questionnaire in this study contains seven items adapted from the foreign language enjoyment questionnaire used in Dewaele and MacIntyre (2014) and Li et al. (2018), measuring classroom-based foreign language enjoyment. This questionnaire also adopts a 7-point Likert scale, ranging from 1 (not at all true of me) to 7 (very true of me). CFA was also performed to confirm the one factor structure of the questionnaire and good model fits were found with $\chi^2 = 20.10$ (df = 8, $p < 0.001$), CFI = 0.99, TLI = 0.98, SRMR = 0.03. The Cronbach's $\alpha$ for the single one factor of the questionnaire was 0.84.

### College English Test-Band 4 (CET-4)

In our present study, CET-4 was used to measure the participants' English learning outcomes. As the most influential English proficiency test throughout colleges in China (Jin, 2008), the CET is administered by the National College English Testing Committee on behalf of the Chinese Ministry of Education (Zheng and Cheng, 2008). CET aims to provide an objective evaluation of a student's overall English proficiency and positively affect EFL teaching at the tertiary level in China (State Education Commission, 1986). The CET-4 test-takers are undergraduate students in China except English majors. Each test takes 125 min to complete, with the total score of 710, containing four parts: writing (15%), listening comprehension (35%), reading comprehension (35%), and translation (15%) (Zheng and Cheng, 2008). The CET has been subjected to rigorous validation processes to ensure its high quality as an assessment tool for undergraduates (Yang and Weir, 1998).

## Data Analysis

Descriptive statistics (e.g., mean, standard deviation, and internal consistencies) for technology-based SRL strategies were calculated to answer the first research question. To answer the second research question regarding how students' self-regulated technology-using English learning strategies were related to English learning achievement, five linear regression analyses (Models 1A–E) and one multiple regression analysis (Model 2) were run with the five types of strategies as predictors and the learning achievement as the outcome variable. For Models 1A–E,

motivational regulation strategies, goal setting, and learning evaluation, social strategies, technology-based English song and movie learning, and technology-based vocabulary learning were regarded as predictors, respectively, in each model. For the one multiple regression analysis, the overall technology-based SRL strategy use was considered to be the predictor. This step was necessary, because the significance of exploring self-regulated technology-using English learning strategies is to see whether they are helpful for learning achievement. Adjusted R-squared values were reported as effect sizes.

To answer the third research question, Pearson product-moment correlation was carried out to assess the relationships between English self-efficacy, English enjoyment, and students' SRL strategies in technology-using conditions. All the analyses were performed with IBM SPSS 24.0 (SPSS Inc., Chicago, IL, United States). Finally, the data were subjected to AMOS using structural equation models (SEM) to investigate the structural relationships between SRL strategies, English language self-efficacy, English enjoyment, and English learning outcomes. The following model fit indices were used for evaluating the model fit (Hu and Bentler, 1999): the chi-square statistic ($\chi 2$) and its degrees of freedom (*df*), along with the associated *p*-value; the Comparative Fit Index (CFI) (a value equal to or greater than 0.90 indicates acceptable model fit); Tucker-Lewis Index (TLI) (a value equal to or greater than 0.90 indicates acceptable model fit); the Root Mean Square Error of Approximation (RMSEA) (a value between 0.05–0.08 indicates good fit); and the Standardized Root Mean-square Residual (SRMR) (a value less than 0.08 indicates good fit).

# RESULTS

## What Are the Reported Types and Frequencies of Chinese University EFL Students' Technology-Based Self-Regulated English Learning Strategies?

Confirmatory factory analysis analysis in this study confirmed the five factor structure of the TSRLSQ: (1) motivational regulation strategies; (2) goal setting and learning evaluation; (3) social strategies; (4) technology-based English song and movie learning; and (5) technology-based vocabulary learning. The means, standard deviations, and Cronbach's alpha reliability coefficients for the five types of SRL strategies were presented in **Table 1**. As suggested by Oxford (1990), in the case of a seven-point Liker scale, a variable mean in the range of 4.9–7.0 is usually considered to be high level, 3.5–4.8 medium level, and 1.0–3.4 low level (Guo and Wei, 2019). As such, among the five types of SRL strategies, technology-based vocabulary learning is the only type of strategies that was reported to be highly frequently used ($M$ = 5.26, $SD$ = 1.21), whereas the other four types of strategies, i.e., motivational regulation strategies ($M$ = 4.36, $SD$ = 1.24), English song and movie learning strategies ($M$ = 4.11, $SD$ = 1.34), goal setting and learning evaluation strategies ($M$ = 3.95, $SD$ = 1.30), and social strategies ($M$ = 3.77, $SD$ = 1.45), were reported to be used at a medium level.

The correlations among different types of technology-based SRL strategies, English language self-efficacy, English enjoyment, and learning achievement were also listed in **Table 1**. The results showed that EFL students' technology-based SRL strategies were positively correlated with each other (rs = 0.33–0.95, ps $< 0.001$). Furthermore, different types of technology-based SRL strategies, English language self-efficacy, and English enjoyment had positive correlations with learning achievement (rs = 0.18–0.38, $p < 0.001$), except technology-based vocabulary learning strategy ($r = 0.08$, $p > 0.05$).

## How Do Different Types of Technology-Based SRL Strategies Contribute to EFL Students' English Learning Outcomes?

Regression analyses of technology-based SRL strategies as predictors of English learning outcomes showed that motivational regulation strategies ($\beta = 0.24$, $p < 0.001$, $\Delta R^2 = 0.06$) (see model 1A in **Table 2**), goal setting and learning evaluation ($\beta = 0.20$, $p < 0.001$, $\Delta R^2 = 0.04$) (see model 1B in **Table 2**), social strategies ($\beta = 0.18$, $p < 0.001$, $\Delta R^2 = 0.03$) (see model 1C in **Table 2**), as well as English song and movie learning ($\beta = 0.21$, $p < 0.001$, $\Delta R^2 = 0.04$) (see model 1D in **Table 2**) were statistically significantly associated with students' English learning outcomes, whereas technology-based vocabulary learning ($\beta = 0.08$, $p > 0.05$, $\Delta R^2 = 0.004$) was not (see model 1E in **Table 2**). Additionally, when all five types of technology-based SRL strategies entered the equation simultaneously, it turned out that motivational regulation strategies ($\beta = 0.24$, $p = 0.006$, $\Delta R^2 = 0.06$) was the only significant predictor of the CET-4 score (see model 2 in **Table 2**).

## To What Extent Do Students' English Language Self-Efficacy and English Enjoyment Correlate With Different Types of Technology-Based SRL Strategies?

Pearson product-moment correlation (**Table 1**) between English language self-efficacy, English enjoyment, and students' technology-based SRL strategies showed that English language self-efficacy ($r = 0.63$, $p < 0.001$) and English enjoyment ($r = 0.64$, $p < 0.001$) positively correlated with the overall technology-based SRL strategies. Specifically, all the four domains of English self-efficacy (speaking, listening, reading, and writing) significantly positively correlated with the five different types of technology-based SRL strategies, among which technology-based vocabulary learning strategies had the weakest correlations with both the overall English language self-efficacy and the four different domains of English language self-efficacy (rs = 0.20–0.25, ps $< 0.001$). English enjoyment also positively correlated with the five different types of technology-based SRL strategies (rs = 0.41–0.61, ps $< 0.001$).

## What Are the Relationships Among Students' English Language Self-Efficacy, English Enjoyment, Technology-Based SRL Strategies, and English Learning Outcomes?

Structural equation models model suggested that the data fit the hypothesized model, $X^2 = 2894.97$ ($df = 1143$, $p < 0.001$); CFI = 0.90; TLI = 0.90; RMSEA = 0.05; SRMR = 0.06 (see **Figure 2**). The completely standardized parameter estimates of the significant correlations among the four variables are presented in **Figure 2**. Consistent with the finding of regression analyses and Pearson correlation analyses (**Table 1**), SRL strategies were statistically significantly associated with students' English learning outcomes ($\beta = 0.29$, $p < 0.001$). English language self-efficacy was statistically significantly linked with SRL strategies ($\beta = 0.35$, $p < 0.001$), and English enjoyment was statistically significantly associated with the use of SRL strategies ($\beta = 0.45$, $p < 0.001$). Moreover, the path from English enjoyment to English language self-efficacy was statistically significant ($p < 0.001$), with a standardized loading of 0.72. These variables explained 6.9% of the variance in EFL undergraduate students' English learning outcomes.

The mediating effect of English language self-efficacy was tested using bootstrapping approach. As shown in **Table 3**, the indirect effect of English enjoyment on technology-based SRL strategies via English language self-efficacy was equal to the product of the coefficients for each of the paths in the mediation chains (i.e., 0.72*0.35 = 0.25). The 95% bias-corrected confidence interval for the mediated effect was between 0.17 and 0.34, with a $p$-value at 0.001 for the two-tailed significance test and the standard error at 0.04. The total effect of English enjoyment on technology-based SRL strategy use was 0.25 + 0.45 = 0.70. After controlling for English language self-efficacy, the direct relationship between English enjoyment and SRL strategies was also significant ($\beta = 0.45$, $p < 0.01$). Therefore, the effect of English enjoyment on technology-based SRL strategies was partially mediated by English language self-efficacy.

**TABLE 1 |** Correlation coefficients between the latent variables.

| | Mean | SD | Cronbach's alpha | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SRL | 4.25 | 1.13 | 0.95 | 1.00 | | | | | | | | | | | | |
| MRS | 4.36 | 1.24 | 0.91 | 0.95*** | 1.00 | | | | | | | | | | | |
| GS | 3.95 | 1.30 | 0.85 | 0.88*** | 0.77*** | 1.00 | | | | | | | | | | |
| SS | 3.77 | 1.45 | 0.87 | 0.85*** | 0.78*** | 0.69*** | 1.00 | | | | | | | | | |
| TE | 4.11 | 1.34 | 0.86 | 0.88*** | 0.78*** | 0.69*** | 0.70*** | 1.00 | | | | | | | | |
| TV | 5.26 | 1.21 | 0.68 | 0.62*** | 0.52*** | 0.53*** | 0.33*** | 0.49*** | 1.00 | | | | | | | |
| Efficacy | 3.36 | 1.33 | 0.97 | 0.63*** | 0.60*** | 0.57*** | 0.58*** | 0.58*** | 0.24*** | 1.00 | | | | | | |
| Speaking | 3.34 | 1.46 | 0.92 | 0.60*** | 0.57*** | 0.53*** | 0.58*** | 0.55*** | 0.20*** | 0.92*** | 1.00 | | | | | |
| Listening | 3.34 | 1.39 | 0.9 | 0.59*** | 0.57*** | 0.53*** | 0.53*** | 0.55*** | 0.21*** | 0.94*** | 0.83*** | 1.00 | | | | |
| Reading | 3.41 | 1.42 | 0.93 | 0.58*** | 0.55*** | 0.53*** | 0.49*** | 0.53*** | 0.25*** | 0.93*** | 0.77*** | 0.84*** | 1.00 | | | |
| Writing | 3.35 | 1.44 | 0.93 | 0.58*** | 0.55*** | 0.52*** | 0.54*** | 0.53*** | 0.25*** | 0.93*** | 0.80*** | 0.81*** | 0.85*** | 1.00 | | |
| Enjoyment | 4.76 | 1.12 | 0.84 | 0.64*** | 0.61*** | 0.55*** | 0.51*** | 0.57*** | 0.41*** | 0.61*** | 0.60*** | 0.55*** | 0.56*** | 0.55*** | 1.00 | |
| CET4 | 402.80 | 54.40 | | 0.23*** | 0.24*** | 0.20*** | 0.18*** | 0.21*** | 0.08 | 0.38*** | 0.37*** | 0.31*** | 0.37*** | 0.36*** | 0.32*** | 1.00 |

****p < 0.001. MRS, motivational regulation strategies; GS, goal setting and learning evaluation; SS, social strategies; TE, technology-based English song and movie learning; TV, technology-based vocabulary learning.*

**TABLE 2 |** Regression models of technology-based SRL strategies as predictors for English learning outcomes.

| | | English learning outcomes | | |
|---|---|---|---|---|
| **Model** | | **$\Delta R^2$** | **SE(B)** | **β** |
| 1A | Motivational regulation strategies | 0.06 | 1.86 (10.64) | 0.24*** |
| 1B | Goal-setting and learning evaluation | 0.04 | 1.79 (8.40) | 0.20*** |
| 1C | Social strategies | 0.03 | 1.62 (6.76) | 0.18*** |
| 1D | Technology-based English song and movie learning | 0.04 | 1.74 (8.44) | 0.21*** |
| 1E | Technology-Based vocabulary learning | 0.004 | 1.96 (3.37) | 0.08 |
| 2 | Motivational regulation strategies | 0.06 | 3.84 (10.68) | 0.24** |
| | Goal-setting and learning evaluation | | 3.00 (2.63) | 0.06 |
| | Social strategies | | 2.71 (−2.56) | −0.07 |
| | Technology-based English song and movie learning | | 2.93 (2.73) | 0.07 |
| | Technology-based vocabulary learning | | 2.36 (−4.26) | −0.10 |

***p < 0.01; ***p < 0.001.*

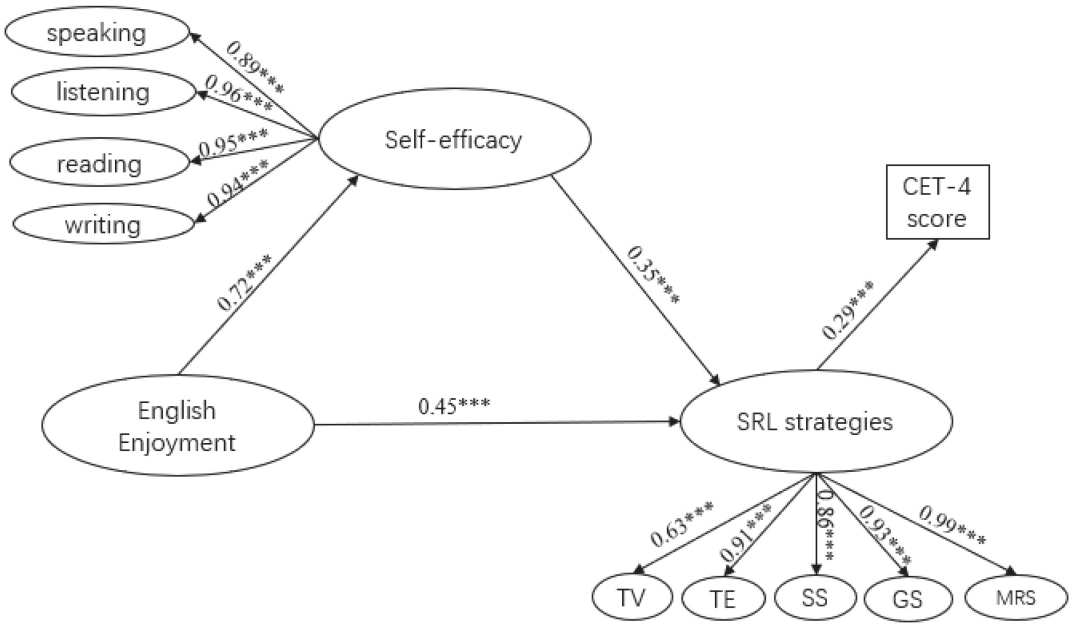


**FIGURE 2 |** Structural model of the relationships between technology-based SRL strategies, English language self-efficacy, English enjoyment, and English learning outcomes (i.e., CET-4). ***$p < 0.001$. SRL strategies, technology-based self-regulated English learning strategies; MRS, motivational regulation strategies; GS, goal setting and planning; SS, social strategies; TE, technology-based English song and movie learning; TV, technology-based vocabulary learning.

The indirect effect of English language self-efficacy on English learning outcomes through technology-based SRL strategies was equal to 0.35*0.29 = 0.10. The 95% bias-corrected confidence interval for the mediated effect was between 0.06 and 0.15, with a *p*-value at 0.001 for the two-tailed significance test and the standard error at 0.02. In addition, the indirect effect of English enjoyment on English learning outcomes via technology-based SRL strategies was 0.45*0.29 = 0.13. The indirect effect of English enjoyment on English learning outcomes through English language self-efficacy and technology-based SRL strategies was 0.72*0.35*0.29 = 0.07. Hence the total effect of English enjoyment on English learning outcomes was equal to 0.13 + 0.07 = 0.20. The 95% bias-corrected confidence interval for the mediated effect was between 0.13 and 0.28, with a *p*-value at 0.002 for the two-tailed significance test and the standard error at 0.04.

**TABLE 3 |** Standardized direct, indirect, and total effects for structural model.

| Predicted variable | Predictor variable | Direct effect | Indirect effect | Total effect |
|---|---|---|---|---|
| Learning outcomes | Technology-based SRL | 0.29** | | 0.29** |
| | English self-efficacy | | 0.10** | 0.10** |
| | English enjoyment | | 0.20** | 0.20** |
| Technology-based SRL | English self-efficacy | 0.35** | | 0.35** |
| | English enjoyment | 0.45** | 0.25** | 0.70** |
| English self-efficacy | English enjoyment | 0.72** | | 0.72** |

*$^{**}p < 0.01$.*

## DISCUSSION

One shortcoming of the previous studies on technology-assisted language learning is that lack of domain- and situation-specific instruments might result in failure to capture patterns of learning strategies that are idiosyncratic to technology-using conditions particularly in an EFL context. Building on SRL research in educational psychology and computer-assisted L2 learning research, this study investigated Chinese university students' technology-assisted SRL strategies in relation to their English language self-efficacy, English enjoyment, and English learning outcomes. The study has confirmed five types of English learning strategies in technology-using conditions among Chinese university EFL students: (1) motivational regulation strategies; (2) goal setting and learning evaluation; (3) social strategies; (4) technology-based English song and movie learning; and (5) technology-based vocabulary learning. We concur with the argument that understanding the SRL strategies that students use has implications for classroom-based language teaching and learning (Zimmerman and Schunk, 2011). For example, technology-based vocabulary learning was reported to be the most frequently used strategy in this study; this strategy, however, unlike other types of SRL strategies, was found to have no positive impact on students' English learning outcomes. In our communication with some study participants during the interviews in the course of questionnaire construction, we noted that learning English vocabulary through use of lexical apps on

mobile phones was prevalent among the students. Given the results of this study, there is a pressing need to engage students in a variety of more meaningful vocabulary learning activities both inside and outside the classroom.

In this study, technology-based social strategies were reported to be the least frequently used ones among all five SRL strategies, which was somewhat unexpected because one of the great educational potentials of technology-assisted language learning environments is that they provide learners with interaction opportunities and easy access to authentic language input via communicating with native speakers (Golonka et al., 2014). One possible explanation might be that the students in this study were required to take CET-4 before graduation, which usually results in tons of mechanical CET-4 drilling exercises rather than social communicative learning activities for the students. Consequently, some students might believe their major goal of learning English is to pass the CET-4, and therefore, their technology-based English learning might be limited to lexical app-based words memorization and practice of multiple choice type of questions. An important implication of this result is thus that students should be encouraged to participate in more social language learning activities such as cooperative learning. Since teachers play an important role in engaging students in social or experiential learning activities, there is a pressing need to integrate communicative approach into the EFL curriculum so as to reduce the teacher prolonged control of the classroom and to maximize student responsibility and involvement in the social learning activities.

The regression analyses and SEM analyses in this study showed a positive relationship between technology-based SRL strategies and students' English learning outcomes, supporting Hypothesis 5. This result is congruent with prior research findings on the positive relationship between SRL strategies and academic achievement in second language acquisition (e.g., Xiao and Yang, 2019; Bai and Wang, 2020; Teng and Zhang, 2020). An important pedagogical implication of this result is that teachers can raise students' awareness of the importance of technology-based SRL strategies, introduce to them a wide variety of learning strategies, and model the use of these strategies in the classroom. In particular, the results of regression analyses highlight the importance of motivational regulation strategies as it is the only significant and positive predictor of CET-4 scores among the five types of technology-based SRL strategies. This evidenced the differential associations different types of technology-based SRL strategies display with English learning outcomes, supporting Hypothesis 1. Our result echoes findings of previous research on the vital role of motivational regulation strategies in students' learning achievement (Wolters, 2003; Teng and Zhang, 2016) and classroom performance (Wolters, 1999). Note that the current TESOL Technology Standards for Language Learners (Healey et al., 2018) are more often focused on decontextualized knowledge and skills in technology use and evaluation of technology-based tools as aids in the development of their language learning competence with little regard to the role of motivation-regulation factors that are critical to sustaining students' use of technology, and that impact on students' learning outcomes. This study suggests that although knowledge and skills in using various technological tools are important, it is equally important to empower students, as agents, to develop and increase their strategic awareness of motivational regulation in optimizing effectiveness of their technology use in the target language learning. As teachers occupy a central role in promoting technology-assisted language learning among students, it is imperative to provide teachers with professional training to enrich their knowledge and skills in the utilization of motivational regulation strategies.

Furthermore, the SEM analyses in this study showed that technology-based SRL strategies mediated the relationship between English enjoyment and English learning outcomes. The findings indicated that students who enjoy English learning are more likely to control their efforts and regulate their learning process, which in turn contributes to the improvement of learning outcomes, supporting Hypotheses 2–4. The positive associations between English enjoyment, technology-based SRL strategies, and learning outcomes provide empirical evidence for the argument that motivational beliefs (e.g., enjoyment for learning) are major determinants of individual's behavioral choices and learning achievement (Wigfield and Eccles, 2000). The results thus echo prior research findings which showed that students with high levels of learning interests tended to use more SRL strategies such as effort regulation, and get higher English test scores than their peers with lower enjoyment for English learning (Bai and Wang, 2020). Our result also lends support to Goetz et al.'s (2008) observation that academic enjoyment is strongly related to domain-specific academic self-concepts which, in turn, impact on students' cognitions, behavior, and ultimate success in the academic domain. While the importance of enjoyment as an essential affective factor in L2 acquisition has been well acknowledged in the literature, its relationships with other learner variables have been under-researched particularly in an SRL context. This study thus adds to the limited existing research by showing that the role that enjoyment plays is two-dimensional in the current study as it has a strong relationship with SRL strategies and that this relationship is mediated by English language self-efficacy. The findings suggest a need to nurture students' English language enjoyment and self-efficacy, which can be realized through providing diversified English learning activities (Bai and Wang, 2020), or through actively developing pedagogical practices that promote a positive self-concept and self-confidence among students so that they find the time spent there is a constant source of satisfaction, enjoyment, and self-efficacy (Dörnyei and Murphy, 2003).

The finding that a strong and positive relationship exists between English language self-efficacy and the technology-based SRL strategies leads us to concur with existing studies (e.g., Kim et al., 2015; Bai and Wang, 2020; Wang and Zhan, 2020) that postulate that students with higher level of self-efficacy tend to be more self-regulated than their peers with low self-efficacy profiles, reinforcing the consensus in the literature that positive beliefs about one's learning capability tend to result in carrying out a learning task more readily, working harder, or persisting longer when learners encountered difficulties. Pedagogically, teachers can emphasize skill development and effective learning methods, as well as provide specific feedback and encouragement to enhance students' confidence and sense of competence in authentic mastery experience (Liem et al., 2008). Particularly

noteworthy is that among the four domains of English language self-efficacy examined in this study, English speaking competence appeared to be the most significant predictor of students' use of a variety of technology-based SRL strategies. An important pedagogical implication of this result is that EFL teachers need to provide students with a fairly wide repertoire of speaking practice opportunities to help them maintain a positive belief about their English speaking competence, which in turn will likely be a vital driving force for students' engagement in a range of SRL activities.

Although our study is the first to empirically investigate the associations among enjoyment, self-efficacy, SRL strategies in one model in relation to students' learning achievement in the second language context, a number of limitations should be acknowledged and direction for future research needs to be provided. First, the present study adopted a cross-sectional design and the results only indicate associations or relationships between variables. As such, inferences about the causal relations among variables which require experimental design are not drawn. In future research that aims to examine causal relationships among these variables, longitudinal design that involves data collections in different points over a period of time should be adopted in order to draw inferences of causal relations. Second, the fact that participants in this study were from a few majors of one university may limit the generalizability of the findings. Future research needs to include students from different types of educational institutions and from a wider range of disciplines so as to gain more robust results concerning the dynamic relationships between enjoyment, self-efficacy, technology-assisted SRL engagement, and students' learning achievement. Third, in this study, data were collected by means of self-reports, which might be susceptible to response bias as is the case with other survey-based investigations. While self-report has been widely used as a valid method for exploring student perceptions and feelings, future research can adopt additional objective measures, such as classroom observations, to minimize the limitations associated with self-reported data. Finally, our research model only explained 6.9% of the variance in EFL undergraduate students' English learning outcomes, and a large portion of the variation remains uncaptured. Future research needs to explore potential variables that were not explored in the current model, such as students' anxiety toward using technologies for English learning and the role of extrinsic motivation in driving students to learn English.

## CONCLUSION

This study contributes to the knowledge about Chinese EFL undergraduate students' SRL strategies in technology-using conditions. The results of the study add to the literature that considers how technology-based SRL strategies are associated with students' language learning achievement. From a theoretical perspective, our research extends SRL theories to technology-using language learning conditions, particularly with respect to the significant role of English enjoyment and English language self-efficacy, and in relation to students' English learning outcomes. Pedagogically, awareness of the complex interrelationships among SRL strategies, English enjoyment, English language self-efficacy, and learning outcomes is helpful for educators to clearly understand what actually motivates and empowers students' self-directed technology use for learning and the quality of this technology-based learning process. It is thus important for educators to create a pleasant and inspiring environment that empowers students in self-regulation of their technology-facilitated English learning practices so that they experience learning success and satisfaction inside and outside the classroom.

While this study is one of the pioneer studies conducted in an EFL context which addressed research gaps from previous SRL studies, it examined associations between technology-based SRL and students' English learning outcomes in relation to only two learner variables, i.e., English language self-efficacy and English enjoyment. We understand the limitation of using cross-sectional data and we cannot draw cause-and-effect conclusions without a longitudinal design. We thus cautioned our readers of generalizing our results to cause-and-effect, especially the reciprocal relationship between enjoyment and self-efficacy (Malanchini et al., 2017). Future studies on technology-enhanced language learning can also investigate some variables that were not contained in the current model, such as students' previous learning experiences and their willingness to use technology for language learning. As the items used in our TSRLSQ are newly proposed, further assessment with other populations is needed to provide evidence of external aspects of the construct validity of responses to the scale. In addition, this study relied solely on the use of self-reported data which is inherently subject to bias, and the findings therefore constitute only an overall picture of the participants' experiences in terms of the measured variables. Future extensions of the research can make use of qualitative observational research to corroborate the statistical evidence reported in this study.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Macau Ethics Assessment Committee. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

ZA and ZG conceived the idea and developed the materials. ZA carried out the data collection. ZA, SL, and ZG took the lead in writing the manuscript. CW, ZG, and HL provided critical feedback. All authors read and approved the final manuscript.

## FUNDING

The research was funded by the University of Macau.

## REFERENCES


Ainley, M., and Hidi, S. (2014). "Interest and enjoyment," in *International Handbook of Emotions in Education*, eds R. Pekrun, and L. Linnenbrink-Garcia, (New York, NY: Routledge), 205–227.

An, Z., Gan, Z., and Wang, C. (2020). Profiling Chinese EFL students' technology-based self-regulated English learning strategies. *PLoS One* 15:e0240094. doi: 10.1371/journal.pone.0240094

Bai, B., and Guo, W. (2019). Motivation and self-regulated strategy use: relationships to primary school students' English writing in Hong Kong. *Lang. Teach. Res.* 1–22. doi: 10.1177/1362168819859921

Bai, B., and Wang, J. (2020). The role of growth mindset, self-efficacy and intrinsic value in self-regulated learning and English language learning achievements. *Lang. Teach. Res.* 1–22. doi: 10.1177/1362168820933190

Bandura, A. (1994). *Self-Efficacy: The Exercise of Control*. New York, NY: Freeman.

Bandura, A. (1997). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Bull.* 84, 191–215. doi: 10.1037/0033-295x.84.2.191

Bandura, A. (2006). "Guide for constructing self-efficacy scales," in *Self-Efficacy Beliefs of Adolescents*, eds F. Pajares, and T. S. Urdan, (Greenwich, CT: Information Age Publishing), 307–337.

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., and Lai, S. L. (2009). Measuring self-regulation in online and blended learning environments. *Int. Higher Educ.* 12, 1–6. doi: 10.4324/9781315814605-1

Bassi, M., Steca, P., Delle Fave, A., and Caprara, G. V. (2007). Academic self-efficacy beliefs and quality of experience in learning. *J. Youth Adolesc.* 36, 301–312. doi: 10.1007/s10964-006-9069-y

Bekleyen, N., and Hayta, F. (2015). "Language learning in new era: do mobile phones help?," in *The Practice of Foreign Language Teaching: Theories and Applications*, ed. A. Akbarov, (New Castle: Cambridge Scholars Publishing), 434–445.

Benson, P. (2001). *Teaching and Researching Learner Autonomy in Language Learning*. London: Longman.

Botero, G. G., Questier, F., and Zhu, C. (2018). Self-directed language learning in a mobile-assisted, out-of-class context: do students walk the talk? *Comput. Assist. Lang. Learn.* 32, 71–97. doi: 10.1080/09588221.2018.1485707

Burston, J. (2015). Twenty years of MALL project implementation: a meta-analysis of learning outcomes. *ReCALL* 27, 4–20. doi: 10.1017/s0958344014000159

Butler, D. L., and Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Rev. Educ. Res.* 65, 245–274. doi: 10.3102/00346543065003245

Carneiro, R., Lefrere, P., and Steffens, K. (2007). *Self-Regulated Learning in Technology Enhanced Learning Environments: A European Review*. Kaleidoscope Network of Excellence. Available online at: http://www.lmi.ub.es/taconet/documents/srlinteles3.pdf

Chang, C., and Chang, C.-K. (2014). Developing students' listening metacognitive strategies using online videotext self-dictation-generation learning activity. *EuroCall Rev.* 22, 3–19. doi: 10.4995/eurocall.2014.3636

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York, NY: Harper and Row.

Demouy, V., Jones, A., Kan, Q., Kukulska-Hulme, A., and Eardley, A. (2016). Why and how do distance learners use mobile devices for language learning? *EuroCall Rev.* 24, 10–24. doi: 10.4995/eurocall.2016.5663

Dewaele, J.-M., John, W., Kazuya, S., and Livia, D. (2017). Foreign language enjoyment and anxiety: the effect of teacher and learner variables. *Lang. Teach. Res.* 22, 676–697. doi: 10.1177/1362168817692161

Dewaele, J. M., and MacIntyre, P. D. (2016). "Foreign language enjoyment and foreign language classroom anxiety: the right and left feet of FL learning?," in *Positive Psychology in SLA*, eds P. D. MacIntyre, T. Gregersen, and S. Mercer, (Bristol: Multilingual Matters), 215–236. doi: 10.21832/9781783095360-010

Dewaele, J. M., and MacIntyre, P. D. (2014). The tow faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Stud. Second Lang. Learn. Technol.*, 4, 237–274. doi: 10.14746/ssllt.2014.4.2.5

Dörnyei, Z., and Murphy, T. (2003). *Group Dynamics in the Language Classroom*. Cambridge: Cambridge University Press.

Eccles, J. S. (2009). Who am I and what am I going to do with my Life? Personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368

Gan, Z., Hu, G., Nang, H., Wang, W., and An, Z. (2020). Feedback behaviour and preference in university academic English courses: associations with English language self-efficacy. *Assess. Eval. Higher Educ.* 1–16. doi: 10.1080/02602938.2020.1810628

Godwin-Jones, R. (2018). *Using Mobile Devices in the Language Classroom. Cambridge Papers in ELT Series*. Cambridge: Cambridge University Press.

Goetz, T., Frenzel, A. C., Hall, N. C., and Pekrun, R. (2008). Antecedents of academic emotions: testing the internal/external frame of reference model for academic enjoyment. *Contemp. Educ. Psychol.* 33, 9–33. doi: 10.1016/j.cedpsych.2006.12.002

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., and Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Comput. Assist. Lang. Learn.* 27, 70–105. doi: 10.1080/09588221.2012.700315

Guo, W., and Wei, J. (2019). Teacher feedback and students' self-regulated learning in mathematics: a study of Chinese secondary students. *Asia Pacific Educ. Res.* 28, 165–175.

Healey, D., Hanson-Smith, E., Hubbard, P., Ioannou-Georgious, S., Kessler, G., and Ware, P. (2018). *TESOL Technology Standards: Description, Implementation, Integration*. Alexandria, VA: TESOL.

Holec, H. (1981). *Autonomy in Foreign Language Learning*. Oxford: Pergamon Press.

Hong, J. C., Hwang, M. Y., Tai, K. H., and Lin, P. H. (2017). Intrinsic motivation of Chinese learning in predicting online learning self-efficacy and flow experience relevant to students' learning progress. *Comput. Assist. Lang. Learn.* 30, 552–574. doi: 10.1080/09588221.2017.1329215

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6, 1–55. doi: 10.1080/10705519909540118

Jin, Y. (2008). Powerful test, powerless test designers? Challenges facing the college English test. *CELEA J.* 31, 3–11.

Jung, H. J. (2015). Fostering an English teaching environment: factors influencing English as a foreign language teachers' adoption of mobile learning. *Informat. Educ.* 14, 219–241. doi: 10.15388/infedu.2015.13

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 187–200. doi: 10.1007/bf02289233

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* 20, 141–151. doi: 10.1177/001316446002000116

Kim, D. H., Wang, C., Ahn, H. S., and Bong, M. (2015). English language learners' self-efficacy profiles and relationship with self-regulated learning strategies. *Learn. Individ. Differ.* 38, 136–142. doi: 10.1016/j.lindif.2015.01.016

Kukulska-Hulme, A. (2016). "Mobile assistance in language learning: acritical appraisal," in *The International Handbook of Mobile-Assisted Language Learning*, eds A. Palalas, and M. Ally, (Beijing: China Central Radio & TV University Press), 138–160.

Lai, C., and Gu, M. (2011). Self-regulated out-of-class language learning with technology. *Comput. Assist. Lang. Learn.* 24, 317–335. doi: 10.1080/09588221.2011.568417

Lai, C., Hu, X., and Lyu, B. (2018). Understanding the nature of learners' out-of-class language learning experience with technology. *Comput. Assist. Lang. Learn.* 31, 114–143. doi: 10.1080/09588221.2017.1391293

Lai, C., Wang, Q., and Lei, J. (2012). What factors predict undergraduate students' use of technology for learning? A case from Hong Kong. *Comput. Educ.* 59, 569–579. doi: 10.1016/j.compedu.2012.03.006

Lake, J. (2013). "Positive L2 self: linking positive psychology with L2 motivation," in *Language Learning Motivation in Japan (Second Language Acquisition)*, eds M. T. Apple, D. Da Silva, and T. Fellner, (Bristol: Multilingual Matters).

Lee, C., Yeung, A. S., and Ip, T. (2016). Use of computer technology for English language learning: do learning styles, gender, and age matter? *Comput. Assist. Lang. Learn.* 29, 1035–1051. doi: 10.1080/09588221.2016.1140655

Li, C., Jiang, G., and Dewaele, J.-M. (2018). Understanding Chinese high school students' foreign language enjoyment: validation of the Chinese version of the Foreign Language Enjoyment scale. *System* 76, 183–196. doi: 10.1016/j.system.2018.06.004

Liem, A. D., Lau, S., and Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemp. Educ. Psychol.* 33, 486–512. doi: 10.1016/j.cedpsych.2007.08.001

MacIntyre, P. D., and Gregersen, T. (2012). "Affect: the role of language anxiety and other emotions in language learning," in *Language Learning Psychology: Research, Theory and Pedagogy*, eds S. Mercer, S. Ryan, and M. Williams, (Basingstoke: Palgrave), 103–118. doi: 10.1057/9781137032829_8

Malanchini, M., Wang, Z., Voronin, I., Schenker, V. J., Plomin, R., Petrill, S. A., et al. (2017). Reading self-perceived ability, enjoyment and achievement: a genetically informative study of their reciprocal links over time. *Dev. Psychol.* 53, 698–712. doi: 10.1037/dev0000209

Mills, N., Pajares, F., and Herron, C. (2007). Self-efficacy of college intermediate French students: relation to motivation and achievement. *Lang. Learn.* 57, 417–442. doi: 10.1111/j.1467-9922.2007.00421.x

Nicol, D. J., and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud. Higher Educ.* 31, 199–218. doi: 10.1080/03075070600572090

Oxford, R. L. (1990). *Language Learning Strategies: What Every Teacher Should Know*. New York, NY: Newbury House.

Oxford, R. L. (2013). *Teaching and Researching Language Learning Strategies*, 2nd Edn. Harlow: Pearson.

Pajares, F. (2009). "Motivational role of self-efficacy beliefs in self-regulated learning," in *Motivation and Self-Regulated Learning: Theory, Research, and Applications*, eds D. H. Schunk, and B. J. Zimmerman, (New York, NY: Routledge), 111–139.

Panadero, E., Jonsson, A., and Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: four meta-analyses. *Educ. Res. Rev.* 22, 74–98. doi: 10.1016/j.edurev.2017.08.004

Peters, M., Weinberg, A., and Sarma, N. (2009). To like or not to like! student perceptions of technological activities for learning French as a second language at five Canadian universities. *Can. Modern Lang. Rev.* 65, 869–896. doi: 10.3138/cmlr.65.5.869

Piniel, K., and Albert, A. (2018). Advanced learners' foreign language-related emotions across the four skills. *Second Lang. Learn. Teach.* 8, 127–148. doi: 10.14746/ssllt.2018.8.1.6

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educ. Psychol Rev.* 16, 385–407. doi: 10.1007/s10648-004-0006-x

Pintrich, P. R., and De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *J. Educ. Psychol.* 82, 33–40. doi: 10.1037/0022-0663.82.1.33

Pintrich, P. R., Smith, D., Garcia, T., and McKeachie, W. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor, MI: University of Michigan.

Raubenheimer, J. (2004). An item selection procedure to maximize scale reliability and validity. *SA J. Ind. Psychol.* 30, 59–64.

Ryan, R. M., Connell, J. P., and Plant, R. W. (1990). Emotions in non-directed text learning. *Learn. Individ. Differ.* 2, 1–17.

Schunk, D. H., Pintrich, P. R., and Meece, M. L. (2008). *Motivation in Education: Theory, Research, and Applications*, 3rd Edn. Upper Saddle River, NJ: Pearson.

Schunk, D. H., and Zimmerman, B. J. (2012). *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. Mullica, NJ: Routledge.

State Education Commission (1986). *College English Teaching Syllabus (For College and University Students of Arts and Science)*. Shanghai: Shanghai Foreign Language Education Press.

Steel, C. H., and Levy, M. (2013). Language students and their technologies: charting the evolution 2006-2011. *Eur. Assoc. Comput. Assist. Lang Learn.* 25, 306–320.

Su, Y., Zheng, C., Liang, J.-C., and Tsai, C. C. (2018). Examining the relationship between English language learners' online self-regulation and their self-efficacy. *Austr. J. Educ. Technol.* 34, 105–121.

Tabachnick, B. G., and Fidell, L. S. (2007). *Using Multivariate Statistics*, 5th Edn. Boston, MA: Allyn & Bacon/Pearson Education.

Teng, L. S., and Zhang, L. J. (2016). Fostering strategies learning: the development and validation of the writing strategies for motivational regulation questionnaire (WSMRQ). *Asia Pacific Edu. Res.* 25, 123–134.

Teng, L. S., and Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: can self-regulated learning strategies-based writing instruction make a difference? *J. Second Lang. Writ.* 48, 1–16.

Toffoli, D., and Sockett, G. (2010). How non-specialist students of English practice informal learning using Web 2.0 tools. *Asp* 58, 1–15.

Tseng, S.-S., and Yeh, H.-C. (2019). The impact of video and written feedback on student preferences of English speaking practice. *Lang. Learn. Technol.* 23, 145–158.

Walker, C. O., and Greene, B. A. (2009). The relations between student motivational beliefs and cognitive engagement in high school. *J. Educ. Res.* 102, 463–472. doi: 10.3200/joer.102.6.463-472

Wang, C., and Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Q.* 51, 931–947. doi: 10.1002/tesq.355

Wang, C., Schwab, G., Fenn, P., and Chang, M. (2013). Self-efficacy and self-regulated learning strategies for English Language learners: comparison between chinese and german college students. *J. Educ. Dev. Psychol.* 3, 173–191.

Wang, H., and Chen, C. W. (2020). Learning English from YouTubers: english L2 learners' self-regulated language learning on YouTube. *Innovat. Lang. Learn. Teach.* 14, 333–346. doi: 10.1080/17501229.2019.1607356

Wang, W., and Zhan, J. (2020). The relationship between english language learner characteristics and online self-regulation: a structural equation modeling approach. *Sustainability* 12:3009. doi: 10.3390/su12073009

Wigfield, A., and Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015

Winke, P., and Goertler, S. (2008). Did we forget someone? Students' computer access and literacy for CALL. *Calico J.* 25, 482–509. doi: 10.1558/cj.v25i3.482-509

Winne, P. H. (2011). "A cognitive and metacognitive analysis of self-regulated learning," in *Handbook of Self-Regulation of Learning and Performance*, eds B. J. Zimmerman, and D. H. Schunk, (New York, NY: Routledge/Taylor & Francis), 15–32.

Wolters, C. (1999). The relationship between high school students' motivational regulation and their use of learning strategies, effort, and classroom performance. *Learn. Individ. Differ.* 11, 281–299. doi: 10.1016/s1041-6080(99)80004-1

Wolters, C. (2003). Regulation of motivation: evaluating an underemphasized aspect of self-regulated learning. *Educ. Psychol.* 38, 189–205. doi: 10.1207/s15326985ep3804_1

Wong, M. S. (2005). Language learning strategies and language self-efficacy: investigating the relationship in Malaysia. *RELC J.* 36, 245–269. doi: 10.1177/0033688205060050

Xiao, Y., and Yang, M. (2019). Formative assessment and self-regulated learning: how formative assessment supports students' self-regulation in English language learning. *System* 81, 39–49. doi: 10.1016/j.system.2019.01.004

Yang, H., and Weir, C. J. (1998). *The CET Validation Study*. Shanghai: Shanghai Foreign Language Education Press.

Yi, M. Y., and Hwang, Y. (2003). Prediction the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *Int. J. Hum. Comput. Stud.* 59, 431–449. doi: 10.1016/s1071-5819(03)00114-9

Zheng, Y., and Cheng, L. (2008). Test review: college english test (CET) in China. *Lang. Test.* 25, 408–417. doi: 10.1177/0265532208092433

Zimmerman, B. J. (2011). "Motivational sources and outcomes of self-regulated learning and performance," in *Handbook of Self-Regulation of Learning and Performance*, eds B. J. Zimmerman, and D. H. Schunk, (Mahwah, NJ: Lawrence Erlbaum), 49–64.

Zimmerman, B. J., and Schunk, D. H. (2001). *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*. Mullica, NJ: Routledge.

Zimmerman, B. J., and Schunk, D. H. (2011). "Self-regulated learning and performance: An introduction and an overview," in *Educational Psychology Handbook Series. Handbook of Self-Regulation of Learning and Performance*, eds B. J. Zimmerman, and D. H. Schunk, (New York, NY: Routledge/Taylor & Francis Group), 1–12.

Zuffianò, A., Alessandri, G., Gerbino, M., Kanacri, B. P. L., Di Giunta, L., Milioni, M., et al. (2013). Academic achievement: the unique contribution of self-efficacy beliefs in self-regulated learning beyond intelligence, personality traits, and self-esteem. *Learn. Individ. Differ.* 23, 158–162. doi: 10.1016/j.lindif.2012.07.010

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# APPENDIX A. TECHNOLOGY-BASED SELF-REGULATED ENGLISH LEARNING STRATEGIES QUESTIONNAIRE

## MRS (Motivational Regulation Strategies)

1. I select and use appropriate technological tools to improve the areas I'm weak in.
2. I use technologies outside the classroom to access authentic materials in English.
3. I search related materials online when I have difficulties in the process of studying English.
4. I seek opportunities through technological resources to practice my oral English.
5. I use technologies to help me sustain/enhance interest in learning English.
6. I use technologies (APPs or websites) to make the English learning task more interesting.
7. I use mobile devices to enhance my willingness to participate in English social events.
8. Sometimes I look through the visual and vivid courseware to arouse my interest in English learning.
9. When I feel bored with learning English, I adopt technological resources to decrease the boredom and increase the enjoyment.

## GS (Goal Setting and Learning Evaluation)

1. I listen to English radio broadcasts (e.g., VOA and BBC) to improve my English proficiency
2. At the beginning of the semester, I set technology-assisted English learning goals.
3. I often monitor my technology-assisted English learning progress.
4. I reflect on the effectiveness of using technologies for English learning.
5. I adjust my English learning plans in response to different technology-assisted learning activities.

## SS (Social Strategies)

1. I seek advice on how to use technologies effectively for English language learning.
2. I seek opportunities to talk with native English speakers through technological tools.
3. When I have problems in English learning, I ask my teacher for help through technological tools.
4. I share my problems with my classmates online so we can solve our problems together.

## TE (Technology-Based English Song and Movie Learning)

1. I "copy" useful words and expressions in English movies or programs.
2. I practice saying new expressions in English movies or programs to myself.
3. I listen to English songs to help me remember words.
4. I use technologies (e.g., English movies) to learn more about English and the culture.
5. I use technologies to connect English learning with my personal interest (e.g., playing English games, or listening and singing English songs).

## TV (Technology-Based Vocabulary Learning)

1. I use lexical apps to help me memorize new words.
2. I use online dictionaries to check English words.
3. I use technologies (e.g., vocabulary apps) to help me persist in my English learning goals.

frontiers in Psychology

ORIGINAL RESEARCH
published: 26 January 2021
doi: 10.3389/fpsyg.2020.594013

# Commercial Video Games in School Teaching: Two Mixed Methods Case Studies on Students' Reflection Processes

*Marco Rüth* and Kai Kaspar*

*Department of Psychology, University of Cologne, Cologne, Germany*

Commercial video games are popular entertainment media and part of students' media reality. While commercial video games' main purpose is not learning, they nonetheless could and should serve as objects of reflection in formal educational settings. Teachers could guide student learning and reflection as well as motivate students with commercial video games, but more evidence from formal educational settings is required. We conducted two mixed methods case studies to investigate students' reflection processes using commercial video games in regular formal high school teaching. In a double lesson, 29 students of a 10th-grade biology course (Study 1) and 17 students of a 12th-grade advanced course on history (Study 2) played and discussed a commercial video game related to the current curricular topic. We examined the reflection processes of students in terms of their reactions to the teachers' game-related statements and questions. Regarding teachers' statements, students discussed several topics related to game enjoyment and the games' representation of topic-related content. Regarding teachers' questions, students discussed multiple goals in each game, how the games represented topic-related content, and how the games could be appropriate for learning. In Study 2, students additionally discussed emotions, stereotypes, violence, and the narrative related to the digital history game. We found that the discussions provided students opportunities to reflect on their game experiences and the current curricular topic as well as to practice media criticism. We further provide quantitative results on students' perceived topic knowledge, on several facets of their learning motivation, and on their acceptance of video games. Overall, our findings illustrate the educational value of using commercial video games as objects of reflection.

**Keywords: game-based learning, reflection, motivation, video game acceptance, commercial video games, guided discovery learning, formal education**

OPEN ACCESS

***Edited by:***
*Mirta Vernice,*
*University of Urbino Carlo Bo, Italy*

***Reviewed by:***
*Marlou Poppelaars,*
*Radboud University, Netherlands*
*Thomas Connolly,*
*University of the West of Scotland, United Kingdom*

****Correspondence:***
*Marco Rüth*
*marco.rueth@uni-koeln.de*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *12 August 2020*
***Accepted:*** *08 December 2020*
***Published:*** *26 January 2021*

***Citation:***
*Rüth M and Kaspar K (2021) Commercial Video Games in School Teaching: Two Mixed Methods Case Studies on Students' Reflection Processes. Front. Psychol. 11:594013. doi: 10.3389/fpsyg.2020.594013*

## INTRODUCTION

More than 80% of school-aged children are video gamers in the United States (13–17 years; Pew Research Center, 2018), and 15- to 19-year-olds play video games more than an hour each day on average (Bureau of Labor Statistics, 2019). Likewise, about 87% of 12- to 19-year-olds in Germany are video gamers and play games more than an hour per day on average

(Pedagogical Media Research Center Southwest, 2019). Thus, for most secondary school students, gaming is a typical leisure activity and part of their everyday life. Beyond their popularity, video games are a cultural good, and the high potential value of using commercial video games in formal education has been articulated for years (e.g., Sandford et al., 2006; van Eck, 2009; Becker, 2017a,b,c; Caldwell et al., 2017; Arnseth et al., 2018). Using commercial video games in school teaching allows teachers to develop students' media literacy, a key topic of formal education (Squire, 2008) and a cross-sectional and cross-curricular goal relevant to all teachers. In this regard, "the meanings and functions of games cut across formal and informal contexts, but [. . .] this can also be a source of discussion and reflection" (Arnseth et al., 2018, p. 124). Video games include serious games, which are primarily designed and effective for learning and retention of knowledge (Wouters et al., 2013), and commercial off-the-shelf games, which are primarily designed for entertainment purposes and the gratification of players' intrinsic needs (Ryan et al., 2006). Importantly, there might be a discrepancy between teachers' assumption that serious games engage students and students' expectations that playing serious games is like playing commercial games they usually play outside school (Arnseth et al., 2018) so that teachers might provide students the so-called "chocolate-covered broccoli". In contrast, using commercial video games could allow teachers to authentically address students' media reality (students' expectations) and allow them to teach subject-specific topics and the cross-sectional topic of media literacy (teachers' goals). Accordingly, we argue that the integration of commercial video games into formal school teaching could partly resolve the aforementioned discrepancy, while it is still unclear what forms these integrations might take.

Teachers reported several observations about how using video games in school teaching positively influenced students' engagement, learning motivation, content knowledge, and subject-specific as well as cross-curricular skills (Huizenga et al., 2017). Notably, positive effects of digital game-based learning are not only based on teachers' individual observations and assumptions but corroborated by several meta-analytic results (Boyle et al., 2012, 2016; Clark et al., 2016; Hainey et al., 2016). Regarding students' expectations, the results of a survey comprising 858 secondary school students indicate that they might accept video games in education (Bourgonjon et al., 2010). This is important because, when playing video games as leisure activity, players seem to engage only in lower levels of reflection (Mekler et al., 2018). In formal school teaching, however, students could be triggered to engage in deeper reflection processes after playing commercial video games to address their media criticism skills and media literacy in general.

Commercial video games relate to learning theories (van Eck, 2009; Becker, 2017b), and several features can make them suitable for learning (Gee, 2005; Wilson et al., 2009). For example, players can actively explore and manipulate content, solve problems and challenges, and receive scores or other means of feedback. Still, teachers who decide to use commercial video games should be aware that students could be confronted with inaccurate or wrong information, for example, regarding scientific principles in biology (Schrader et al., 2016) or regarding stories with narrative biases in history (McCall, 2016). Accordingly, commercial video games should not be viewed as mere content providers for formal education. Instead, using commercial video games can trigger student reflection, for instance, via conflicts or competitions in games (Kiili et al., 2011).

While integrating commercial video games in education seems promising, a common key issue is *how* to best align gaming with existing formal educational settings (Mayer, 2019). In line with previous works that highlighted benefits of using commercial video games in education (van Eck, 2009; Becker, 2017c; Caldwell et al., 2017), effective integrations into formal educational settings require appropriate instructional approaches. Arnseth et al. (2018) explicated that teachers need to enable students to reflect on what they do and experience in video games. Similarly, Huizenga et al. (2017) argued that teachers should allow students to interact in the classroom by means of questioning and discussing their game experiences. Indeed, teacher-provided forms of guidance had the largest effect on students' learning outcomes in digital game-based contexts (Clark et al., 2016). Moreover, meta-analytic findings outline that guided discovery approaches can promote learning and direct reflection processes more than unassisted discovery (Alfieri et al., 2011).

Reflection processes have strong relations to learning and include goal setting, action, and feedback (Seale and Cann, 2000). Reflection processes are also related to playing video games, while playing can be understood as an ongoing cyclic process of acting within the game, getting a reaction, evaluating the reaction, and reflecting on it (Stephenson-Mittlböck, 2012). Reflection processes can take place during gaming (reflection-in-action) and after gaming (reflection-on-action) (Schön, 1987; Brockbank and McGill, 2007b). After gaming, players can reflect on their game experiences by bringing them to mind, thinking about them, and evaluating them against their initial playing strategy, knowledge, or hypotheses (Kiili, 2007). Thereby, students can learn via generalization of experiences or gain a new understanding following a paradigm shift (Brockbank and McGill, 2007c). When playing video games as leisure activity, however, players do hardly reach the corresponding level of reflection to gain new insights, yet they do enjoy reflecting on their game experiences (Mekler et al., 2018). Still, laboratory studies indicate that some game elements can trigger reflections on real-life topics such as mortality (Chittaro and Sioni, 2018). Further, scores in a transfer test were higher if student reflection was guided during digital game-based learning in terms of self-explanations (Johnson and Mayer, 2010), and explanatory feedback, moreover, reduced students' errors and misconceptions (Moreno and Mayer, 2005). The latter studies provide mostly quantitative (Moreno and Mayer, 2005; Johnson and Mayer, 2010) or qualitative results (Chittaro and Sioni, 2018) based on controlled experimental conditions with single play sessions. Taken together, it thus seems necessary to systematically investigate students' reflection processes following a single play session in the ecologically valid setting of formal school teaching.

In the following, we present two case studies using commercial video games based on a guided discovery approach to teach media criticism and to foster students' reflection processes.

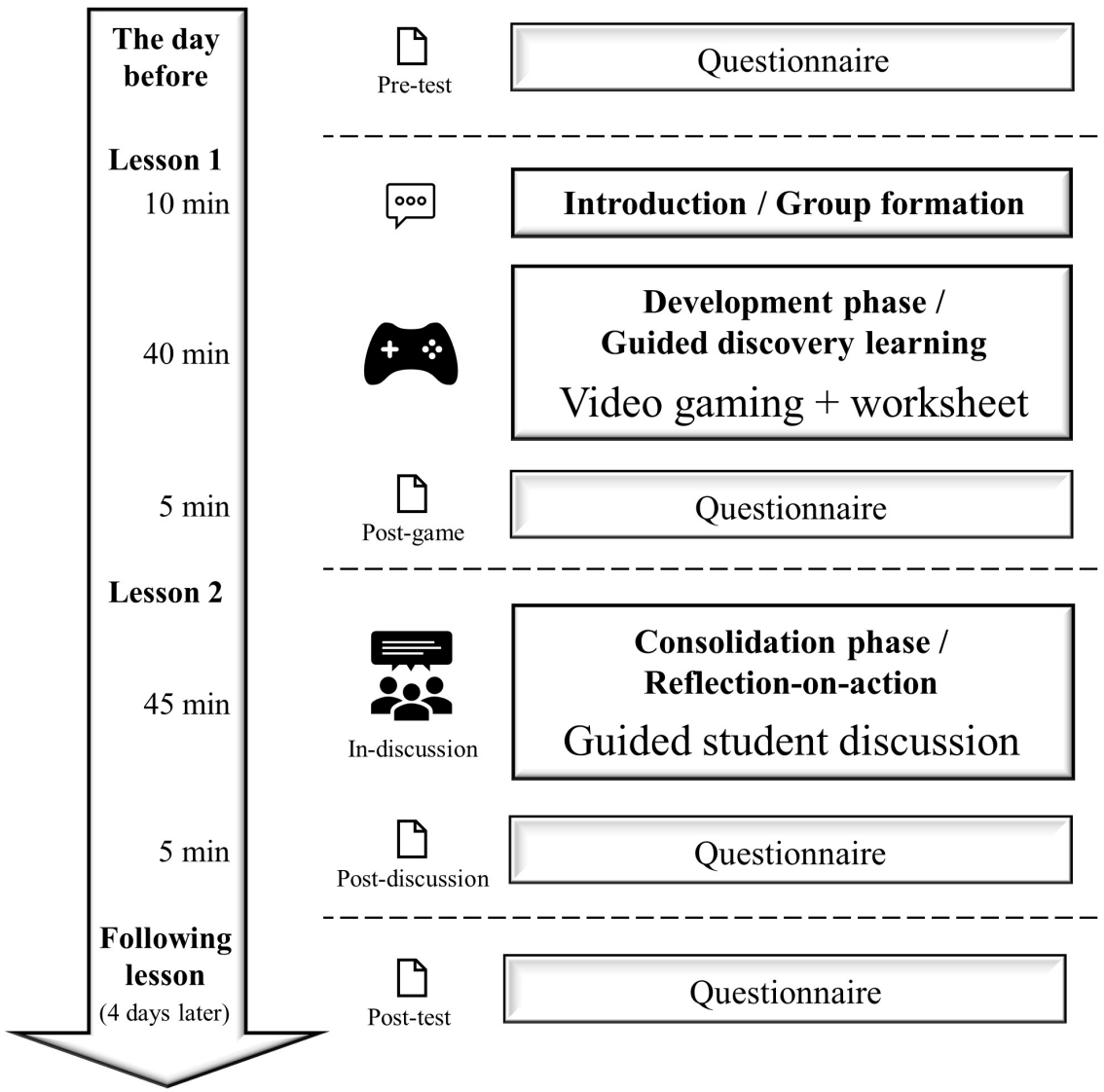


**FIGURE 1 |** Overview of the study procedure.

To understand how commercial video games serve as objects of reflection in formal educational settings, we integrated two different commercial video games into two similar formal high school settings. Both settings were regular double lessons, and students had already addressed the topics in previous lessons. In line with curricular guidelines, students' learning objectives were to analyze topic-related content and to reflect on the topic. To investigate students' reflection processes, we used measurements the day before the lessons (pre-test), after video gaming (post-game), during a discussion following video gaming (in-discussion), after the discussion (post-discussion), and at the beginning of the following lesson (post-test) (see **Figure 1**).

In formal educational settings, student discussions allow teachers to debrief students after gaming (reflection-on-action) so that students can share and complement their game experiences and lines of thoughts (Peters and Vissers, 2004). More specifically, teacher-led student discussions can provide students a joint phase of reflection-on-action as well as guidance in terms of reflective support, that is "to assist students in reflecting on the learning process and the knowledge acquired" (De Jong and Lazonder, 2014, p. 375). Worksheets can further structure learning activities and facilitate reflection processes with video games in formal high school teaching (Panoutsopoulos and Sampson, 2012). After gaming, teachers can guide reflection processes by making statements and asking questions related to students' game experiences, while the overall question is what students discuss:

RQ1: What will students discuss regarding their game experiences in reaction to teachers' game-related statements and questions?

While worksheets *can* guide student learning, students *have to* pursue games' goals to make progress in games. Since goals also play a role in the process of reflection and gaming in terms of goal setting and playing strategy (Seale and Cann, 2000; Kiili, 2007), we question which goals students pursue (RQ1a). Moreover, the content in commercial video games is not designed as learning content *per se* (van Eck, 2009; Schrader et al., 2016) and games can transport biased narratives (McCall, 2016) so that it is important to understand how students reflect on the content of games (RQ1b). How students evaluate games' appropriateness for school teaching (RQ1c) is also of interest, since students' perceived usefulness of video games can have a strong positive effect on students' preference for video games (Bourgonjon et al., 2010). Taken together, we address the following research questions:

RQ1a: Which goals of the game will students discuss?
RQ1b: Which content of the game will students discuss?
RQ1c: Which aspects of games' appropriateness for school teaching will students discuss?

Importantly, effective instruction should include student learning and student motivation (Slavin, 1994). Regarding student learning, both video gaming and student discussions might affect students' knowledge by means of social interactions between students (Pea, 1993) or by means of students exchanging knowledge and making sense out of individual game experiences together (Palincsar, 1998). We thus question if a gaming phase (RQ2a) and a discussion phase (RQ2b) foster learning via reflection processes (e.g., generalization of experiences or new understandings) in terms of higher students' perceived topic knowledge. We further differentiate between students' perceived influence of video gaming and of the discussion (RQ2c):

RQ2a: How will a video gaming phase influence students' perceived topic knowledge?
RQ2b: How will a guided discussion phase following video gaming influence students' perceived topic knowledge?
RQ2c: Which influence will students attribute to video gaming, compared to a subsequent guided student discussion, on their perceived topic knowledge?

Regarding student motivation, students were found to be more motivated using video games than in non-game conditions (Clark et al., 2016), which suggests that video games can be promising incentives in formal educational settings. While the main purpose of commercial video games is to fulfill intrinsic personal needs and to elicit enjoyment and motivation (Ryan et al., 2006; Boyle et al., 2012), we consider the issue of providing students "chocolate-covered broccoli" and question if the games still serve their entertainment purpose in formal educational settings. Following Rüth and Kaspar (2020), we therefore address main facets of students' learning motivation in terms of students' interest, relevance, competence, satisfaction, and volition. In specific, we here examine students' learning motivation regarding the lesson's topic, video gaming, and discussing: first, since video gaming and discussing a commercial video game could trigger learning and reflection processes and provide new insights or perspectives to students, we explore if students' learning motivation regarding the lesson's focal topic will change (RQ3a). Second, we assess students' learning motivation regarding video gaming (RQ3b). Third, we examine students' learning motivation regarding discussing (RQ3c):

RQ3a: Will students' learning motivation regarding the lesson's focal topic (in terms of topic interest, topic commitment, personal relevance of the topic, social relevance of the topic, and volition to learn about the topic) change from pre- to post-test?
RQ3b: Which level of learning motivation regarding video gaming (in terms of game enjoyment, game interest, perceived competence, and satisfaction) will students report at post-game?
RQ3c: Which level of learning motivation regarding discussing (in terms of personal interest, personal relevance, and motivation to participate) will students report at post-discussion?

Finally, decisions on integrating commercial video games into formal educational settings should consider whether students accept video games in such educational settings. Secondary school students' (age 12–20) preference for using video games in school teaching partly relies on the usefulness and learning opportunities they see in using video games and if they know how to use video games (Bourgonjon et al., 2010). Yet, the survey study of Bourgonjon et al. (2010) suggests to also consider the actual use of video games in formal educational settings, which seems to be neglected in several studies on video game acceptance (Wang and Goh, 2017). Therefore, we here investigate students' acceptance of video games when these are used in formal school teaching:

RQ4a: Will students' acceptance of video games in school teaching change from pre- to post-test?
RQ4b: Which general level of acceptance of video games will students report at pre- and post-test?

## STUDY 1

It has been argued that particularly evolution is a topic that provides many possibilities for scientific misconceptions and misinterpretations (Herrero et al., 2014) so that it is important to foster student reflection on commercial video games addressing evolution. In Study 1, we integrated the commercial video game *Spore* (Maxis, 2008) into a double lesson of biology teaching on evolution. We decided to integrate this video game into formal school teaching based on teacher's expertise, game reviews, a pedagogical review, and previous game-related conceptual and empirical research: Critics were mostly positive, while gamers had many user discussions, specifically in the year of release (second most discussed PC Game of 2008, Metacritic, 2020a). A pedagogical review outlined Spore as a casual game that allows creating one's own world, with hardly any consequences of one's mistakes (Spielbar, 2020a). Previous research criticized Spore conceptually for how it presents evolution and that playing it

could result in misconceptions about evolution (Bohannon, 2008; Bean et al., 2010; Ching, 2012; Schrader and McCreery, 2012; Schrader et al., 2016). Schrader and McCreery (2012) concluded that "for many educators, SPORE represents a failure to build serious games", but it "provides an interesting case to examine learning with a game" (pp. 18–19). In other words, that Spore represents the scientific topic of evolution does not make it a serious game, and unlike for serious games, learning the content that Spore conveys is even not recommended. Still, learning with Spore might be promising when teaching material or student discussions complement the game experiences: A case study found a higher increase in students' examination scores and more engagement with course material when students played Spore along with other teaching materials compared to students who received traditional instruction in a regular upper-level college class (Poli et al., 2012). Herrero et al. (2014) highlighted several main learning principles included in Spore and provided qualitative evidence on student learning in an extracurricular workshop in which students discussed. Overall, integrating Spore into formal school teaching constitutes an interesting case to investigate how students discuss the topic evolution and reflect on their game experiences.

## Materials and Methods

### Participants

The sample comprised 29 students from a 10th-grade biology course at a German comprehensive school (high school level) ($M_{age}$ = 15.52, $SD_{age}$ = 0.63; 16 female). In the analyses, differences in sample size are due to some students being absent at some times of measurement (for an overview, see **Supplementary Table 1**). Based on a 5-point scale including verbal markers (1 = "never", 5 = "very frequently") and ratings from 26 students present at pre-test, students reported to play video games occasionally in their leisure time ($M$ = 2.65, $SD$ = 1.38, 26.93% never) but seldom in school teaching ($M$ = 1.69, $SD$ = 0.68, 42.31% never), about evolution ($M$ = 1.27, $SD$ = 0.72, 84.62% never), and the game Spore ($M$ = 1.12, $SD$ = 0.59, 96% never, 4% frequently).

### Design and Procedure

We used a one-group pre–post mixed methods design and collected data at five time points: the day before the lesson (pre-test), during the lesson (post-game, in-discussion, post-discussion), and at the beginning of the following lesson that was 4 days later (post-test). We depict the procedure in **Figure 1** as well as the complete measurement plan including all dependent variables, time points, and references to individual research questions in **Supplementary Table 1**. In a semistructured interview before the study, the biology teacher stated personal interest in video games and some experience with using video games in school teaching. Similar to Rüth and Kaspar (2020), the collaboration with the teacher was initiated via a broad announcement on using video games in formal school teaching in a regional network of teachers.

We employed a strong program of triangulation (Flick et al., 2012) by combining quantitative and qualitative methods (methodological triangulation) and data (data triangulation). On the one hand, we audiorecorded and transcribed the discussion phase (RQ1a–c) and digitized students' responses to open questionnaire items (RQ2a–c). On the other hand, in the questionnaires, we also used symmetric direct self-rating scales with five levels to measure students' perceived topic knowledge (RQ2a and RQ2b), learning motivation (RQ3a–c), and acceptance of video games (RQ4a and RQ4b). We used direct ratings to reduce acquiescence response bias (Saris et al., 2010) and to provide students intuitive self-ratings. Each scale was accompanied by verbal markers (e.g., "not at all", "rather not", "moderately", "rather", and "very"). All questionnaires were in paper–pencil format. Similar measures were used to evaluate the integration of another commercial video game into formal school teaching (Rüth and Kaspar, 2020).

The study took place in a regular double school lesson lasting 120 min. In the first lesson, the teacher shortly outlined the lesson plan, assigned students to groups of four, and provided students with spoken and written (worksheet) task instructions. All groups played the video game for about 40 min, while the teacher ensured that each student played for a similar duration (10 min) and completed a worksheet (development phase). In the second lesson (consolidation phase), the teacher introduced the discussion phase by asking students to rate three different statements about the game and then moderated a semistructured discussion (guided student discussion).

During the double lesson, two carefully instructed observers filled out a structured observation protocol for event sampling. Following Rüth and Kaspar (2020), the purpose of this protocol was to document organizational (e.g., time delays or technical issues) or behavioral aspects (e.g., inattentive or demotivated students), which might be of relevance to assess the validity of our data. According to the observation protocols, some students were late or loud at the beginning of the first lesson but appeared very motivated and concentrated during video gaming, with remaining conversations being only about game experiences. Further, when the double lesson was over, students would have liked to continue playing the game.

### Teaching Materials

Teaching materials consisted of a lesson plan, worksheets, and teaching guidelines for the gaming phase and the discussion phase. The worksheets and the discussion were meant to foster student reflection by asking them to think about the game's goals and to which degree the game addresses or simplifies topic-related aspects, takes a specific viewpoint on the topic, or presents evolutionary processes wrongly. We developed the teaching materials in close collaboration with the teacher.

We used the commercial video game Spore (Maxis, 2008; age rating of 12) and left its original game narrative unchanged: Students first watched the cinematic intro and played the build-in tutorial as well as the first two out of five game phases. In phase 1 (emergence of living cells), players control a simple creature, experience that the environment of the game can be beneficial or harmful for their creature, and choose one out of three diets (carnivore, herbivore, or omnivore) for their creature. In phase 2 (evolution to terrestrial creatures), players face more complex appearances and behaviors than in phase 1,

**TABLE 1 |** Content of the semistructured discussion phase in Study 1.

| Category and subcategories | *n* | Teacher's questions and students' example statements |
|---|---|---|
| **Game's goals** | | What is the general goal of the game? What has allowed or hindered you to reach it? |
| Development of the creature | 2 | "The goal of the game was, as a carnivore, to eat a lot of meat and to keep growing, to evolve, or, as an herbivore, to eat many plants, again to evolve and to become bigger" (S17) |
| Becoming the dominant species | 1 | "Trying to become more and more dominant and to ally with species within the own environment, or to displace them and to try becoming the most dominant species on the planet" (S5) |
| **Game's content** | | Are the evolutionary factors mutation, recombination, and selection represented in the game? If so, how? If not, why not? At which points in the game were, in favor of the gameplay, evolutionary processes represented differently or in a simplified way? Were there game scenes, in which evolutionary processes were perhaps even represented incorrectly? |
| Content-based criticism | 3 | "For example, when one bought these body parts, that was deliberately changed to have fun" (S10)<br>"Mutation did not take place randomly [. . .] That was the factor, which was changed due to game enjoyment and technical reasons" (S5) |
| Content identification | 3 | "Selection I would say [. . .] it happened that one of us died twice in a row. The following time, we said 'okay, that was not that beneficial, we have to do it differently"' (S5)<br>"At the beginning, without any defense or body parts for defense, one died much faster. Thus, if one had not put those parts on, one would not have survived" (S22)<br>"At the start, mutation took place and the creature evolved and has thereby adapted. And that leads from mutation to recombination" (S8) |
| **Games' appropriateness for school teaching** | | Which changes would you make to the game if you were an evolutionary biologist? To what extent do you think video games are suitable for teaching biological knowledge? |
| Constructive critique | 8 | "I would design it much livelier, much more realistic" (S7)<br>"I would completely remove the creature editor and keep everything random" (S27)<br>"I would make the evolutionary factors more clearly" (S8) |
| Games' potentials | 2 | "I think that learning with video games is very good in general, because it is not that boring as writing in school. I think one can learn much better with games, because one is keener on learning, and games really show that. And if there is only someone in front and explains it, it is probably more difficult for that person than for a game, which can really visualize it" (S10)<br>"I would agree with S10 that it is rather good that one can visualize such things more easily with games. [. . .] I would rate this a little higher with respect to learning, if it is well done, than, for instance, a documentary" (S5) |
| Games' limitations | 2 | "One aspect one should care about during game development is that the content is disseminated correctly and that there are no errors included, because of which content is disseminated incorrectly" (S5)<br>"What would be missing for me are all those technical terms. That is all well and good, but what is the use of it, if I have no idea what it is called?" (S7) |

*Categories and example statements are based on responses of 11 students, while in total, 26 students participated in the discussion.*

since creatures gain lower extremities and can join a group of creatures, respectively. In both phases, players collect experience points that allow them to modify their creature. Each group of students played the game using a laptop computer and an external computer mouse. The repetitive background sound in the game was reproduced by a pair of speakers in the front of the classroom.

## Measures

### *Guided student discussion (RQ1)*

At the start of the discussion phase, the teacher triggered students' reflection-on-action by asking them to react to three statements: "Playing the game Spore was fun", "The game vividly illustrates evolutionary processes", and "The game Spore should be used in biology lessons". Following each statement, students positioned themselves on an imaginary scale in the classroom ranging from 0 to 10 (RQ1). During the subsequent semistructured discussion, the teacher asked one question about the game's goal (RQ1a), three questions addressing the game's content (RQ1b), and two questions addressing games' perceived appropriateness for school teaching (RQ1c). Question wordings are depicted in **Table 1**.

### *Students' perceived knowledge about the topic evolution (RQ2)*

We assessed how students perceived the influence of gaming (RQ2a) and discussing (RQ2b) on their knowledge about evolution via self-ratings. At pre-test (the day before the lesson), post-game, and post-test (following lesson, four days later), we asked "How highly do you rate your current knowledge of the topic evolution?" (self-rated topic knowledge). In addition, regarding the specific influence of gaming on topic knowledge (RQ2a), we asked "How much did playing the game on the whole support you in increasing your knowledge of the topic evolution?", and students also provided open responses ("Please describe in your own words how playing the game influenced your knowledge of the topic evolution.") at post-game and post-test. In order to assess the specific content that students learned by video gaming, they were asked to recall up to three aspects at post-test ("What did you learn by playing the game? Please write down up to three aspects."). Regarding the perceived influence of the guided discussion on topic knowledge (RQ2b), students provided self-ratings ("How much did you learn about the topic evolution during the discussion?") and open responses ("Please describe in your

own words how the discussion influenced your knowledge of the topic evolution.") at post-discussion. Finally, we compared the perceived influence of gaming versus discussing on topic knowledge (RQ2c) by means of open responses at post-discussion ("Please describe in your own words how playing the game in comparison with the discussion influenced your knowledge of the topic evolution.").

#### *Students' learning motivation regarding lesson's topic, gaming, and discussing (RQ3)*

Regarding the lesson's topic (RQ3a), we assessed students' learning motivation in terms of interest in the topic, topic commitment, personal relevance of the topic, social relevance of the topic, and volition to learn about the topic. For item wordings, see **Table 2**. Learning motivation regarding video gaming was operationalized in terms of game enjoyment, game interest, perceived competence, and satisfaction (RQ3b). Item wordings are depicted in **Table 3**. Learning motivation regarding discussing was operationalized in terms of personal interest, personal relevance, and motivation to participate in the discussion (RQ3c). Items are depicted in **Table 4**.

#### *Students' acceptance of video games (RQ4)*

At pre- and post-test, students rated their acceptance of video games in school teaching (RQ4a). We also explored students' acceptance of video games as a leisure activity and as significant part of life (RQ4b). Items are depicted in **Table 5**.

### Data Analysis

For each of the following quantitative and qualitative analyses, we included all available data from each time point, while data were missing completely at random (MCAR) in both studies according to Little's MCAR test ($ps > 0.999$).

With reference to other case studies on game-based learning in school teaching (Berg Marklund and Alklind Taylor, 2016), we conducted a thematic analysis of students' reactions to the statements of the teacher (RQ1). We followed a theoretical/analyst-driven approach related to our research question and considered guidelines for thematic analyses (Braun and Clarke, 2006) to formulate codes and semantic themes, as illustrated by means of text excerpts.

Students' responses during the semistructured phase of the guided student discussion (RQ1a–c) and students' open responses in the questionnaires (RQ2a–c) were summarized and categorized following content-structured qualitative content analysis (Kuckartz, 2016). Three main categories were derived from the research questions (RQ1a = "game's goals", RQ1b = "game's content", and RQ1c = "games' appropriateness for school teaching"), and subcategories were created based on content analysis and categorized by two independent coders. Interrater agreement was assessed in terms of Cohen's kappa (Cohen, 1960) and was perfect in Study 1 (Cohen's $\kappa = 1.00$) and good in Study 2 (Cohen's $\kappa = 0.75$). Cases of disagreement were resolved through discussion.

To test potential changes in ratings over time, we used *t*-tests for dependent samples (RQ2a, RQ2b, RQ3a, and RQ4a). To test if the means of the ratings were different from the scales' midpoints (RQ3b and RQ3c), we used one sample *t*-tests. We report effect sizes in terms of Cohen's *d*. In line with Cohen (1988), common thresholds for small, medium, and large effect sizes are 0.2, 0.5, and 0.8, respectively. The alpha level for all statistical tests was 0.05.

## Results

### Content of the Discussion Phase (RQ1)

The discussion phase started with the three statements of the teacher. Following the teacher's first statement ("Playing the game Spore was fun"), the teacher asked those students with high and low ratings to elaborate on their ratings. The main theme was "game enjoyment". One student (S5) elaborated on a high rating of 10, then a fellow student (S6) and the teacher (T) joined in:

> S5: We had enough fun while, while, uh, designing our creature, simply because there were some comic factors where we then just almost fell off the chair with laughter, but. . .
> S6: We have something funny to do right now, so. . .
> T: I am curious.
> S6: Yes, we will see (Laughter of the surrounding students).

Students talked about to continue playing, and based on the observation protocol, indeed most students would have liked to continue playing. However, some students had only limited fun gaming (ratings of 1 or 2):

> S8: Because when one plays it with friends now, like in a group of four, it was a bit of fun, because we made fun of it a bit and so on. But if one imagines that one also plays it alone and in such a way—that is no fun.

Overall, students either did or did not enjoy gaming, as illustrated by the teacher's statement: "Well, a [rating of] ten. Now we have a big emptiness, now I'm going to go there (...) [to the students with the rating] 'I don't agree at all'". Students also discussed "game features", in particular that the game strongly motivated to continue playing.

Following the teacher's second statement ("The game vividly illustrates evolutionary processes"), the students were undecided. The main theme was the "implementation of evolutionary processes":

> S5: One just, I don't know, goes there, looks for someone else from one's species to mate with, then one modulates one's creature as one likes. [. . .] This means that the factor 'mutation' is missing.
> S10: And one buys one's characteristics. That's not like in nature. That's just, well, one chooses how one looks afterward, so to speak, and it's just not like that somehow. Yes.
> S11: (interjects) But how else could they have done that?
> S10: (simultaneously with S12) By chance.
> S5: That one randomly gets some things and later one can only choose between those things.

In the last excerpt, students first discussed evolutionary processes to be missing or unrealistic in the game but

then suggested potential game changes and shifted from problems to solutions. Similarly, the theme "game character development" emerged, and students discussed both issues and solutions.

Following the teacher's third and final statement ("The game Spore should be used in biology lessons"), students discussed the theme "game's potential role in biology teaching":

**TABLE 2 |** Influence of lessons on students' learning motivation regarding lessons' topic in both studies.

| | Study 1 | | | | | Study 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Measure** | **Pre-test *M* (*SD*)** | **Post-test *M* (*SD*)** | ***t*** | ***p*** | ***d*** | **Pre-test *M* (*SD*)** | **Post-test *M* (*SD*)** | ***t*** | ***p*** | ***d*** |
| Topic interest: "How interested are you in [topic]?" | 3.67 (1.02) | 3.33 (0.80) | −1.67 | 0.110 | 0.37 | 4.25 (0.86) | 4.38 (0.72) | 0.81 | 0.432 | 0.20 |
| Topic commitment: "How much are you concerned with [topic] in your leisure time?" | 1.57 (0.81) | 1.95 (0.86) | 1.79 | 0.088 | 0.39 | 2.50 (0.73) | 2.75 (0.86) | 1.46 | 0.164 | 0.37 |
| Personal relevance of the topic: "How important is [topic] for you?" | 2.95 (0.86) | 3.29 (0.72) | 1.92 | 0.069 | 0.42 | 3.75 (0.77) | 4.00 (0.73) | 1.46 | 0.164 | 0.37 |
| Social relevance of the topic: "How important do you think [topic] is for other people?" | 3.52 (0.68) | 3.24 (0.62) | −1.37 | 0.186 | 0.30 | 3.31 (1.01) | 3.31 (0.95) | 0.00 | >0.999 | 0.00 |
| Volition to learn about the topic (three items; Study 1: $\alpha = 0.78$, Study 2: $\alpha = 0.66$): e.g. "How important to you is learning about [topic]?" | 3.37 (0.77) | 3.33 (0.57) | −0.18 | 0.863 | 0.04 | 4.13 (0.53) | 4.02 (0.65) | −0.68 | 0.510 | 0.17 |

*Values depict mean (standard deviation) of 21 biology students (Study 1) and 16 history students (Study 2). [topic] is a placeholder for "evolution" (Study 1) and "First World War" (Study 2). Statistical results refer to dependent t-tests comparing the means at pre- and post-test.*

**TABLE 3 |** Students' learning motivation regarding video gaming in both studies.

| | Study 1 | Study 2 |
|---|---|---|
| **Measure** | **Post-game *M* (*SD*)** | **Post-game *M* (*SD*)** |
| Game enjoyment (four items adapted from Reinecke et al., 2012; Study 1: $\alpha = 0.91$, Study 2: $\alpha = 0.94$): e.g. "How much did you enjoy playing the game?" | 3.65** (0.92) | 3.81** (1.06) |
| Game interest: "How interested are you in the game?" | 3.81** (1.33) | 4.00** (1.17) |
| Perceived competence (five items adapted from the Player Experience of Need Satisfaction questionnaire by Ryan et al., 2006; Study 1: $\alpha = 0.71$, Study 2: $\alpha = 0.94$): e.g. "How successful did you feel while playing the game?" | 3.94*** (0.57) | 3.75** (0.99) |
| Satisfaction: Game preference | | |
| Game recommendation: "How much would you recommend the game to your friends?" | 3.04 (1.11) | 3.59* (1.06) |
| Game preference for school teaching: "How much would you like to play the game again in school teaching?" | 4.35*** (1.06) | 4.12** (1.36) |
| Game preference in leisure time: "How much would you like to play the game again in your leisure time?" | 3.12 (1.37) | 3.35 (1.41) |
| Satisfaction: Game evaluation | | |
| Game graphics: "How much do you like the graphics in the game?" | 2.77 (0.95) | 3.88** (1.05) |
| Game music: "How much do you like the music in the game?" | 3.00[a] (0.76) | 4.18*** (1.07) |
| Overall game rating: "How much do you like the game overall?" | 3.96*** (0.87) | 3.94**[b] (1.24) |
| Game's appropriateness for school teaching: "How much do you think the game is appropriate for use in school teaching?" | 3.65** (1.09) | 3.71* (1.21) |

*Values depict mean (standard deviation) of 26 biology students (Study 1) and 17 history students (Study 2). Asterisks indicate the results of one sample t-tests comparing the mean with the scale's midpoint of 3. [a]n = 15; [b]n = 16. *p < 0.05, **p < 0.01, ***p < 0.001.*

**TABLE 4 |** Students' learning motivation regarding discussing in both studies.

| | Study 1 | Study 2 |
|---|---|---|
| **Measure** | **Post-discussion *M* (*SD*)** | **Post-discussion *M* (*SD*)** |
| Personal interest in the discussion: "How interesting was the discussion about [topic] for you?" | 3.00 (1.06) | 3.88*** (0.78) |
| Personal relevance of the discussion: "How relevant was the discussion about [topic] for you?" | 3.42* (0.90) | 3.94*** (0.75) |
| Motivation to participate in the discussion: "How much were you motivated to participate in the discussion about [topic]?" | 2.96[a] (1.17) | 3.41 (1.06) |

*Values depict mean (standard deviation) of 26 biology students (Study 1) and 17 history students (Study 2). [topic] is a placeholder for "evolution" (Study 1) and "First World War" (Study 2). Asterisks indicate the results of one sample t-tests comparing the mean with the scale's midpoint of 3. [a]n = 25. *p < 0.05, ***p < 0.001.*

**TABLE 5 |** Students' acceptance of video games in both studies.

| | Study 1 | | Study 2 | |
|---|---|---|---|---|
| **Measure** | **Pre-test *M* (*SD*)** | **Post-test *M* (*SD*)** | **Pre-test *M* (*SD*)** | **Post-test *M* (*SD*)** |
| Acceptance of video games | | | | |
| . . .in school teaching: "How much are you in favor of using video games in school teaching?" | 2.90 (1.30) | 2.95 (1.12) | 3.63 (1.31) | 3.50 (1.37) |
| . . .as a leisure activity: "How important are video games for you as a leisure activity?" | 2.48 (1.50) | 2.43 (1.33) | 3.44 (1.82) | 3.44 (1.79) |
| . . .as significant part of life: "How significant are video games in your life?" | 2.10 (1.37) | 2.24 (1.14) | 2.94 (1.53) | 3.13 (1.50) |

*Values depict mean (standard deviation) of 21 biology students (Study 1) and 16 history students (Study 2).*

S18: I think it's just fun and I think if one has fun together, then one learns better. So, if one has to make some exercise sheets all the time, I think one doesn't learn so much, because it's not so much fun. [. . .]

S14: I would use it especially for younger children, because then they have a little bit of fun and learn playfully, so maybe we are already a little too old for that. [. . .]

S5: I think that if we do it like now, for two hours in one piece, that's okay for one lesson, I also know how it [the game] goes on. [. . .] The first two [phases] are biological, the following ones could be directly used in politics or business [education]. These are the first two phases, which are biologically based, thereafter it's not so good for the topic.

S22: Well, I think the game was fun, but one didn't really learn something from it. Because what we did in class before, one learned the technical terms and how it happens, but there, that was just a game. One has not really learned anything. Yes, but as I said before, if one gets through the topic, playing something like that again at the end is like watching a film, I think it's really good.

S8: Yes, I agree with S22 [. . .] It would only bring me something if I should somehow find out where the evolution and the mutation and the recombination takes place and why it is like that.

To sum up, students once more discussed the relevance of game enjoyment for learning, but they also reflected on the game's target group. A student who knew the game (S5) then argued that the gaming duration sufficed and that the game's later phases could be appropriate for teaching other subjects. Although students thought that no topic-related learning occurred, some favored gaming at the end of a teaching unit or given topic-related search tasks.

After the statements of the teacher, the semistructured discussion phase followed. Many responses of the students to the teacher's questions in this phase were assigned to the following subcategories: Students identified two goals in the game (RQ1a) and identified and criticized topic-related content in the game (RQ1b). When being asked to take an expert's perspective (RQ1c), students mostly provided constructive critique and moreover discussed games' potentials and limitations. **Table 1** depicts the categories, number of statements assigned to the categories, the teacher's questions, and example statements of the students.

## Influence of Gaming and Discussing on Students' Perceived Topic Knowledge (RQ2)

Regarding the influence of the video gaming phase on students' perceived topic knowledge (RQ2a), no change was found between self-rated topic knowledge at pre-test and post-game, $t(18) = -0.44$, $p = 0.667$, $d = 0.10$ (see **Figure 2A**). Students' perceived impact of video gaming on topic knowledge did also not change from post-game to post-test, $t(22) = -1.50$, $p = 0.148$, $d = 0.31$ (see **Figure 2B**). The minority of students had the impression that playing the game increased their topic knowledge (post-game: 33.33%; post-test: 20.83%) (see **Figure 2C**). Further, at post-test, few students recalled one (4.17%) or three topic-related aspects (12.50%) they learned via video gaming, and about every second student did not recall anything (45.83%). The other students' statements were about that the game was fun (8.33%), could improve skills of collaboration, communication, and tactics in general (4.17%), or disregarded topic-related content (4.17%). Of the students, 20.83% gave no open response.

Regarding the influence of the guided discussion phase on students' perceived topic knowledge (RQ2b), we found no change in self-rated topic knowledge from post-game to post-test, $t(19) = -0.37$, $p = 0.716$, $d = 0.08$ (see **Figure 2A**). Students' perceived impact of the discussion on topic knowledge was numerically slightly below the scale's midpoint ($M = 2.80$, $SD = 0.82$), and only few students openly stated that discussing increased their topic knowledge (7.69%) (see **Figure 2C**).

With respect to the question whether video gaming or discussing had a stronger influence on students' topic knowledge (RQ2c), some students indicated to have learned more through video gaming (19.23%) or discussing (11.54%), respectively. However, most responding students stated that there was no influence at all (42.31%), while some students stated a similar influence (3.85%) or gave no response (23.08%).

## Students' Learning Motivation Regarding Lesson's Topic, Gaming, and Discussing (RQ3)

Measures of students' learning motivation regarding the topic of the lesson did not change from pre- to post-test, all $|ts| \leq 1.92$, $ps \geq 0.069$, $ds \leq 0.42$ (RQ3a). **Table 2** shows the results in detail. Students' game enjoyment, game interest, perceived competence, game preference for school teaching, and overall game rating as well as the perceived game's appropriateness for school teaching were above the scales' midpoints, all $ps \leq 0.005$ (RQ3b). **Table 3** shows the results in detail. The perceived personal relevance of the discussion was also above

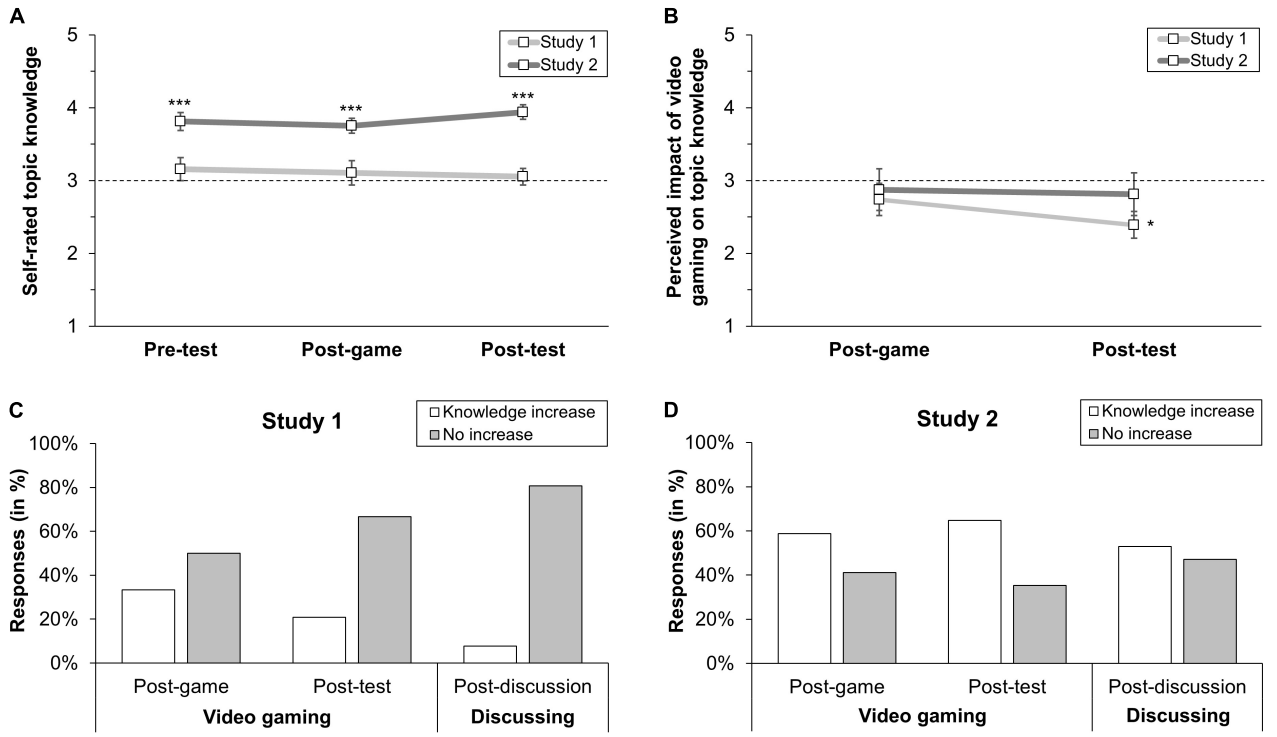


**FIGURE 2 |** Effects of video gaming and discussing on students' perceived topic knowledge in both studies. **(A)** Mean self-rated topic knowledge in Study 1 and Study 2. **(B)** Mean perceived impact of video gaming on topic knowledge in Study 1 and Study 2. **(C)** Open responses on perceived impact of video gaming and discussing on topic knowledge in Study 1 (missing responses not displayed). **(D)** Open responses on perceived impact of video gaming and discussing on topic knowledge in Study 2. Vertical lines indicate the standard error of the mean. Asterisks in **(A,B)** indicate the results of one sample *t*-tests comparing the means with the scales' midpoints of 3 (dashed horizontal lines). * $p < 0.05$. *** $p < 0.001$.

the scale's midpoint ($p = 0.025$) (RQ3c). **Table 4** shows the results in detail.

### Students' Acceptance of Video Games in School Teaching (RQ4)

Students' acceptance of video games in school teaching did not change from pre- to post-test, $t = 0.21$, $p = 0.833$, $d = 0.05$ (RQ4a). Students' acceptance of video games in school teaching was about the scale's midpoint and numerically higher than students' acceptance of video games as a leisure activity and as significant part of life (RQ4b). Detailed results are depicted in **Table 5**.

## Discussion

In Study 1, we integrated the commercial video game Spore into a regular double lesson of a 10th-grade Biology course, in line with the curricular topic and learning objectives. We focused on students' reflection processes after video gaming (reflection-on-action) by means of analyzing a guided student discussion. Using text excerpts, we illustrated students' reactions to the teacher's three statements regarding the video game Spore. First, we found that playing Spore was fun, although we observed that students gave either very high or very low ratings. Yet, even a rather negative rating was put into perspective by saying that playing with friends, as in this case study, was fun, but playing alone would not be fun. Second, students were rather undecided regarding the statement that the game's illustrations of evolutionary processes were vivid but gave constructive critique about how the game could be improved. The third statement suggested using Spore in biology lessons, and one student who was familiar with the game noted that prolonging the gaming phase would hardly provide additional value for biological topics. The other students' statements indicated that students agreed with the way the game was integrated into school teaching and considered it as an object of reflection rather than a learning tool. While these three statements already triggered responses on several aspects of students' game experiences, students raised more points during the subsequent semistructured discussion. In particular, students found that topic-related technical terms were completely missing in the game and that visual representations of evolutionary factors were simplified. In line with this, students highlighted that the game should be more realistic in terms of evolutionary theory and graphics. Students further discussed that games help in visualizing topics and can be more engaging than learning by other means of school teaching.

We also examined complementary measures of student learning, learning motivation, and acceptance of video games. Students' self-rated topic knowledge was hardly influenced by video gaming and discussing. Open responses of the majority of students also indicated that neither video gaming nor discussing influenced the topic knowledge. Students' learning motivation

did not change regarding the lesson's topic. Students' game preference for school teaching and the perceived appropriateness of the game for school teaching were above the scales' midpoints. Students' perceived personal relevance of discussing was also above the scale's midpoint. Finally, students' acceptance of video games in school teaching did not change and was numerically higher than students' acceptance of video games as leisure activity and as significant part of life.

# STUDY 2

Digital history games provide students opportunities to interact with numerous possible interpretations of historical narratives and events (Kee, 2011). Since a game might display only one interpretation, one should foster critical reflection in history teaching in terms of encouraging active and reflective play, treating a game critically as one interpretation, and discussing, debriefing, and evaluating games (McCall, 2016). To gain complementary results on using commercial video games in formal high school teaching, we integrated Valiant Hearts: The Great War (Ubisoft Montpellier, 2014) into a double lesson of history teaching. As in Study 1, we grounded this decision on teacher's expertise, game reviews, a pedagogical review, and previous game-related conceptual and empirical research. Overall, both critics and user scores were positive (Metacritic, 2020b). A pedagogical review highlighted that the game deals responsibly with the topic violence and rather illustrates the fate of ordinary soldiers than showing exaggerated heroism (Spielbar, 2020b). Importantly, previous research criticized the use of stereotypes and oversimplifications in Valiant Hearts (Boltz, 2019) and outlined the importance of discussions about the game to reveal its commemorative character (Rughiniș and Matei, 2015). Overall, in line with curricular guidelines, we integrated Valiant Hearts into school teaching to scrutinize how students discuss the topic First World War and reflect on their game experiences.

## Materials and Methods

### Participants

The sample comprised 17 students from a 12th-grade advanced course on history at a German comprehensive school (high school level) ($M_{age}$ = 17.45, $SD_{age}$ = 0.50; eight female). The 16 students present at pre-test stated to play video games frequently in their leisure time ($M$ = 3.75, $SD$ = 1.48, 6.25% never) but seldom in school teaching ($M$ = 1.63, $SD$ = 0.62; 43.75% never), about the First World War ($M$ = 2.13, $SD$ = 1.31, 50% never), and the game Valiant Hearts ($M$ = 1.56, $SD$ = 1.26, 81.25% never, 6.25% occasionally, 6.25% frequently, and 6.25% very frequently).

### Design and Procedure

The design and procedure were identical to that of Study 1, except that the teacher assigned students to groups of two. The teacher ensured that each student played for a similar duration (20 min). We collaborated with a history teacher with personal interest in video games and some experience with using video games in school teaching. The collaboration with the teacher was initiated by the same announcement as in Study 1. The first author used the same observation protocol as in Study 1 and noticed technical issues at the beginning of the first lesson causing a short delay. During video gaming, students' conversations were about the puzzles in the game, and the teacher or students occasionally provided a hint to groups unable to solve a puzzle for a while.

### Teaching Materials

We developed similar teaching materials as in Study 1, in close collaboration with the teacher. We used Valiant Hearts: The Great War (Ubisoft Montpellier, 2014), a single player adventure and puzzle video game dealing with historical events of the First World War (age rating of 12). The game sequence comprised the declaration of war, the beginning of the war (from the perspective of a French soldier), the First Battle of the Marne (from the perspective of an US-American soldier), and the Battle of Neuve Chapelle (again from the perspective of the French soldier). To proceed in the game, players have to solve puzzles and interact with non-player characters. At various points in the game, players can access optional textual and pictorial information about facts and items related to the First World War. Each group of students played the game using a mobile tablet and in-ear headphones.

### Measures

The measures were almost identical to Study 1, except from the teacher's statements and topic-related items. The teacher stated the following three statements: "The game is boring", "Comic drawings do not do justice to the seriousness of the topic", and "The game Valiant Hearts does not belong to history lessons". During the subsequent semistructured discussion, the teacher asked one question about the game's goal (RQ1a), three questions addressing the game's content (RQ1b), and one question addressing games' perceived appropriateness for school teaching (RQ1c). Question wordings are depicted in **Table 6**. Unlike in Study 1, we analyzed four additional questions that the history teacher asked: "Which emotions does the game arouse, and by what?", "Does the game use stereotypes, and what for?", "How does the game position itself to violence and the experience of violence?", and "What historical narrative does this game convey?". We reformulated topic-related items by replacing the term "evolution" by "First World War".

## Results

### Content of the Discussion Phase (RQ1)

The discussion phase started with the three statements of the teacher. Following the teacher's first statement ("The game is boring"), only few students agreed. The first themes that emerged were "game's narrative" and "game enjoyment":

S1: Yes, I don't know, this is only from a French point of view I would say. [. . .] And that's always just from left to right, so, I didn't find it so exciting.
S12: Yes, actually I agree with S1.
T: Then, why don't you stand at seven or eight?
S12: Because in a certain way it was not that boring, because one had to do something. It wasn't that I just watched a film that might be boring [. . .]

In regard to this first excerpt, S1 criticized the game's narrative in terms of a rather one-sided perspective and reported little

**TABLE 6 |** Content of the semistructured discussion phase in Study 2.

| Category and subcategories | *n* | Teacher's questions and students' example statements |
|---|---|---|
| **Game's goals** | | What is the general goal of the game? What has allowed or hindered you to reach it? |
| Knowledge transfer | 2 | "Well, I think there were multiple goals. On the one hand, that the historical context was represented and taught" (S5)<br>"The goal of the game was simply to play through the time from 1914 to 1918 and to collect general information about everyday life, about that time, that year, so that one deepens one's knowledge and also learns something new" (S2) |
| Survival | 1 | "I think the goal of the game was also simply to survive, because one faces many situations throughout the game, where one, I'd say, would simply die" (S4) |
| Change in perspective | 2 | "That one gets a different point of view in the game" (S9) |
| **Game's content** | | At which points did the game, in favor of gameplay, represent historical events in an altered or simplified manner? Were there game scenes, in which historical events were perhaps even represented incorrectly? What perspective does the game take? |
| Content-based criticism | 1 | "So, the declaration of war was actually not explained in detail" (S10) |
| Content simplification via gameplay | 3 | "It was displayed in the upper right corner when the enemy was reloading the weapon, so that one knew when and when not to run" (S6) |
| Content representation | 3 | "I think it is good that the perspective from a normal person was taken. That could have been absolutely anybody during the First World War, with respect to the three perspectives taken in the game, and not a glorious hero who rushes over the battleground annihilating everything" (S4) |
| **Games' appropriateness for school teaching** | | Which changes would you make to the game if you were an historian? |
| Constructive critique | 9 | "I think that these historical aspects, one can basically read in the book, one can simply click away. [. . .] It could be explained using cutscenes, so that in case of events one does not have to read" (S11)<br>"Putting everything into a game—that would go beyond its scope. Because one cannot create a scene that takes an hour to understand the whole context" (S12)<br>"I would change the perspective somewhat, such that we look at it differently and that we cannot just run forward or run backward. Maybe that one can see the crowds of people from above [. . .] So, I forget that the masses of people who died, ran toward each other, and killed each other, that has not become so clear" (S10)<br>"If you find a mixture, so that it is graphically good, that one also wants to play this game, and this information is still in there, then one could learn the topic quite well" (S2) |
| Games' limitations | 2 | "I think one cannot learn enough with such a game, because in our first exam, we wrote a lot about the reasons, which were contained too little to feel well prepared and to take the exam after playing the game, what is not really the point of playing it, but yes" (S12)<br>"I think it is perhaps not so suitable that one would be prepared for an exam. [. . .] Maybe to get started, but not *per se*" (S14) |

*Categories and example statements are based on responses of 9 students, while in total, 17 students participated in the discussion.*

game enjoyment. After the teacher (T) questioned why S12 agreed with S1 while providing a lower rating (7.5 versus 5), S12 elaborated on game enjoyment and used a media comparison (games versus films) to outline that games allow for active usage. Students with ratings close to zero continued discussing:

S2: I found the game relatively good for learning, still now. Because time also simply passed now much faster than usual, simply because it was more interesting to play through it just like that. [. . .] But now we already knew a lot about the topic anyway, so I don't think it expanded our knowledge. But if we had done that at the beginning, so, also played at home in leisure time, then one could have taken something from that.

S3: Well, I thought that it was actually fun, that controls were limited, and one actually had a mission what to do. And it was actually fun (Teacher: Nice). So, for the fact that I am not a computer fan.

S4: I really liked the game. I also thought it was really good for learning. But I also looked at the entries, I skimmed over them, so I didn't just leave them out completely. [. . .] Because these entries were shown exactly when one asked oneself questions. [. . .]

S5: I wanted to add that the game can, I would say, also be played to the end, that is, to the conclusion. I think it's also good to, I'd say, refresh one's knowledge a bit, so that it gets reactivated. But I also think that some details have been mentioned that were perhaps not discussed in class.

Regarding the last excerpt, the theme "game for learning" occurred several times: S2 suggested that starting a teaching unit with this game could increase topic knowledge, S4 outlined the educational value of the contextual yet optional information provided by the game, and S5 argued how the game could be beneficial for learning in several ways. Related to the theme "game enjoyment", S2 referred to other school lessons and felt as if time passed faster because gaming was more interesting. As a non-gamer, S3 agreed and favored the simple controls and that goals of the game were clear.

Following the second statement of the teacher ("Comic drawings do not do justice to the seriousness of the topic"), one student (S5) was undecided (rating of 5):

S5: Because I think that although the music that was played there did indicate a certain seriousness, I don't think that the drawing itself can really convey this, that is to say, convey it to this extent, as it actually was or as we had now discussed.
T: Comic drawings do not do justice to the seriousness of the subject. The others all stand there almost at zero.
S6: One also has to remember that the game is played on a tablet and possibilities are not so huge for high-end graphics. [. . .]
S4: I think the graphic style didn't really influence that at all, because I think the colors, they were pretty gray and pretty sad. It was always like that, very serious and the comic graphics didn't really distract me. [. . .] The comic graphics don't play such a role when the colors are quite dark, and the music is really sad.

In sum, the students did not consider the comic style as inappropriate to display the topic, as emphasized by the teacher's statement. Students discussed the main theme "game graphics" and considered that graphics might also be simplified so that the game can be used on devices with less processing power (such as tablets). Students moreover outlined that the seriousness of the topic was mainly conveyed via the music and the color scheme of the game.

Following the teacher's third and final statement ("The game Valiant Hearts does not belong to history lessons"), the students discussed several aspects:

S7: Yes, in my opinion it just shows the different points of view of the different soldiers. It wasn't just France now, there was also Germany at the end, so if one had continued playing that, the USA was there of course. [. . .] So, I think that fits in there quite well.
S3: I also found the side information to be very informative, because one doesn't discuss everything in history class.
S8: I think one can really learn something from that. Well, first of all, we have the topic and so I think, one can play such a game now and then.

To sum up, students again discussed the themes "game's narrative" in terms of multiple viewpoints (S7) and "game for learning" in terms of the game's learning content (S3) and curriculum relatedness (S8). Overall, students found the game appropriate for history lessons.

Following the teacher's statements, the teacher continued to ask questions regarding the game's goals (RQ1a), game's content (RQ1b), and games' appropriateness for school teaching (RQ1c). As in Study 1, we summarized students' answers by means of subcategories and example statements, see **Table 6**: Students discussed knowledge transfer, survival, and change in perspective as goals of the game (RQ1a) and discussed the game's content in terms of content-based criticism, content simplification via gameplay, and content representation (RQ1b). Concerning games' appropriateness for school teaching (RQ1c), students mainly provided constructive critique and discussed limitations, in particular that playing the game would not suffice to prepare for the regular course exam.

Then, the teacher continued to discuss the additional questions on emotions, stereotypes, violence, and the game's narrative, for each of which we identified themes discussed by the students. Regarding the teacher's question on emotions ("Which emotions does the game arouse, and by what?"), students discussed the theme "inappropriate emotions":

S3: It wasn't really conveyed that it [the topic] is that serious and that it was supposed to evoke grief [. . .]
T: Okay. So, the emotions that you actually expected, they weren't there?
S3: Yes.
S12: I thought the game actually evoked the wrong emotions. I think one gets a lot of fun and one thinks 'yeah, cool', but that's just not what the First World War was. It's about that one doesn't really understand, sometimes not at all, which emotions are supposed to emerge, but simply this grief or despair or so. [. . .]
S9: So, I agree with S12, because when I had the conversation [in the game], there you [the observing student] also felt like beating up the one there or, I don't know, throwing a stick of dynamite somewhere or something. And, well, this is conveyed in a completely different way, as if one was in the mood for war, but this [the game] is actually only meant to show how bad it was back then. But that doesn't come across when one plays it. [. . .]
S5: Even if the emotions were not the right ones, the music and certain scenes that were being played back, where one couldn't play—the cutscenes—emotions are rather shown in that way.

In relation to the last excerpt, S3 experienced different emotions during gaming than expected regarding the topic. S12 and S9 once more articulated the theme "game enjoyment" and that gaming evoked fun instead. With regard to the topic, this could limit the understanding (S12) or convey an inappropriate mood (S9). Finally, S5 agreed but highlighted that the music and cutscenes of the game also communicated emotions. The theme "reflections on game experiences" also emerged:

S4: I think one has to differentiate very strongly between what one feels when one plays and what one feels when one really thinks about what actually happened. Well, games should be fun, that's the point of it, it's a medium of entertainment. If a game would not be fun, nobody would play it, that is logical. [. . .] And if one just thinks about what just happened, what one did, for just a second, then it can bring about grief and a bit of pity, if one really thinks about what just happened. And one didn't really notice that people died there, for example, when one ran along there, because one was just concentrated on surviving and that was mostly the

case during the war. But when one has thought about it afterward, one feels a little sad. [. . .]

S12: Well, say I'm talking to S2 about it [the game] and then, uh, we wouldn't both start reflecting on it, but just say, 'yes, I killed that one, haha, it was really funny, I kind of knocked him out with one punch'. I don't think that we teenagers or in general one would sit down there and think about it all over again. [. . .] I don't think that one would really become aware of it, because that's just, they all say 'it's just a game', so, I don't have to die for real.

Summing up, S4 made the point that it is important to distinguish between emotions during gaming and after gaming. S12 then outlined a hypothetical example, which indicates that teenaged students might not exchange serious considerations after gaming (reflection-on-action) on their own.

Regarding the teacher's question on stereotypes ("Does the game use stereotypes, and what for?"), students first named stereotypic properties of game characters (e.g., "wine" and "baguette" were related to French people). Then, the teacher elaborated on the original question: "One would not have had to do it. What does the game do that for? Why do the developers do this?", and students discussed:

S2: In any case, it also clarifies the character of these, of these squads, by somehow bringing in this French war music, or so, so one notices that these are the French.

S6: I think the developers have done that to make it a bit more humorous. I don't think it was about making the Germans look more like Germans. I think the game did that quite well.

S4: One also has to consider that the game almost contains no spoken language. [. . .] And since the game should also work for younger people, [. . .] one has to make it clear that even someone who doesn't know about history knows 'I see, they are Germans, oh, now I'm going to play a French' — that this is also correctly understood. And also to make it more humorous, I also agree with that.

In sum, the students discussed the theme "functions of stereotypes". They mentioned that stereotypic properties facilitated to correctly identify game characters' nationalities (e.g., French, German, and US-American), which could support younger people and people without history knowledge to understand the game's narrative. Students also discussed that stereotypes can add some humor.

Following the teacher's question on violence ("How does the game position itself to violence and the experience of violence?"), the students discussed:

S6: One only has to knock them out, the goal is just to get through the game with as little violence as possible. That makes the game pretty good. [. . .]

S2: If one wants to teach the eighth graders something and then somehow, I don't know, bring something really brutal into it and make the graphics realistic, then I think they are really disturbed and don't come to school [. . .] or some of them prefer to come to school, that's also possible. [. . .]

S4: One should keep in mind that although one doesn't use violence oneself, not really, one's opponents indeed do. That they, they really come with machine guns and bombs.

To summarize, the main theme was "elements of violence", and students mentioned that there was relatively little violence in the game and that, as players of the game, one used little violence. Students also discussed the theme "effects of violence" and that adding more brutality and more realistic graphics to the game could negatively affect younger students.

Regarding the teacher's question on the game's narrative ("What historical narrative does this game convey?"), the teacher also asked students to relate the game's content to content of previous lessons:

T: Can you relate this to the graves and monuments, with which monument would you more likely associate this and with which not at all? Why?

S4: Well, I think, when one plays the game, one should also make one's own opinion of how one thinks it is. And about the monuments, I think of the monuments that we looked at in class, it doesn't fit so well, because the monuments always made it very clear that war is something good or bad. [. . .]

S1: I think that the [monument of a] mourning woman fits quite well. Because in the beginning one sees how the family is torn apart and how they have to stay alone. So, I think it fits, because the monument was aimed at all nations. And I think that in turn also fits into the game.

To sum up, the theme "normative ethics" emerged, while students did not think that the game clearly communicated whether war was good or bad. They also discussed the theme "related course content" by referencing the game's narrative to monuments the students had discussed in previous lessons.

### Influence of Gaming and Discussing on Students' Perceived Topic Knowledge (RQ2)

Regarding the influence of the video gaming phase on students' perceived topic knowledge (RQ2a), there was no change in students' self-rated topic knowledge from pre-test to post-game, $t(15) = -0.44$, $p = 0.669$, $d = 0.11$ (see **Figure 2A**). Students' perceived impact of video gaming on topic knowledge was numerically slightly below the scale's midpoint at post-game and post-test (see **Figure 2B**). Students' open responses revealed that playing the game increased most students' perceived topic knowledge (post-game: 58.82%; post-test: 64.71%) (see **Figure 2D**). Based on the free recall, students remembered one (47.06%), two (17.65%), or three topic-related aspects (17.65%) they learned by playing the game; the other students stated that playing motivated for further investigations (5.88%) or could avoid boredom (5.88%). Of the students, 5.88% gave no response.

Regarding the influence of the discussion phase on students' perceived topic knowledge (RQ2b), we found no change in self-rated topic knowledge from post-game to post-test, $t(15) = 1.86$,

$p = 0.083$, $d = 0.47$ (see **Figure 2A**). Students' perceived impact of the discussion on topic knowledge was numerically slightly below the scale's midpoint ($M = 2.71$, $SD = 0.92$). About half of the students (52.93%) stated that discussing increased their knowledge in terms of an exchange of player perspectives (29.41%), content-related depth (11.76%), and reactivation of knowledge (11.76%); the other students (47.06%) stated no knowledge increase (see **Figure 2D**).

At post-discussion, a higher number of students indicated to have learned more through video gaming (58.82%) than through discussing (11.76%), while the other students stated a similar influence (11.76%) or no influence (17.65%) (RQ2c).

### Students' Learning Motivation Regarding Lesson's Topic, Gaming, and Discussing (RQ3)

Measures of students' learning motivation regarding the lesson's topic did not change from pre- to post-test, all $|ts| \leq 1.46$, $ps \geq 0.164$, $ds \leq 0.37$ (RQ3a) (see **Table 2**). Students' game enjoyment, perceived competence, and all satisfaction measures were above the scales' midpoints (all $ps \leq 0.037$) (RQ3b) (see **Table 3**). Students' personal interest and personal relevance of discussing was also above the scales' midpoints (all $ps \leq 0.001$) (RQ3c) (see **Table 4**).

### Students' Acceptance of Video Games in School Teaching (RQ4)

Students' acceptance of video games in school teaching did not change from pre- to post-test, $t = -0.81$, $p = 0.432$, $d = 0.20$ (RQ4a). Students' acceptance of video games in school teaching was numerically higher than the scale's midpoint and numerically higher than students' acceptance of video games as a leisure activity and as significant part of life (RQ4b). The detailed results are depicted in **Table 5**.

## Discussion

In Study 2, we integrated the commercial video game Valiant Hearts: The Great War (Ubisoft Montpellier, 2014) into a regular double lesson of a 12th-grade advanced course on history, in line with the curricular topic and learning objectives. The reflection processes of the students were triggered by the teacher's three statements about the game experiences. First, a student found the game rather boring due to a one-sided narrative, while other students discussed several reasons why they were excited by playing the game: time went by faster, controls were simple, the game's goals were clear, and the game provided contextual learning content. Second, being asked about an illustrative feature of the game (comic drawings), most students found this feature appropriate to convey the serious topic of the First World War and outlined that the music and the color scheme of video games also contribute to an overall game experience. Third, students rather argued in favor of Valiant Hearts belonging to history lessons, since the game shows different viewpoints, provides topic-related information, and the students saw an overall educational value of playing such games occasionally in school teaching. During the subsequent semistructured discussion, students criticized the events displayed in the game as partially imprecise or simplified, found that historical facts were missing, that attending the learning content is optional in the game, and that playing such games could not suffice for exam preparation. In turn, they discussed that particularly important facts should become mandatory to prevent that players are missing those.

In reaction to the teacher's additional question on emotion, students discussed how different multimedia features of the game evoked emotions and raised two important points: On the one hand, students reflected on the game being an entertainment product that evoked fun as a positive emotion, instead of the expected negative emotions related to the First World War—such as grief and despair. On the other hand, another student outlined the importance to distinguish between emotions during gaming and after gaming, similar to the key differentiation between reflection-in-action and reflection-on-action that we introduced in this case study. Moreover, a student argued that reflection might not even take place by itself and without reflection triggers, which is in line with findings on how players reflected on game experiences from leisure time (Mekler et al., 2018). Stereotypes are part of the game (Boltz, 2019), and students discussed two functions: Stereotypes could facilitate correct assignment of game characters to groups or nationalities and might add some humor to the game. Students also realized and favored that the game does not endorse violence, since players make significantly less use of violence than players' opponents. Finally, evaluating the game's narrative allowed students to reactivate knowledge by relating the game's content to previous lessons' content. Overall, while students actively discussed multiple relevant aspects related to the game after playing it, we would like to highlight that students themselves pointed out the value of such a discussion phase, also since reflection might not take place by itself and not during gaming.

Regarding student learning, students' perceived topic knowledge was hardly influenced by video gaming and discussing. Still, most students openly stated that the game increased their topic knowledge, and most students recalled one to three topic-related aspects they learned through gaming. Students furthermore stated that discussing allowed for exchanging player perspectives, content-related depth, and reactivating knowledge. More students stated that they learned more through video gaming than through discussing. Students' learning motivation regarding the lesson's topic did not change. Students' ratings for playing the game again in school teaching and game's appropriateness for school teaching were above the scales' midpoints, while qualitative results suggest that some students doubted the game's appropriateness for school teaching. Students' personal interest as well as relevance of discussing were also above the scales' midpoints. Finally, students' acceptance of video games in school teaching did not change and was numerically higher than students' acceptance of video games as leisure activity and as significant part of life.

# GENERAL DISCUSSION

It has been discussed for years *that* commercial video games can be used for learning (Boyle et al., 2016; Clark et al., 2016) and for formal education (Gee, 2005; Squire, 2008; van Eck, 2009;

Hainey et al., 2016), while our case studies addressed the current need for more research on *how* to integrate these games into formal educational settings (Mayer, 2019). Our case studies illustrate how commercial video games might serve as objects of reflection in formal school teaching, in line with the curricular topic and learning objectives: Student learning and reflection can be guided by means of worksheets during a gaming phase, and a following discussion phase moderated by a teacher can allow students to critically discuss and reflect on their game experiences. With regard to this focal question, we next condense the students' reflections on their game experiences to formulate several suggestions in regard to using commercial video games as objects of reflection.

Concerning student reflection, we found that some central statements of a teacher about the quality of the game can already trigger student reflections, for instance reactions on the games' level of entertainment, their appropriateness of visualizing a lesson's topic, and their suitability for teaching. More specifically, students may discuss multiple goals of the games (RQ1a), how the games represent learning content (RQ1b), and evaluate the game from an expert's perspective (RQ1c). First, students can identify multiple goals in games so that one could trigger and guide their reflections not only on different strategies of playing games (Kiili, 2007) but also on relations between their game experiences and real events, behaviors, or decisions. Reflective guidance (e.g., by means of verbal instructions and worksheets as in our studies) can support students to pursue learning objectives instead of performance goals or no goals, which may reduce task complexity and increase fun in game-based learning contexts (Nebel et al., 2017). Second, playing and discussing games in school teaching can allow students to share experiences on whether and how games represent content related to a subject matter in a simplified way. Third, rather than relying on individual game experiences, playing games in school teaching can allow students to formulate, discuss, and reflect on constructive critiques and own ideas for modifications of the games together with fellow students and teachers. That most of the students' statements belonged to this category also points to the potential of letting students design own games (Prensky, 2008; An, 2016) as well as to the benefits of perspective taking and using games as "objects-to-think-with" (Dishon and Kafai, 2020, p. 1). Notably, it seems valuable to also address more general topics such as emotions, stereotypes, violence, and narratives when discussing video games. More generally, we may speculate that student discussions could foster students' media literacy in terms of procedural knowledge, that is, how to analyze commercial video games' goals and content and how to identify their potentials and limitations for learning. Taken together, we conclude that integrations of commercial video games into formal educational settings could include a subsequent discussion phase to let students share game experiences and reflect on the games' goals, content, and appropriateness for school teaching.

Concerning student learning, only qualitative results of Study 2 provided partial evidence that video gaming (RQ2a) and discussing (RQ2b) increased students' topic knowledge. Furthermore, only in Study 2 students stated that they learned more through video gaming than through the discussion (RQ2c). Overall, while our non-significant quantitative results are in line with previous meta-analytic findings suggesting single play sessions being insufficient to increase learning outcomes (Wouters et al., 2013; Clark et al., 2016), mainly the qualitative results of Study 2 offered partial evidence that a single play session, followed by a guided student discussion, could increase students' topic knowledge. Moreover, most corresponding effect sizes were small in both studies. Only in case of the non-significant influence of the discussion phase on students' self-rated topic knowledge in Study 2, the effect size was moderate. This might partly rely on the game itself or on the fact that the discussion contained additional questions on several game- and content-related topics, as compared to Study 1. While in Study 1 more students responded that gaming influenced their topic knowledge after gaming than four days later (post-game versus post-test), in Study 2, the responses were more stable across time. Notably, students in both studies found the games appropriate for school teaching, whereas students in Study 2 noted that playing the game might not suffice for exam preparation. Taken together, guided discovery learning through video gaming alone seems to hardly benefit students' topic knowledge in formal educational settings.

Concerning students' learning motivation, we considered main factors of learning motivation (Keller, 2016) regarding the lessons' topics, video gaming, and discussing. First, students' learning motivation regarding the lessons' topics did not change (RQ3a), which relates to other studies in which single play sessions did not change students' motivation (Wouters et al., 2013). Second, students' learning motivation regarding video gaming was relatively high in both studies (RQ3b) in terms of game enjoyment, perceived competence, game interest, game preference for school teaching, and games' appropriateness for school teaching. In specific, high game enjoyment indicates that students were intrinsically motivated, and high perceived competence indicates that students were able to handle the games' challenges and controls (Ryan et al., 2006). Third, students' learning motivation regarding discussing (RQ3c) was about the scales' midpoints in Study 1 and above the scales' midpoints in Study 2. Overall, our results do not show that our way of using and discussing commercial video games for a double lesson changes students' learning motivation but could allow students to experience high enjoyment, competence, and satisfaction.

Concerning students' acceptance of video games, we provide results from the context of use that are often lacking in studies on video game acceptance (Bourgonjon et al., 2010; Wang and Goh, 2017). In our studies, students were undecided (Study 1) or tended to accept video games in school teaching (Study 2), while acceptance in school teaching was highest (versus as leisure activity and as significant part of life) (RQ4). Considering that students gave high ratings for game preference for school teaching and game's appropriateness for school teaching (RQ3b), it overall seems that students could welcome the use of commercial video games in school teaching.

Taken together, our case studies demonstrate *how* commercial video games can be used as objects of reflection in formal educational settings. Importantly, we employed mixed methods to thoroughly investigate how playing and discussing commercial

video games in formal educational settings might affect learning and reflection processes. Our results indicate that at least some commercial video games are ineffective in disseminating declarative knowledge and should not be regarded nor be used as content providers. Rather such commercial video games could serve as objects of reflection in formal educational settings, for instance, by having teachers moderating student discussions on games' goals, content, and appropriateness for school teaching. Considering that teachers with gaming experience and technological competence can facilitate successful integrations of commercial video games (van Eck, 2009), other commercial video games could be used in similar ways to teach evolution (Leith et al., 2016), history (Schrier, 2014), or other topics of formal education. While commercial video games are not frequently used in high school teaching to teach twenty-first century skills (Qian and Clark, 2016), more such integrations could allow to foster students' reflection processes or related critical thinking skills. Overall, appropriate uses and discussions of video games could facilitate to meet the media reality of many students as well as to foster key media literacy skills (Squire, 2008).

## Limitations and Future Research

Both presented case studies are conceptually similar and provide results from a basic (Study 1) and advanced high school course (Study 2). Both case studies were of a specific duration (one double lesson) and contained participants in a specific age range (15- to 18-year-olds). We moreover investigated convenience samples with rather small sample sizes. However, we would like to highlight that our samples reflect sizes of regular school classes and that effects should also be detectable at the level of school classes to be of practical relevance. Although our measures were not standardized, we used a rigorous translation–back translation process to transform validated scales into direct rating scales, resulting in acceptable internal consistency in terms of Cronbach's alpha (Tavakol and Dennick, 2011). Still, our results do only rely on students' self-reports and do not include objective test results such as, for instance, results from a test on topic knowledge or on media literacy competence. While we triangulated qualitative and quantitative methods and data, we did not consider any teacher ratings nor in-game measures of student learning. Overall, limited generalizability and comparability of related results in this research field is a common yet critical problem and partly can be traced back to the issue of a missing unequivocal terminology (cf. Rüth and Kaspar, 2017). In this regard, there may be conceptual overlaps between the terms commercial video games and serious games, the latter aiming at maximizing learning and motivation effects by blending entertainment and learning (Breuer and Bente, 2010). Indeed, while the appropriateness of the term serious game is a matter of debate in game science (Klabbers, 2018), in practice, one can identify aspects that commercial video games share with serious games (Ulrich and Helms, 2017) and assess their educational value (Becker, 2017a; Rüth, 2017). Using this rationale, we noted in advance that learning from the game in Study 1 is problematic due to its biased representations of evolution (Schrader et al., 2016), but guided learning with the game is conceivable (Schrader and McCreery, 2012). Similarly, we noted that the game we used in Study 2 was a commercial video game that contained optional textual and pictorial information related to the topic. In this sense, Study 2 was not only complementary in terms of investigating another topic and a more advanced course but also in terms of using a commercial video game with more elaborate learning features than the game used in Study 1. Similarly, future research might benefit from evaluating games' features to further examine commercial video games' roles and effects.

To scrutinize how commercial video games could be best integrated into formal school teaching and other formal learning settings, further research is needed. Future investigations could go beyond the dimension of reflection-on-action and use other measures of reflective learning (Brockbank and McGill, 2007a) to provide more insights, for instance, on different levels of reflection processes in game-based learning (Mekler et al., 2018). Such investigations could also evaluate effects based on intermediate results such as game scores (formative evaluation) or based on final results such as the self-reports in our pre–post design (summative evaluation). Overall, several measures before, during, and after gaming as well as other aspects of the study design can be considered (e.g., All et al., 2016; Rüth, 2017). Measurement time points yet should be chosen carefully, and future studies could use measures at different time points, for instance, to investigate whether and how the perceived influence of gaming on learning changes over time. To test appropriate instructional approaches, one could adapt guidelines for lesson plans, for example, on teaching evolution (Nelson, 2012; Southerland and Nadelson, 2012) or history (McCall, 2016; Carvalho, 2017). Notably, several studies on game-based learning are missing links to a learning theory (Wu et al., 2012) so that future research could aim at supplementing video gaming with another instructional method (such as the guided discovery learning of our case studies) and let students work in groups, both being ways to increase learning (Wouters et al., 2013). However, regarding instructional guidelines, a specific learning instruction might not be beneficial for learning when using commercial video games (Hawlitschek and Joeckel, 2017). Future studies could adapt instructional approaches across multiple lessons and investigate longitudinal effects of video gaming and discussing on student reflection as well as student learning and motivation. While we focused on students' statements and self-ratings, one could also investigate how students create and share artifacts (e.g., screenshots or notes) and trace players' hypotheses and decisions more precisely (Schrader and McCreery, 2012). Further, such subjective measures could be complemented with teacher ratings of student performance and objective test results as well as in-game metrics. For instance, students' self-reported dance skill and game score were found to change differently over time across four game-based lessons in a formal educational setting (Rüth and Kaspar, 2020). One could also ask teachers to think aloud, for instance, retrospectively on their experiences about teaching with the game or concurrently when playing a game to explore how they evaluate its educational value. Furthermore, video gaming can be integrated into formal education via digital environments

as in flipped classroom settings (e.g., students play games at home and discuss them in the classroom) or distance learning settings (e.g., teachers guide game sessions or discussions via video conferences). One could then examine video recordings from such settings, for instance, to investigate the quantity and quality of student–student or teacher–student interactions and to compare those between game-based and non-game settings. Finally, there are other lines of game-based research (cf. Mayer, 2019), and one might compare how students reflect on video games containing different features or on video games and other media, such as interactive e-books or interactive videos.

To conclude, we presented two case studies to demonstrate how commercial video games could be integrated as objects of reflection into formal school teaching. We illustrated how students reflected on their game experiences and on commercial video games' goals, content, and appropriateness for school teaching. It overall seems that teachers who provide students guidance during gaming and discussing allow students reflection processes that might foster their media literacy skills. In this regard, using commercial video games as objects of reflection appears to be a promising approach to foster student reflection in future formal education.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

Ethics board approval was not required for the study on human participants in accordance with the local legislation and institutional requirements and as the video games' integration into school teaching was in accordance with the formal school curriculum. Written informed consent to participate in this study was provided by the students' legal guardians/next of kin and all students voluntarily participated in this study.

## AUTHOR CONTRIBUTIONS

MR and KK designed the study and interpreted the results. MR conducted the literature review, organized and supervised the data collection, performed the statistical analyses, and drafted and revised the manuscript. KK supervised the statistical analyses, revised the manuscript, and acquired the funding. Both authors contributed to the article and approved the submitted version.

## FUNDING

This work was partly funded by the Federal State Government of North-Rhine Westphalia (NRW, Germany) via the Grimme Research College (Project ID: D-73795-K-710, Grant Recipient: KK). The funding source was not involved in any aspect of the project and did not have any input on the interpretation or publication of the research results.

## ACKNOWLEDGMENTS

We would like to thank the teachers for their collaboration as well as Sandra Schwab and Mareike Westfal for their support with data collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.594013/full#supplementary-material

## REFERENCES

Alfieri, L., Brooks, P. J., Aldrich, N. J., and Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *J. Educ. Psychol.* 103, 1–18. doi: 10.1037/a0021017

All, A., Castellar, E. P. N., and Van Looy, J. (2016). Assessing the effectiveness of digital game-based learning: best practices. *Comput. Educ.* 92, 90–103. doi: 10.1016/j.compedu.2015.10.007

An, Y.-J. (2016). A case study of educational computer game design by middle school students. *Educ. Technol. Res. Dev.* 64, 555–571. doi: 10.1007/s11423-016-9428-7

Arnseth, H. C., Hanghøj, T., and Silseth, K. (2018). "Games as tools for dialogic teaching and learning," in *Games and Education: Designs in and for Learning*, eds H. C. Arnseth, T. Hanghøj, T. D. Henriksen, M. Misfeldt, R. Ramberg, and S. Selander (Leiden: Brill), 123–139. doi: 10.1163/9789004388826_008

Bean, T. E., Sinatra, G. M., and Schrader, P. G. (2010). Spore: spawning evolutionary misconceptions? *J. Sci. Educ. Technol.* 19, 409–414. doi: 10.1007/s10956-010-9211-1

Becker, K. (2017a). "Evaluating games," in *Choosing and Using Digital Games in the Classroom. A Practical Guide*, ed. K. Becker (Cham: Springer), 151–173. doi: 10.1007/978-3-319-12223-6_6

Becker, K. (2017b). "Digital game-based learning: learning with games," in *Choosing and Using Digital Games in the Classroom. A Practical Guide*, ed. K. Becke (Cham: Springer), 25–61. doi: 10.1007/978-3-319-12223-6_2

Becker, K. (2017c). "Digital game pedagogy: teaching with games," in *Choosing and Using Digital Games in the Classroom. A Practical Guide*, ed. K. Becke (Cham: Springer), 63–97. doi: 10.1007/978-3-319-12223-6_3

Berg Marklund, B., and Alklind Taylor, A.-S. (2016). Educational games in practice: the challenges involved in conducting a game-based curriculum. *Electron. J. Learn.* 14, 122–135.

Bohannon, J. (2008). Flunking SPORE. *Science* 322:531. doi: 10.1126/science.322.5901.531b

Boltz, L. O. (2019). "Nervousness and maybe even some regret: videogames and the cognitive-affective model of historical empathy," in *Exploring the Cognitive, Social, Cultural, and Psychological Aspects of Gaming and Simulations*, ed. B. R. Dubbels (Hershey, PA: IGI Global), 228–251. doi: 10.4018/978-1-5225-7461-3.ch008

Bourgonjon, J., Valcke, M., Soetaert, R., and Schellens, T. (2010). Students' perceptions about the use of video games in the classroom. *Comput. Educ.* 54, 1145–1156. doi: 10.1016/j.compedu.2009.10.022

Boyle, E. A., Connolly, T. M., Hainey, T., and Boyle, J. M. (2012). Engagement in digital entertainment games: a systematic review. *Comput. Hum. Behav.* 28, 771–780. doi: 10.1016/j.chb.2011.11.020

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., et al. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Comput. Educ.* 94, 178–192. doi: 10.1016/j.compedu.2015.11.003

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Q. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Breuer, J., and Bente, G. (2010). Why so serious? on the relation of serious games and learning. *J. Comput. Game Cult.* 4, 7–24.

Brockbank, A., and McGill, I. (2007a). "Methods and assessment of reflective learning," in *Facilitating Reflective Learning in Higher Education*, 2nd Edn, eds A. Brockbank and I. McGill (Berkshire, UK: Open University Press), 164–205.

Brockbank, A., and McGill, I. (2007b). "Reflection and reflective practice," in *Facilitating Reflective Learning in Higher Education*, 2nd Edn, eds A. Brockbank and I. McGill (Berkshire, UK: Open University Press), 85–107.

Brockbank, A., and McGill, I. (2007c). "What is learning? – a review of learning theories," in *Facilitating Reflective Learning in Higher Education*, 2nd Edn, eds A. Brockbank and I. McGill (Berkshire, UK: Open University Press), 38–63.

Bureau of Labor Statistics (2019). *Average Daily Time Spent Playing Games and Using Computer for Leisure Per Capita in the United States in 2019, by Age Group*. Available online at: https://www.statista.com/statistics/502149/average-daily-time-playing-games-and-using-computer-us-by-age/ (accessed November 06, 2020).

Caldwell, K. E. H., Osterweil, S., Urbano, C., Tan, P., and Eberhardt, R. (2017). ""I just don't know where to begin": designing to facilitate the educational use of commercial, off-the-shelf video games," in *Serious Games and Edutainment Applications: Volume II*, eds M. Ma and A. Oikonomou (Cham: Springer), 625–648. doi: 10.1007/978-3-319-51645-5_27

Carvalho, V. M. (2017). Videogames as tools for social science history. *Historian* 79, 794–819. doi: 10.1111/hisn.12674

Ching, D. (2012). Passion play: will wright and games for science learning. *Cult. Stud. Sci. Educ.* 7, 767–782. doi: 10.1007/s11422-012-9456-5

Chittaro, L., and Sioni, R. (2018). Existential video games: proposal and evaluation of an interactive reflection about death. *Entertain. Comput.* 26, 59–77. doi: 10.1016/j.entcom.2018.01.004

Clark, D. B., Tanner-Smith, E. E., and Killingsworth, S. S. (2016). Digital games, design, and learning: a systematic review and meta-analysis. *Rev. Educ. Res.* 86, 79–122. doi: 10.3102/0034654315582065

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Measurement* 20, 37–46. doi: 10.1177/001316446002000104

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. New York, NY: Routledge Academic.

De Jong, T., and Lazonder, A. W. (2014). "The guided discovery learning principle in multimedia learning," in *The Cambridge Handbook of Multimedia Learning*, 2nd Edn, ed. R. E. Mayer (New York, NY: Cambridge University Press), 371–390. doi: 10.1017/cbo9781139547369.019

Dishon, G., and Kafai, Y. B. (2020). Making more of games: cultivating perspective-taking through game design. *Comput. Educ.* 148:103810. doi: 10.1016/j.compedu.2020.103810

Flick, U., Garms-Homolová, V., Herrmann, W. J., Kuck, J., and Röhnsch, G. (2012). "I can't prescribe something just because someone asks for it...": using mixed methods in the framework of triangulation. *J. Mixed Methods Res.* 6, 97–110. doi: 10.1177/1558689812437183

Gee, J. P. (2005). Learning by design: good video games as learning machines. *Learn. Digit. Media* 2, 5–16. doi: 10.2304/elea.2005.2.1.5

Hainey, T., Connolly, T. M., Boyle, E. A., Wilson, A., and Razak, A. (2016). A systematic literature review of games-based learning empirical evidence in primary education. *Comput. Educ.* 102, 202–223. doi: 10.1016/j.compedu.2016.09.001

Hawlitschek, A., and Joeckel, S. (2017). Increasing the effectiveness of digital educational games: the effects of a learning instruction on students' learning, motivation and cognitive load. *Comput. Hum. Behav.* 72, 79–86. doi: 10.1016/j.chb.2017.01.040

Herrero, D., del Castillo, H., Monjelat, N., Garca-Varela, A., Checa, M., and Gómez, P. (2014). Evolution and natural selection: learning by playing and reflecting. *J. New Approaches Educ. Res.* 3, 26–33. doi: 10.7821/naer.3.1.26-33

Huizenga, J. C., Ten Dam, G. T. M., Voogt, J. M., and Admiraal, W. F. (2017). Teacher perceptions of the value of game-based learning in secondary education. *Comput. Educ.* 110, 105–115. doi: 10.1016/j.compedu.2017.03.008

Johnson, C. I., and Mayer, R. E. (2010). Applying the self-explanation principle to multimedia learning in a computer-based game-like environment. *Comput. Hum. Behav.* 26, 1246–1252. doi: 10.1016/j.chb.2010.03.025

Kee, K. (2011). Computerized history games: narrative options. *Simul. Gaming* 42, 423–440. doi: 10.1177/1046878108325441

Keller, J. M. (2016). Motivation, learning, and technology: applying the ARCS-V motivation model. *Particip. Educ. Res.* 3, 1–15. doi: 10.17275/per.16.06.3.2

Kiili, K. (2007). Foundation for problem-based gaming. *Br. J. Educ. Technol.* 38, 394–404. doi: 10.1111/j.1467-8535.2007.00704.x

Kiili, K., Ketamo, H., and Lainema, T. (2011). "Reflective thinking in games: triggers and constraints," in *Leading Issues in Games Based Learning*, ed. T. Connolly (Reading, UK: Academic Publishing International), 178–193.

Klabbers, J. H. (2018). On the architecture of game science. *Simul. Gaming* 49, 207–245. doi: 10.1177/1046878118762534

Kuckartz, U. (2016). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung [Qualitative content analysis: Methods, practice, computer assistance]*, 3rd Edn. Weinheim: Beltz Juventa.

Leith, A. P., Ratan, R. A., and Wohn, D. Y. (2016). The (de-)evolution of evolution games: a content analysis of the representation of evolution through natural selection in digital games. *J. Sci. Educ. Technol.* 25, 655–664. doi: 10.1007/s10956-016-9620-x

Maxis (2008). *Spore [video game]*. Redwood City: Electronic Arts, Inc.

Mayer, R. E. (2019). Computer games in education. *Annu. Rev. Psychol.* 70, 531–549. doi: 10.1146/annurev-psych-010418-102744

McCall, J. (2016). Teaching history with digital historical games: an introduction to the field and best practices. *Simul. Gaming* 47, 517–542. doi: 10.1177/1046878116646693

Mekler, E. D., Iacovides, I., and Bopp, J. A. (2018). ""A game that makes you question. . . ": exploring the role of reflection for the player experience," in *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, (New York, NY: ACM), 315–327. doi: 10.1145/3242671.3242691

Metacritic (2020a). *Spore for PC Reviews*. Available online at: https://www.metacritic.com/game/pc/spore (accessed November 06, 2020).

Metacritic (2020b). *Valiant Hearts: The Great War for iPhone/iPad Reviews*. Available online at: https://www.metacritic.com/game/ios/valiant-hearts-the-great-war (accessed November 06, 2020).

Moreno, R., and Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *J. Educ. Psychol.* 97, 117–128. doi: 10.1037/0022-0663.97.1.117

Nebel, S., Schneider, S., Schledjewski, J., and Rey, G. D. (2017). Goal-setting in educational video games: comparing goal-setting theory and the goal-free effect. *Simul. Gaming* 48, 98–130. doi: 10.1177/1046878116680869

Nelson, C. E. (2012). "Why don't undergraduates really "get" evolution? what can faculty do?," in *Evolution Challenges: Integrating Research and Practice in Teaching and Learning about Evolution*, eds K. S. Rosengren, S. K. Brem, E. M. Evans, and G. M. Sinatra (New York, NY: Oxford University Press), 311–347. doi: 10.1093/acprof:oso/9780199730421.003.0014

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annu. Rev. Psychol.* 49, 345–375. doi: 10.1146/annurev.psych.49.1.345

Panoutsopoulos, H., and Sampson, D. (2012). A study on exploiting commercial digital games into school context. *Educ. Technol. Soc.* 15, 15–27.

Pea, R. D. (1993). "Practices of distributed intelligence and designs for education," in *Distributed Cognitions: Psychological and Educational Considerations*, ed. G. Salomon (Cambridge: Cambridge University Press), 47–87.

Pedagogical Media Research Center Southwest (2019). *JIM-Studie 2019. Jugend, Information, Medien. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger in Deutschland. [JIM Study 2019. Youth, Information, Media. Basic study on media use among 12- to 19-year-olds in Germany]*. Available online at: https://www.mpfs.de/fileadmin/files/Studien/JIM/2019/JIM_2019.pdf (accessed November 06, 2020).

Peters, V. A., and Vissers, G. A. (2004). A simple classification model for debriefing simulation games. *Simul. Gaming* 35, 70–84. doi: 10.1177/1046878103253719

Pew Research Center (2018). *Percentage of Teenagers who Play video Games in the United States as of April 2018, by gender*. Available online at: https://www.

statista.com/statistics/454309/teen-video-game-players-gender-usa/ (accessed November 06, 2020).

Poli, D., Berenotto, C., Blankenship, S., Piatkowski, B., Bader, G. A., and Poore, M. (2012). Bringing evolution to a technological generation: a case study with the video game SPORE. *Am. Biol. Teach.* 74, 100–103. doi: 10.1525/abt.2012.74.2.7

Prensky, M. (2008). Students as designers and creators of educational computer games: who else? *Br. J. Educ. Technol.* 39, 1004–1019. doi: 10.1111/j.1467-8535.2008.00823_2.x

Qian, M., and Clark, K. R. (2016). Game-based Learning and 21st century skills: a review of recent research. *Comput. Hum. Behav.* 63, 50–58. doi: 10.1016/j.chb.2016.05.023

Reinecke, L., Tamborini, R., Grizzard, M., Lewis, R., Eden, A., and Bowman, N. D. (2012). Characterizing mood management as need satisfaction: the effects of intrinsic needs on selective exposure and mood repair. *J. Commun.* 62, 437–453. doi: 10.1111/j.1460-2466.2012.01649.x

Rughiniş, R., and Matei, Ş (2015). "Play to remember: the rhetoric of time in memorial video games," in *Human-Computer Interaction: Interaction Technologies. HCI 2015. Lecture Notes in Computer Science*, ed. M. Kurosu (Cham: Springer), 628–639. doi: 10.1007/978-3-319-20916-6_58

Rüth, M. (2017). "Spielerisches lernen besser bewerten: effektivität und effizienz von computerspielen [better evaluation of game-based learning: on effectiveness and efficiency of video gaming]," in *Spielend lernen! Computerspiele(n) in Schule und Unterricht. Schriftenreihe zur digitalen Gesellschaft NRW (Band 5) [Game-based learning! Video Games in school and school teaching. Publication series of the digital society NRW]*, eds W. Zielinski, S. Aßmann, K. Kaspar, and P. Moormann (Munich: kopaed), 39–53.

Rüth, M., and Kaspar, K. (2017). The E-learning setting circle: first steps toward e-learning theory development. *Electron. J. Learn.* 15, 94–103.

Rüth, M., and Kaspar, K. (2020). Exergames in formal school teaching: a pre-post longitudinal field study on the effects of a dance game on motor learning, physical enjoyment, and learning motivation. *Entertain. Comput.* 35:100372. doi: 10.1016/j.entcom.2020.100372

Ryan, R. M., Rigby, C. S., and Przybylski, A. (2006). The motivational pull of video games: a self-determination theory approach. *Motiv. Emot.* 30, 344–360. doi: 10.1007/s11031-006-9051-8

Sandford, R., Ulicsak, M., Facer, K., and Rudd, T. (2006). *Teaching with Games: Using Commercial off-the-Shelf Computer Games in Formal Education*. Bristol, UK: Futurelab.

Saris, W., Revilla, M., Krosnick, J. A., and Shaeffer, E. M. (2010). Comparing questions with agree/disagree response options to questions with item-specific response options. *Surv. Res. Methods* 4, 61–79. doi: 10.18148/srm/2010.v4i1.2682

Schön, D. (1987). *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*. San Francisco, CA: Jossey-Bass.

Schrader, P. G., and McCreery, M. (2012). "Are all games the same?," in *Assessment in Game-Based Learning. Foundations, Innovations, and Perspectives*, eds D. Ifenthaler, D. Eseryel, and X. Ge (New York, NY: Springer), 11–28. doi: 10.1007/978-1-4614-3546-4_2

Schrader, P. G., Deniz, H., and Keilty, J. (2016). Breaking SPORE: building instructional value in science education using a commercial, off-the shelf game. *J. Learn. Teach. Digit. Age* 1, 63–73.

Schrier, K. (2014). "Using digital games to teach history and historical thinking," in *Learning, Education and Games. Volume One: Curricular and Design Considerations*, ed. K. Schrier (Canada: ETC Press), 73–91.

Seale, J. K., and Cann, A. J. (2000). Reflection on-line or off-line: the role of learning technologies in encouraging students to reflect. *Comput. Educ.* 34, 309–320. doi: 10.1016/S0360-1315(99)00052-4

Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: a model of instructional effectiveness. *Intl. J. Educ. Res.* 21, 141–157. doi: 10.1016/0883-0355(94)90029-9

Southerland, S. A., and Nadelson, L. S. (2012). "An intentional approach to teaching evolution," in *Evolution challenges: Integrating Research and Practice in Teaching and Learning About Evolution*, eds K. S. Rosengren, S. K. Brem, E. M. Evans, and G. M. Sinatra (New York, NY: Oxford University Press), 348–372. doi: 10.1093/acprof:oso/9780199730421.003.0015

Spielbar (2020a). *Spore*. Available online at: https://www.spielbar.de/spiele/145238/spore (accessed November 06, 2020).

Spielbar (2020b). *Valiant Hearts: The Great War*. Available online at: https://www.spielbar.de/spiele/147007/valiant-hearts-the-great-war (accessed November 06, 2020).

Squire, K. D. (2008). "Video-game literacy: a literacy of expertise," in *Handbook of Research on New Literacies*, eds J. Coiro, M. Knobel, C. Lankshear, and D. J. Leu (New York, NY: Routledge), 635–670.

Stephenson-Mittlböck, K. (2012). "Personality development through immersion into intermediate areas of digital role-playing games," in *Computer Games and New Media Cultures. A Handbook of Digital Games Studies*, eds J. Fromme and A. Unger (New York, NY: Springer), 233–248. doi: 10.1007/978-94-007-2777-9_15

Tavakol, M., and Dennick, R. (2011). Making sense of Cronbach's alpha. *Intl. J. Med. Educ.* 2:53. doi: 10.5116/ijme.4dfb.8dfd

Ubisoft Montpellier (2014). *Valiant Hearts: The Great War [video game]*. Paris: Ubisoft Entertainment S. A.

Ulrich, F., and Helms, N. H. (2017). Creating evaluation profiles for games designed to be fun: an interpretive framework for serious game mechanics. *Simul. Gaming* 48, 695–714. doi: 10.1177/1046878117709841

van Eck, R. (2009). "A guide to integrating COTS games into your classroom," in *Handbook of Research on Effective Electronic Gaming in Education*, ed. R. E. Ferdig (Hershey, PA: IGI Global), 179–199. doi: 10.4018/978-1-59904-808-6.ch011

Wang, X., and Goh, D. H.-L. (2017). Video game acceptance: a meta-analysis of the extended technology acceptance model. *Cyberpsychol. Behav. Soc. Netw.* 20, 662–671. doi: 10.1089/cyber.2017.0086

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., et al. (2009). Relationships between game attributes and learning outcomes: review and research proposals. *Simul. Gaming* 40, 217–266. doi: 10.1177/1046878108321866

Wouters, P., van Nimwegen, C., van Oostendorp, H., and van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *J. Educ. Psychol.* 105, 249–265. doi: 10.1037/a0031311

Wu, W.-H., Chiou, W.-B., Kao, H.-Y., Hu, C.-H. A., and Huang, S.-H. (2012). Re-exploring game-assisted learning research: the perspective of learning theoretical bases. *Comput. Educ.* 59, 1153–1161. doi: 10.1016/j.compedu.2012.05.003

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 03 February 2021
doi: 10.3389/fpsyg.2021.591708

# Teachers' Conceptions of Teaching Chinese Descriptive Composition With Interactive Spherical Video-Based Virtual Reality

*Mengyuan Chen*, Ching-Sing Chai, Morris Siu-Yung Jong and Michael Yi-Chao Jiang*

*Department of Curriculum and Instruction & Centre for Learning Sciences and Technologies, The Chinese University of Hong Kong, Shatin, China*

Phenomenographic research about teachers' conception of teaching has consistently revealed that teachers' conception of teaching influence their classroom practices, which in turn shape students' learning experiences. This paper reports teachers' conceptions of teaching with regards to the use of interactive spherical video-based virtual reality (ISV-VR) in Chinese descriptive composition writing. Twenty-one secondary teachers in Hong Kong involved in an ISV-VR-supported Chinese descriptive writing program participated in this phenomenographic study. Analyses of the semi-structured interviews establish seven conception categories that are specifically related to the use of ISV-VR for descriptive Chinese composition writing: (1) offering students more observational opportunities; (2) improving students' writing skills; (3) promoting students' learning participation and motivation; (4) shifting learning from teacher-centric to student-centric, (5) enhancing collaborative learning among students; (6) cultivating students' positive values and moral character, and (7) shaping students' self-identity as "writers." The concurrent and convenient access to the ISV-VR resources was for the teachers an enriched and supportive environment for them to cultivate students' writer identity. In addition, it was discovered that the structural relationships of the conceptions may be better organized along three axes of continuum: conception's orientation, teaching attention locus, and understanding of writing. These categories form a hierarchy from skill-oriented to community-oriented, and finally to identity-oriented conception. The findings may provide researchers and practitioners with novel insight into the teaching of composition writing in the contexts of L1 acquisition supported by virtual reality technology.

**Keywords: virtual reality, writing education, teacher conception, phenomenography, qualitative research method**

OPEN ACCESS

***Edited by:***
*Monica Pellerone,*
*Kore University of Enna, Italy*

***Reviewed by:***
*Joni Tzuchen Tang,*
*National Taiwan University of Science and Technology, Taiwan*
*Judit García-Martín,*
*University of Salamanca, Spain*

****Correspondence:***
*Mengyuan Chen*
*chenmy@link.cuhk.edu.hk*

***Specialty section:***
*This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology*

***Received:*** *05 August 2020*
***Accepted:*** *12 January 2021*
***Published:*** *03 February 2021*

***Citation:***
*Chen M, Chai C-S, Jong MS-Y and Jiang MY-C (2021) Teachers' Conceptions of Teaching Chinese Descriptive Composition With Interactive Spherical Video-Based Virtual Reality. Front. Psychol. 12:591708. doi: 10.3389/fpsyg.2021.591708*

## INTRODUCTION

Literacy has always been a key competency and one of the key intellectual infrastructural elements that students need to develop (Earnshaw, 2007; Bazerman, 2009; MacArthur et al., 2016). Among the four language skills that constitute literacy, writing is recognized not only as a means of communication but also a means to improve reading skill, comprehension, and critical thinking

(How and Larkin, 2013; Teng, 2016; Kong, 2018; Wang and Matsumura, 2019). Many language educators consider writing as the pinnacle of language education (Earnshaw, 2007; Bazerman, 2009; MacArthur et al., 2016).

In addition to the focus on writing, many contemporary language education curriculum emphasize the development of students' creativity, critical thinking, and positive values and attitudes (Choi, 2016; Macalister and Nation, 2019; Mickan and Wallace, 2019). Similarly, the Chinese language education (CLE) curriculum guide in Hong Kong also highlights the importance of improving students' communication, collaboration, self-learning skills, and culture learning supported by Information and Communication Technologies (ICT) (Curriculum Development Council of Hong Kong, 2017). These emphases can be broadly classified as 21st-century teaching and learning practices (Chai et al., 2019).

Despite the above emphases, the school's daily writing exercises are centered on knowledge presentation (take notes, write summaries, complete worksheets, etc.) (Teale and Yokota, 2000). Students may not prefer such pedagogical practices that fail to engage them cognitively and socially with technological supports (Chai et al., 2019). In addition, such practices may not address the perpetual problem of students' weakness in composition writing in general, and in descriptive writing in particular (Holliway, 2004; Akincilar, 2010; Carter, 2015). This study leverages on the interactive spherical video-based virtual reality (ISV-VR) technology to transform the writing pedagogy for descriptive writing. As teachers are the key agents in all education reform (Jong, 2019); and studies showed that teachers' conceptions of writing instruction play a role in how they implement writing programs within their classrooms (Lin, 2016; Kong, 2018; Wang and Matsumura, 2019); this study aims to explore how teachers conceive and hence enact descriptive writing pedagogy employing the affordances of ISV-VR technology. This study contributes to the current literature by describing a range of conceptions of teaching that may empower or limit teachers' use of virtual reality (VR) technology for the teaching of composition writing. Knowledge of teachers' conceptions may also provide useful insight for teacher educators and policymakers to enhance educational practices (Kelly and Beth, 2017).

# LITERATURE REVIEW

## Challenges in Teaching Descriptive Composition

Among various genres of composition, essays involving substantial descriptions are usually among the assignments that students must complete in composition classes (Holliway, 2004; Akincilar, 2010; Carter, 2015). According to Babayigit and Stainthorp (2011), descriptive writing is a comprehensive process that relies on the integration of different levels of processing skills such as working memory and transcription-related abilities. Successful descriptive composition affords readers a vivid environment to situate the themes of the essay through the writer's words. Experienced writers can create "pictures" with words (Holliway, 2004) but many novice writers fail to visualize the environment with words. Wilhelm (2008) stated that unless students "see" something in their mind, they will not be able to write about it. Researchers have reported that teaching secondary school students how to depict vivid and rich pictures for readers to fully understand the text is always a pedagogical challenge (e.g., Holliway, 2004; Akincilar, 2010; Babayigit and Stainthorp, 2011; Carter, 2015). This could be due to students' lack of vocabulary or writing skills, but it could also be due to pedagogical gaps in teaching.

In addition to the general writing skills including the choice of vocabulary, sentence structure, grammar and mechanics (Nair and Sanai, 2018; Chen et al., 2019), there are some descriptive techniques commonly used in Chinese landscape literature. The advocated descriptive writing techniques has been listed in the curriculum guide. According to Curriculum Guide for Chinese Language Education and Learning (Primary 1 to Secondary 6), the writing techniques of descriptive composition include the static descriptive method, which refers to a method of writing that creates vivid and concrete images of people and objects by describing them in a static state; and the dynamic description, which refers to a method of describing the moving state of people and objects to create concrete and lifelike images for the readers. In addition, the walking description method requires authors to observe scenes while walking, and describe the changing scenes; and the fixed-point description method is to choose an appropriate and fixed perspective to observe scenes, and to describe the object such as from top to bottom or left to right. Lastly, the sensory description method means that the author writes about people or scenes base on what he sees, hears, tastes, smells and touches (Curriculum Development Council of Hong Kong, 2017).

Despite the clear articulation of such techniques, a recent report from the high-stakes diploma of secondary education examination in Hong Kong shows that the lack of detailed descriptions, monotonous expressions, and unaffectionate descriptions are major problems with Hong Kong students' Chinese compositions (Hong Kong Examination and Assessment Authority, 2018). The report attributed these problems as the result of students failing to make observations in daily life and the lack of an in-depth understanding of their communities. Similar problems in students' performance in descriptive writing have been commonly reported (Holliway, 2004; Akincilar, 2010; Babayigit and Stainthorp, 2011; Carter, 2015; Huang et al., 2020). The challenges may lie in facilitating students' understanding about how description are connected with other parts of the essay, and the socio-historical significance of the described environment. Teaching students to observe their environment sensitively so as to build personal affective connections with the communities seems to be a pedagogical gap for CLE. ISV-VR is a new educational tool that supports learning and teaching activities in which "students' observation" is an important pedagogical component (e.g., Geng et al., 2019; Chien et al., 2020; Jong et al., 2020). It also greatly saves the cost and time of VR courseware development (Ozkeskin and Tunc, 2010; Mohiuddin et al., 2016; Repetto et al., 2018). Moreover, the technical skills of producing ISV-VR content is easy to master,

thus allowing school teachers to develop immersive materials according to their own teaching needs (Chen et al., 2019; Chien et al., 2020). Given the affordances of ISV-VR that may address the pedagogical challenge of fostering descriptive writing, the potential of ISV-VR supported Chinese writing was introduced to the teachers and relevant training for the technical skills; and resources with possible lesson activities was provided for the teachers to engage students in immersive appreciation of various places in Hong Kong. This study looks into the teachers' conceptions of designing and using ISV-VR to teach descriptive writing (for convenience, VRDW).

## Affordances of VR in Writing Instruction

Advancements in ICT have created the demand for innovative use of digital technology in the language education context (Lin, 2015; Kozar, 2016; Liu et al., 2017; Taghizadeh and Yourdshahi, 2019). VR technology is a potential tool to remedy the lack of opportunities for students to gather writing materials through detailed observation of phenomena. VR enables students to immerse themselves in a virtual environment comparable to the authentic world (Ozkeskin and Tunc, 2010; Lin and Lan, 2015; Mantziou et al., 2018; Repetto et al., 2018). It provides a safe and self-directed environment that enables students to observe places at length, with or without their teachers' guidance in the classroom.

Literature from immersive language learning approaches highlights the need for a rich and authentic environment to provide the socio-cultural context for learners to develop nuanced understandings of language in use (Cummins, 2000; Marsh et al., 2000; Swain and Lapkin, 2005). Research in embodied cognitive science also attests that situated immersive learning can stimulate students' learning motivation, enhance the learning experience, and encourage knowledge transfer (Marsh et al., 2000; Wilson, 2002; Swain and Lapkin, 2005; Steele et al., 2018). The content presented through interactive spherical video is more realistic than the 3D animation environment. The images employed in ISV-VR are filmed in the real world, and thus to present the complexity, diversity and details of the real world (Lin et al., 2019; Chang et al., 2020). Developers can also add a variety of human-computer interactions to the VR environment (Geng et al., 2019; Chien et al., 2020). The authors' previous research results of VRDW indicated that the VRDW courses could improve students' engagement and learning motivation in writing courses, and their writing performance regarding to the writing skills and themes of compositions (Chan et al., 2019; Chao et al., 2020). Findings from another study adopted spherical video-based VR (SVVR) to teach descriptive writing also showed that the SVVR writing approach could improve students' writing performance in terms of content and appearance as well as their creativity tendency and writing self-efficacy, while also reducing their cognitive load (Huang et al., 2020).

Despite the aforementioned affordances, Lin and Lan's (2015) review of the use of VR in language learning indicated that only 3 of the 29 studies they reviewed studied the teachers' perceptions or awareness of VR. These studies were focused on second language and other language skills instead of writing. A recent search (dated 26th November 2020) using the most comprehensive database (i.e., the Google Scholar) with the search term "composition writing" AND "Chinese language" AND " virtual reality" yielded eight returns, none of which investigate how teachers' conception or beliefs could influence how they conceive the role of VR for descriptive composition writing. It seems important for educational researcher to understand the conceptions that shape teachers use of the VR technology for writing curriculum because teachers' conception of technology has been shown to influence their practices (Alshahrani and Al-Shehri, 2012; Lameras et al., 2012; Lin, 2015; Kozar, 2016; Hsieh and Tsai, 2017). The current study thus sought to explore teachers' conceptions of VR usage in L1 instruction with a special focus on the teaching of writing.

## Teacher's Conceptions of Writing Instruction

"Conceptions" were defined by Pratt (1992) as specific meanings associated with phenomena that mediate the conceiver's reactions to situations involving these phenomena. In the context of educational research, conceptions of teaching and learning are studied employing the methods from phenomenography, which studies how people experience and understand phenomena in the world around us (Marton and Pong, 2005). Phenomenographic analysis methods can be used to study meaning, understanding, and conceptual variation. Rather than describing the real situation of the world, phenomenography focuses more on studying people's experience from the perspective of the participants. The core of phenomenographic study is the concept of "essence." For studying people's experiences of a certain aspect of reality, the "essence" here refers to the common and internal subjective meanings of that aspect (Paakkari et al., 2011).

In the context of education, conceptions of teaching have been developed as a distinctive field of research (Thompson, 1992). The teaching conception refers to a coherent system of knowledge and beliefs about teaching and related phenomena (Vermunt and Vermetten, 2004). Through phenomenographic research, conceptions of writing instruction have been described as ranging along the continuum of behaviorist and constructivist orientation in some literature (e.g., Lipson et al., 2000; McCarthey and Mkhize, 2013; Newell et al., 2014). Teachers who hold behaviorist-orientated conceptions are likely to emphasize the teaching of individual writing skills (Lipson et al., 2000). Teachers with constructivist-orientated conceptions tend to cultivate students' rhetorical style and voice (Lipson et al., 2000; McCarthey and Mkhize, 2013). Teachers' conceptions can also form a spectrum and can be placed into a hierarchy ranging from superficial to sophisticated, teacher-focused to student-focused, skill-oriented to meaning-oriented, and product-focused to process-focused (e.g., Boulton-Lewis et al., 2001; Kong, 2018; Wang and Matsumura, 2019).

These studies provide us with a deeper understanding of language teachers' conceptions of writing education. Recent studies indicate that language teachers conceive that using information and communication technology could foster students' thinking skills and promote learning beyond school (Alshahrani and Al-Shehri, 2012; Liu et al., 2017). However, none

of these studies focus on teachers' conceptions of using VR to teach first language writing. These studies advocate that more nuanced studies are needed in this research field. The current study thus proposes two research questions: (1) what are the teachers' conceptions of VRDW? (2) what are the variations in teachers' conceptions of VRDW?

# MATERIALS AND METHODS

In consideration of the students' performance gaps and the current affordances of VR technology, this study launched a VRDW learning program for six Hong Kong secondary schools. The program aimed to offer students near real-life observation opportunities through ISV-VR with added pedagogical features to foster closer connections between the students and various Hong Kong communities. Communities in this study are defined as the places and the people living there.

## School Background and Participants

Twenty-one grade 7–9 secondary Chinese teachers from a school in Hong Kong participated in the current study. The school is well equipped with technologies and the leadership is enthusiastic about integrating ICT into the curriculum. The Chinese language department participated in the VRDW program. Teachers designed the teaching plan and the VR learning content, and implemented the teaching plan. All Chinese teachers who participated in the VRDW were approached and they agreed to participate. They have experienced adopting the VRDW for at least one design and implementation cycle. The background information of the participants is reported in **Table 1**. Seven of them are male (33%). The average age of participants is 32.3 (SD = 8.7). Six of them have a Master's degree (29%) while the rest graduated as Bachelor (71%). Sixteen (76%) of them have a postgraduate diploma in education.

## Data Collection

The data in this study were collected through semi-structured interviews. The interviews were conversations around interview questions with the purpose of communicating and documenting the participant's experience (Moustakas, 1994). The questions for interviews were adapted from previous research which explored teachers' conceptions of adopting technology in teaching (Lameras et al., 2012; Hsieh and Tsai, 2017):

1. For what reason do you decide to use VR in teaching?
2. What is the role of VR in your teaching?
3. What are the differences between your VR writing class and your previous writing class?
4. Could you describe a VR writing lesson that you think is successful or unsuccessful?
5. Please give a formal definition of "VR supported descriptive writing learning."

During the interviews, the interviewers asked the participants to give examples or more detailed explanations when the interviewers were not clear with the participants' ideas (Creswell, 2003). The teachers were interviewed in Mandarin or Cantonese. The researchers also observed the classes before conducting the interviews in order to understand more fully and accurately about the teachers' discourse, and made appropriate subtle adjustments to the interview questions. The participants were interviewed individually after they finished the VR curriculum implementation, and each interview lasted between 45–60 min. After the interviews, the audio recordings were transcribed verbatim.

## Data Analysis

This study adopted the phenomenographic approach to analyze the teacher interview transcripts (Marton and Pong, 2005). Phenomenology mainly includes two aspects: looking for variations in teachers' conceptions and the structural relationship between these conceptions. In this study, we study teachers' conceptions of VRDW from the two aspects. After transcription, the data were managed in Nvivo (Version 12) to facilitate the creation and manipulation of codes.

The transcripts were analyzed iteratively. The researchers first read the transcripts several times to familiarize themselves with the participants' thoughts. Then used several keywords to mark the main meanings of teachers' views on using ISV-VR to teach Chinese writing. Next, the researchers compared and contrasted the similarities and differences between the main meanings to form the structural relationship of teachers' conceptions. Finally, based on the structural relationship, the researchers constructed a hierarchical structure of teachers' conceptions of VRDW. It includes describing categories in more details, delineating categories based on dimensions, generating outcome space which shows the logical connections among categories of description (Marton and Pong, 2005). For example, the following passage was marked to belong to category 1 as "breaking time and space limitations" to summarize its main meaning.

"The important thing is that VR can break the constraints of time and space. For example, VR allows us to watch the night scene during the day, which is the advantage of this technology."

For another two examples, we used "safe and effective way to replace field trips" to represent the main idea of the following two passages.

"In fact, using VR in writing class is more like an electronic version of field trips. Actually, I can't take my whole class to have field trips often for safety and time problems. Such opportunities are very rare for students. VR can increase their opportunities for observation. So basically I think this is a 'field trip in the classroom'."

"Using VR can save a lot of time and is more convenient than traditional field trip."

The following two passages was marked as "observe opportunities that hard to get in daily life" to summarize their main meanings.

"Students who use VR can look more carefully without causing misunderstanding and embarrassment, because many students will be embarrassed to observe in real life in crowded places, or observe other people."

"There are things and sights in ISV-VR that I haven't ever seen before. So, I think this is a good opportunity for students

**TABLE 1 |** Composition of the participants.

| | Gender | Age | Chinese teaching seniority (year) | E-learning experience (year) | Degree | PGDE* |
|---|---|---|---|---|---|---|
| T1 | Male | 41–45 | 11–20 | 11–20 | Bachelor | Y |
| T2 | Female | 20–25 | 1–3 | 1–3 | Bachelor | Y |
| T3 | Female | 31–35 | 4–10 | 1–3 | Bachelor | Y |
| T4 | Female | 41–45 | 11–20 | <1 | Master | Y |
| T5 | Female | 20–25 | <1 | <1 | Bachelor | N |
| T6 | Female | 20–25 | <1 | <1 | Bachelor | N |
| T7 | Female | 26–30 | 1–3 | 1–3 | Bachelor | Y |
| T8 | Female | 41–45 | 11–20 | <1 | Master | Y |
| T9 | Female | 36–40 | 11–20 | 4–10 | Bachelor | Y |
| T10 | Male | 26–30 | <1 | 1–3 | Bachelor | Y |
| T11 | Female | 20–25 | <1 | <1 | Bachelor | Y |
| T12 | Male | 36–40 | 11–20 | 11–20 | Master | Y |
| T13 | Female | 26–30 | 4–10 | 1–3 | Bachelor | Y |
| T14 | Male | 20–25 | <1 | <1 | Bachelor | N |
| T15 | Female | 20–25 | <1 | <1 | Bachelor | Y |
| T16 | Male | 26–30 | 1–3 | 1–3 | Master | N |
| T17 | Male | 26–30 | 1–3 | 1–3 | Bachelor | Y |
| T18 | Female | 41–45 | 11–20 | 1–3 | Bachelor | Y |
| T19 | Female | 46–50 | 11–20 | 1–3 | Master | Y |
| T20 | Female | 26–30 | 4–10 | 1–3 | Master | Y |
| T21 | Male | 41–45 | 11–20 | 4–10 | Bachelor | N |

**PGDE refers to Postgraduate Diploma in Education.*

to observe things what they don't have the chance to see in their daily life."

Categories of the teachers' conceptions of VRDW emerged through the comparison and contraction of these key meanings. These examples and some other passages all emphasized using ISV-VR as a way to let students have more chances to observe and experience the world out of the classroom. Thus, a category of "offering students more observational opportunities" was created.

Member checks were employed to improve the trustworthiness of this study. The researchers engaged the participants to review the researchers' interpretation of their experiences and to suggest adjustments to better capture their perspectives. After forming the initial categories, we sought feedback from an 18 years veteran Chinese writing teacher and an educational researcher who have rich research experience in educational phenomenology and Chinese writing education for communicative validity check (Åkerlind, 2005). The categories were then improved and finalized based on the feedback.

# RESULTS

Seven categories of teachers' conceptions of teaching descriptive writing with ISV-VR generated through the analysis. The results show that the participating teachers viewed VRDW as a hierarchy of conceptions that includes (1) offering students more observation opportunities, (2) improving students' writing skills, (3) promoting students' learning participation and motivation, (4) shifting learning from teacher-centric to student-centric, (5) enhancing collaborative learning among students, (6) cultivating students' positive values and moral character, and (7) shaping students' self-identity as "writers." The structural relationship of teachers' conceptions are illustrated through three key aspects: (1) conceptions' orientation, (2) teaching attention locus, and (3) understanding of writing. In the following sections, we first describe the seven categories in detail, followed by their structural relationship and hierarchy, and finally, present the distribution of the categories.

## Description of Categories

### Offering Students More Observational Opportunities

In this category, ISV-VR is described as a tool that affords the means for students to have more observational opportunities. Teachers view using ISV-VR as a safe and efficient way to construct a convenient and comfortable observation environment to replace the "field trip." It allows students, especially those who are unable to participate in field trips, to observe the landscape more carefully, and enabling students to go where they cannot in real life.

T1: In fact, using ISV-VR in writing class is more like an electronic version of "field trip," and it's more efficient and safe. Actually, I can't take a whole class of students to have field trips often. Such opportunities are very rare for students. While ISV-VR can offer students more opportunities to observe the world. So basically I think this is a "field trip in the classroom."

T16: There are things and sights in ISV-VR that I haven't ever seen before. So, I think this is a good opportunity for

students to observe things what they don't have the chance to see in their daily life.

T6: Using VR can save a lot of time and is more convenient than traditional field trip.

T4: Students who use VR can look more carefully without causing misunderstanding and embarrassment, because many students will be embarrassed to observe in real life in crowded places, or observe other people.

In this category, teachers are inclined to substitute the physical field trip with ISV-VR experiences. The field trip is a well-established pedagogical means for immersive learning. Nonetheless, it is a difficult teaching activity to implement due to logistical and safety considerations. Teachers in this category focus on the advantages of VR technology in circumventing the time and space limitations and in providing a safe and efficient learning environment. This conception is grounded in the primary technological affordances of VR. Description about in-depth processing of the meaning of what students observed is lacking.

### Improving Students' Writing Skills

In the second category, teachers believe the purpose of using ISV-VR is to improve students' writing skills. ISV-VR served as an information-rich environment that can stimulate detailed observation. Teachers believe that by observing in ISV-VR, students can acquire more writing materials and practice various descriptive skills in writing. Besides, the rich writing materials provided by ISV-VR can stimulate students to engage in more detailed descriptions, so that the content of their composition is more detailed and vivid.

T7: I think the main use of ISV-VR is to allow students to find more details that they usually don't notice or they can't imagine. They can find these in ISV-VR, and make these things their writing materials so that their composition can be vivid and concrete. Besides, they can also have some new subjects to write about.

T10: When you read the students' work, you will notice that the pictures they have in their minds are not clear, so their description is monotonous. But after using ISV-VR and with my encouragement and guidance in the classroom, students try to use more descriptive techniques in their compositions, so the content of the composition is much more specific and authentic. I think students' writing is actually better than last semester.

T21: I notice that in the VR writing classroom, some students asked me, "Teacher, I see the sky is very vey blue, how should I write?" I told him that he could use words like 'azure blue' or 'bright blue' instead of 'very blue.' I think VR can help students to describe more vivid and realistic pictures by using rich words.

T8: One writing skill students need to learn is the "walking description method," and another writing skill is the "fixed-point description method." There are videos and static images in VR, students can practice using these two writing skills.

In this category, teachers begin to link ISV-VR observation with the content of the writing, and their focus is on the students' writing with more things and with more varied techniques. The teachers believe that writing is a process from input to output while observing in VR can increase students' input to writing, so as to improve their writing performance.

### Promoting Students' Learning Participation and Motivation

In the third category, ISV-VR is conceptualized as a vehicle to promote students' learning participation and motivation, and a tool to change students' attitudes toward writing. Teachers believe that the freshness brought by this cutting-edge technology attracts students to participate in learning activities more actively. Teachers also get more responses from their students, and the interaction between teachers and students is enhanced so that the learning effect is reportedly improved. Moreover, the benefits brought by ISV-VR technology to students' writing will improve students' confidence in writing, thus further enhance their motivation to study and write.

T11: When students observed a certain location in ISV-VR, they sometimes exclaimed: "Wow, it's so beautiful, I want to go there on Saturday!" They talked about their feelings, because they were touched by the scenery, and I'm pleased about that. I think it's so great to stimulate their interest in observation and writing.

T18: The writing ability of my students is relatively weak. But after using ISV-VR, some of them became serious about writing. If they haven't finished watching ISV-VR in school, I would give them the website links of the panorama pictures, and they watched the pictures at home. I feel that students are willing to spend time and effort to prepare for their writing.

T13: When students observing in the VR, I would walk around the class, and some students would ask me a lot of questions about scenes they observed in VR, and they are more involved in the classroom.

T5: In fact, I am very surprised by the effect of the class, because my students are usually very quiet in classroom, there are 20 girls and 2 boys in my class. In the past, I kept asking them questions, but they didn't communicate with me very well. However, in the VR classes, I hope they can answer the questions when I ask them, and they all did as I expected.

In this category, teachers generally focus on the students' learning attitude. Teachers say they are pleased that students' attitudes toward writing have changed. They are more concerned with the affective dimension of learning to write rather than the technical means-end aspect of writing. The students' learning attitude and motivation are foregrounded rather than the stipulated learning outcomes to be achieved.

### Shifting Learning From Teacher-Centric to Student-Centric

This category encapsulates the teachers' view of the utilization of VRDW as shifting learning from teacher-centric to student-centric. In traditional writing classrooms, teachers' role is to teach the selected articles for reading comprehension and explicate the embedded writing techniques for students to emulate. The VRDW creates a shared immersive environment where students can pursue their own perspectives while the teachers facilitated by providing personalized guidance to stimulate students to think deeper. From this perspective, the ISV-VR platform provides students with opportunities to form personalized connections with the places and people they observed. The difference between this category and other categories is that teachers are actively aware of their roles as understanding facilitators. This includes strategies for questioning and discussion to develop students' personal thinking and understanding. This change has also increased the interaction between teachers and students.

T16: My (teaching) habit is that I don't give students too much guidance first. I ask them to observe in the ISV-VR and then write down what they are interested in. after that, I give them some ideas, I give each of them as different advice as possible.

T14: In my ISV-VR class, my role is still a teacher, but what I teach is not content anymore. In the past, when I was teaching, I mainly started with my point of view, but now students don't use my point of view anymore. I mainly lead them to associate what they observed in VR with their life experience, to imagine and reflect a little bit.

T17: When observing in VR, students can freely choose the viewpoint from 360 degrees, pause at any time to observe the details of the scene they are interested in, or shuttle between the scenes. Students decide the scene to observe, the order of observation, and the scene to be described.

T9: I find that in the VR class, there is no need for me to give them a topic or a scene to describe, because students decide what to write on their own, then I help them with the details of writing and guide then to think deeper.

The concept of teachers in this category shows that teachers are no longer the only source of knowledge in the classroom. They realize that the protagonist in the ISV-VR classroom has become students, and the teacher's role has changed from lecturing to guiding. This shift is essential to open up space for students to assume the responsibility of being the author of their composition. Teaching is to promote students' learning and understanding through cooperation with them. The cooperation includes helping, motivating, asking questions, and discussing with students. With this in mind, the teachers explored as much as possible the students' various understandings and ideas and expand these differences. VR triggered diverse responses from students and consequently more personalized responses from the teachers. This contributes to a more student-centric learning rather than the more monologic teaching that is based solely on reading passages.

### Enhancing Collaborative Learning Among Students

The differences in what the students experienced in the ISV-VR environment, both as individuals and as groups, can be shared and compared in classroom discussion sessions and this has provided diverse perspectives for rich discussions and learning to occur. Based on the differences, some teachers conceived VRDW as a pedagogical environment to promote collaborative learning. This is achieved through having a shared 3D environment that affords different and individual angles of observation that triggered different experiences. Collaboration in VRDW includes learning from each other's perspectives via listening and commenting, which led to more refined ideas.

T15: I want students to share what they had seen in ISV-VR, and to work together on the worksheet. Then appraise and select the best work from their group. In this process, students finally came up with very good answers, and they can learn a lot from each other.

T13: I want the students to learn how to observe from the perspectives of multi-angle and multi-sensory. So I divided four groups and asked each group to focus on a specific sensory perspective to observe. I found that if the students focused on a certain angle, they could observe more deeply, and that also create points for peer discussion. At last, when the groups reported their results, every group can learn from each other.

T2: I rarely organized collaborative learning in traditional writing classes in the past. However, VR offers students a learning environment where they can observe collaboratively. Students can choose their point of view 360 degree in the VR scene, or they can pause to observe carefully. So I found that the details observed by students are very different, so I encouraged students to observe collaboratively to enrich their findings in VR.

T18: Using VR can enhance the sharing of experiences among students. If some students have had some personal experience, other students will listen to this student's sharing after watching VR. For example, a classmate shared his experience of shopping at a street stall: "I once bought a vacuum cleaner at a similar street stall. It was originally sold for 120 yuan. I said it had no power cord and asked the shopkeeper to sell it for 20 yuan. The shopkeeper even agreed!" After hearing this anecdote, other students felt more real about the chaotic of the street market.

In this category, in addition to being facilitators, the teachers use ISV-VR to build collaborative learning environments. The teachers intentionally select angles of observation and create points for discussion. They required the students to work in groups to summarize the observational findings and finished the learning tasks collectively. For teachers in this category, the ISV-VR is viewed as a means to engender social collaboration for deeper meaning-making afforded by diverse observational experiences.

### Cultivating Students' Positive Values and Moral Character

The sixth category of cultivating students' positive values and moral character refers to teachers leading students to fully and profoundly understand the cultural and educational connotation of literary works. This conception drives the teachers to focus on building emphatic understanding between reading and writing through the embodied experience in ISV-VR that contextualizes the guided reading of the selected texts. Teachers chose these passages based on the educational connotations and the cultural content they contained, including appeals to cultural significance, positive values, and moral character through reflections on the real world. More than literary achievement and skills development, teachers pay more attention to the cultivation of students' empathy to understand the emotions of other authors and their perception of cultural connotations, so as to increase students' relatedness to the community.

The pedagogical activities help students to have a deeper understanding of the educational connotations and the cultural content that the authors delivered. In this process, the teachers employ the VRDW to facilitate students' associative thinking and their ability to think from multiple perspectives, especially from the perspectives of the authors. The selected authors in turn provide the experiential framing that is culturally valued. The VR materials invoke experiential references for students to interact and understand the authors' perspectives.

T17: It is mainly for students to understand that people have different perspectives of life and different understandings of things, and show students how other authors write about Shatin (an area of Hong Kong), so as to cultivate students' thinking of understanding the world from multiple perspectives. It is difficult for students to get rid of their preconceived ideas, but the combination of VR and reading text will make it easier for students to develop empathy and understand the connotation of education and humanistic emotions in literary works.

T12: Finally, in fact, the most important thing is cultivating students' positive values and attitudes, and to understand the deep meaning of the reading passage, that is, the author's perspective expressed through the text. So the first step is for students to understand other author's ideas and how they view the author's ideas. It is more about emotional and moral education.

T19: For example, this is written by a boy in my class. After watching VR, he imagines that he is a cat in an old neighborhood. He describes the changes in the neighborhood where he lives in the tone of a cat, and laments the fast-paced changes in the city and the strong contrast between the slow-paced life in the old neighborhood.

T12: For example, in the article "Father and Son" describes the Wanyi Reservoir. It writes that the author went to the Wanyi Reservoir with his father when he was a child, and then he took his son with him to the same place when he grown up. Authors expressed his affections of family, through text analysis and observation of Wangi Reservoir in VR, students can experience the feelings described by the author in the text and thus to cultivate students' gratitude to their loved ones.

In this category, the texts and observations in ISV-VR are described as sources of knowledge or ideas, and students became active meaning creators by citing authoritative sources with social value. In selecting passages for reading comprehension, teachers chose articles that reflect deep and critical issues that shape the author's choice of what to describe the places based on the themes established in each unit (e.g., the disjoint between new and old communities, the gap between rich and poor, the diversity of the city life). Students think more deeply under the guidance of the teacher and produce new knowledge. In this process, the students' thinking is also developed and improved.

### Shaping Students' Self-Identity as "Writers"

Based on the previous categories, in this category, teachers led students to compare and contrast the literary works with reference to the scenes that students observed in ISV-VR, so as to enable students to further develop their unique views of the landscapes, people and communities. In other words, detailed observation of the scenes and studying how the scenes are depicted by established writers create deep understanding about how objective scenes are subjectively transformed into meaningful sociocultural artifacts that can engage readers in a unique and emotive time-space created by the authors. The teachers aim to develop students' voices in their writing through reading and they led the students to think about their own relationship with others and communities. Students then write about their unique experiences, opinions, and affections with reference to the viewed VR landscapes. This conception seems closely reflects Norton's (1997) conception of language and identity depicted as "how students understand their relationship to the world" (Norton, 1997, p. 410). It is arguably the core of being a writer and the teachers intend to facilitate students to discover possible themes and to express how these ideas from both the selected passages and VR materials are connected to themselves and their understanding of the world.

T13: So I think what's going on in ISV-VR is that students are learning how to observe like a writer, and to associate with what's going on. Yes, the connection between observation and association, which is equally important, and it's actually the hardest thing to teach. But only by associating with their own life experiences can students form their own unique perspectives.

T17: Students may have different ideas about a community, there can be criticism, also some personal outpouring. So there's a lot of thinking. Like a lot of ideas and associations that come out of it. For example, what are people's impressions of this area in the past? What should be my relationship with the community? These questions could inspire them to think about the relationship between people and communities and reinforce their writer identity.

T13: Before using VR, there was no way to teach students how to observe like a writer. So how to teach students to see from a writer's point of view? It need teachers to show them how to observe, for example, a wall here is mottled, why it has become so, it turns out that these details can reflect the long history of ancient buildings. When students know how to observe like a writer, they can have new findings in daily life and their own viewpoints.

T12: Because some of the scenes in VR are the places where students live, students find that they never observed the places they passed by everyday. In fact, the ultimate goal is to enable students to take the initiative to observe in real life and realize that the whole world is closely related to them. For example, no matter what the teacher says or the text describes how the rocks on the beach were formed through the vicissitudes of time, the students have no experience about it. Only when they see the traces left by the change of the stone can they feel the changes in the vicissitudes of time. And then they want to visit on the spot, I think this is the most meaningful, because these feelings are real.

In this category, teachers realize that descriptive writing has a function of establishing personal sociocultural bonds between the student writers and the places, which transform descriptions into meaning making. Excellence in descriptive writing necessarily reflects meaningful space, time and communal relationships (Hyland, 2002). Students study the Hong Kong landscape literature and observe the corresponding landscapes and living environment via ISV-VR. The sensitization of the context is reinforced through the reading of related passages and receiving teachers' guidance. The concurrent and convenient access to the scene was for the teachers an enriched and supportive environment for them to cultivate students' writer identity. The teachers thus nurture students' connection to local culture and community. Teachers in this category emphasizes students' role as writer and facilitate students' development of individual meaning making, which gives rise to students' sense of identity as writers.

## Relationships Among Categories

The relationship among categories and how the categories are arranged are reflected in a defined structure (Marton and Pong, 2005). By comparing and contrasting the similarities and differences between categories, three key aspects of the relationships emerged from the data analysis: (1) conception's orientation, (2) teaching attention locus, and (3) understanding of writing. As the category level rises, the complexity and inclusiveness of experiencing VRDW will also increase. Åkerlind et al. (2005) pointed out that the hierarchical structure is not based on value judgments, but represents that certain categories include other categories and are therefore more complicated than others. **Table 2** shows the key aspects that illustrate the relationship between the categories defined in this study.

**TABLE 2** | Key aspects of experiencing VRDW.

| Key aspects | Category 1 | Category 2 | Category 3 | Category 4 | Category 5 | Category 6 | Category 7 |
|---|---|---|---|---|---|---|---|
| | Offering students more observational opportunities | Improving students' writing skills | Promoting students' learning participation and motivation | Transforming learning from teacher-centric to student-centric | Enhancing collaborative learning among students | Cultivating students' positive values and moral character | Shaping students' self-identity as "writers" |
| Conception's orientation | Skill-oriented | Skill-oriented | Community-oriented | Community-oriented | Community-oriented | Identity-oriented | Identity-oriented |
| Teaching attention locus | Content | Content | Meaning/concept | Meaning/concept | Meaning/concept | Culture | Culture |
| Understanding of writing | Cognitive activity | Cognitive activity | Social activity | Social activity | Social activity | Cultural activity | Cultural activity |

The "conception's orientation" refers to the teachers' underlying understanding of VRDW. The first two categories mark the teachers' conceptions as "skill-oriented." The teachers' purpose for the application of ISV-VR in writing instruction is to improve students' writing skills, implicitly referencing performances measured by examination. The orientation of conceptions in the 3, 4, and 5 categories shifts to "community." Teachers believe that ISV-VR technology can provide a social learning environment for students. This shift positioned the VRDW to include socialization as part of the process of writing, which is arguably a richer conception of writing. The orientation of conception ultimately shifts to creating a socio-cultural "writer identity" in the latter two categories, as teachers believe the use of ISV-VR technology can help students develop their voice in writing. This shift emphasizes the subjectivity of writing through personal meaning making, which builds a unique writer identity.

As to the second aspect, teaching attention locus, the teachers' attend to three areas of teaching that is hierarchical. In categories 1 and 2, teachers instruct their students to focus on quantity of the "content," which enables students to make observations and enrich their composition by using more descriptive writing skills. In categories 3, 4, and 5, teachers focus on promoting students' engagement and collaboration that ISV-VR provides. The teaching attention was on creating 21st century learning experiences (Chai et al., 2019). In the category 6 and 7, teachers focus on the ultimate aim of writing curriculum: cultivating writer who can construct unique and socioculturally meaningful essays. They seek to cultivate the students' understanding and arouse their interest in local culture and care for their communities. This works by allowing them to observe local representative landscapes and read literary works depicting these landscapes.

The third aspect is the teachers' understanding of writing. In the 1 and 2 categories of teachers' conception of teaching descriptive writing with ISV-VR, teachers regard descriptive writing as a cognitive activity. Descriptive writing is viewed as product-oriented or text-focused. At this point, they consider ISV-VR as a tool to enrich the content of students' compositions. In the 3, 4, and 5 categories, teachers consider descriptive writing to be a social activity. Apart from individual learning, teachers emphasized more on the interactional aspect of writing classes. In the 6 and 7 categories, teachers consider descriptive writing to be a cultural activity. Students learned literary works with cultural significance, enhanced their understanding of local culture through observations in VR, and expressed their ideas of local culture in writing. Students can acknowledge involvement in a particular community through descriptive writing (Hyland, 2002). Moving across the categories in this hierarchy also marks a shift from focusing on teachers to focusing on students, and from emphasizing products to emphasizing processes. Emphasizing products means that teachers' learning focuses on analysis of students' performance, execution, success or final results, which can be measured by indicators such as the use of rhetoric and descriptive writing skills. Emphasizing products means that teachers' learning focuses on the orchestration, dynamics and deployment of the students' mental process, and other variables that modulate the acquisition of observation abilities (García-Martín and García-Sánchez, 2018).

## Distribution of Categories

Teachers' individual experiences may distribute across categories. Therefore, the frequencies of each teacher's conceptions are tabulated in **Table 3** to gain a better understanding of how teachers' conceptions of VRDW spread across the different categories. The frequency was measured according to how many times the participants stated a certain idea. To see the characteristics of each teacher's conceptions more intuitively, the frequencies were further converted into symbols in **Table 4**. If a teacher's conception is identified to a category, then this category will be marked with a "√," and the most frequently mentioned categories were marked with a "▲." For example, in **Table 3**, the VRDW conceptions for T5 are distributed in categories 1 to 6, where category 2 is the most frequently mentioned category (18 counts). Therefore, categories 1 to 6 of T5 are marked with "√," and Category 2 is marked with "▲."

As can be seen in **Table 4**, improve students' writing skills (Category 2) appeared in every teacher's conceptions ($n$ = 21) follow by promoting students learning participation and motivation (Category 3, $n$ = 20), the frequency of shaping students' self-identity as "writers" (Category 7) is the least ($n$ = 5). Besides, the highest frequency of the teachers' mentioned category was improving students' writing skills (Category 2, $n$ = 9) followed by transforming learning from teacher-centric to student-centric (Category 4, $n$ = 6). The majority of teachers' most sophisticated conceptions of VRDW is cultivating students' positive values and moral character (Category 6, $n$ = 8).

# DISCUSSION

This study aims to explore the variations of teachers' teaching conceptions in the context of VRDW. The results offer an understanding of how language teachers assimilate and accommodate new technology in writing education by revealing specifically different ways that teachers experienced VRDW. Teachers' experience of adopting innovation in education is a concern among educators (Alshahrani and Al-Shehri, 2012; Newton and Beverton, 2012; Hsieh and Tsai, 2017; Liu et al., 2017). In this study, seven conception categories described along with teacher interview excerpts formed a spectrum for understanding the teachers' conceptions. The structural relationship and hierarchy further explained the internal connection between the conceptions. Finally, the distribution of the categories presented a better understanding of the spread of the teachers' individual experiences of VRDW.

In the context of language learning, VR technology has been advocated for a long time. However, it has been pointed out that few studies explored from the perspective of teachers (Lin and Lan, 2015). Studies also revealed limitations in teachers' conceptions of ICT-supported language learning (Alshahrani and Al-Shehri, 2012; Newton and Beverton, 2012; Liu et al., 2017). It advocates that we need more nuanced studies on teachers' conceptions of involving VR technology in writing

**TABLE 3 |** Frequency distribution of teachers' conceptions of VRDW.

| | Category 1 | Category 2 | Category 3 | Category 4 | Category 5 | Category 6 | Category 7 |
|---|---|---|---|---|---|---|---|
| T1 | 2 | 4 | 4 | 6 | 0 | 1 | 0 |
| T2 | 1 | 3 | 4 | 1 | 2 | 5 | 1 |
| T3 | 6 | 2 | 2 | 3 | 2 | 2 | 1 |
| T4 | 2 | 2 | 7 | 4 | 0 | 0 | 0 |
| T5 | 2 | 1 | 6 | 2 | 3 | 1 | 0 |
| T6 | 2 | 3 | 7 | 1 | 3 | 0 | 0 |
| T7 | 0 | 4 | 8 | 12 | 0 | 0 | 0 |
| T8 | 0 | 4 | 10 | 0 | 0 | 0 | 0 |
| T9 | 4 | 6 | 2 | 0 | 0 | 0 | 0 |
| T10 | 1 | 4 | 8 | 3 | 1 | 1 | 0 |
| T11 | 0 | 4 | 12 | 0 | 1 | 0 | 0 |
| T12 | 3 | 0 | 5 | 2 | 0 | 20 | 17 |
| T13 | 2 | 11 | 7 | 6 | 5 | 12 | 3 |
| T14 | 1 | 1 | 7 | 3 | 2 | 1 | 0 |
| T15 | 1 | 4 | 6 | 3 | 2 | 1 | 0 |
| T16 | 5 | 2 | 6 | 8 | 0 | 5 | 0 |
| T17 | 0 | 1 | 1 | 7 | 5 | 3 | 2 |
| T18 | 0 | 3 | 6 | 9 | 0 | 0 | 0 |
| T19 | 2 | 3 | 6 | 1 | 2 | 8 | 0 |
| T20 | 3 | 3 | 4 | 7 | 0 | 1 | 0 |
| T21 | 1 | 5 | 8 | 0 | 3 | 0 | 0 |

**TABLE 4 |** Frequency distribution of teachers' conceptions of VRDW.

| | Category 1 | Category 2 | Category 3 | Category 4 | Category 5 | Category 6 | Category 7 |
|---|---|---|---|---|---|---|---|
| T1 | √ | √ | √ | | √▲ | √ | |
| T2 | √ | √ | √ | √ | √ | √▲ | √ |
| T3 | √▲ | √ | √ | √ | √ | √ | √ |
| T4 | √ | √▲ | √ | | √ | | |
| T5 | √ | √▲ | √ | √ | √ | √ | |
| T6 | √ | √▲ | √ | √ | √ | | |
| T7 | | √ | √ | | √▲ | | |
| T8 | | √▲ | √ | | | | |
| T9 | √ | √ | √▲ | | | | |
| T10 | √ | √▲ | √ | √ | √ | √ | |
| T11 | | √▲ | √ | √ | | | |
| T12 | √ | √ | | | √ | √▲ | √ |
| T13 | √ | √ | √ | √ | √ | √▲ | √ |
| T14 | √ | √▲ | √ | √ | √ | √ | |
| T15 | √ | √▲ | √ | √ | √ | √ | |
| T16 | √ | √ | √ | | √▲ | √ | |
| T17 | | √ | √ | √ | √▲ | √ | √ |
| T18 | | √ | √ | | √▲ | | |
| T19 | √ | √ | √ | √ | √ | √▲ | |
| T20 | √ | √ | √ | | √▲ | √ | |
| T21 | √ | √▲ | √ | √ | | | |
| | √ = 16 | √ = 21 | √ = 20 | √ = 12 | √ = 17 | √ = 13 | √ = 5 |
| | ▲ = 1 | ▲ = 9 | ▲ = 1 | ▲ = 0 | ▲ = 6 | ▲ = 4 | ▲ = 0 |

education. Studying teachers' conceptions of VRDW enables educators to gain a sense of how teachers think about teaching goals, activities, strategies, tasks, and processes in this context. In addition, examining teachers' conceptions can also provide insight for teacher trainers and policymakers to promote teachers' professional development. The results of this study may provide insight for better integrating new technology (especially ISV-VR) into the school curriculum.

Consistent with previous research, the more commonly reported conceptions are similar to the surface conceptions found in this study. In category 1 and category 2, teachers considered ISV-VR as enriching writing material and improving students' writing skills. From the aspect of the conception's orientation, the surface conceptions found in the current research were skill-oriented. They are similar to the orientation of surface conceptions that Wang and Matsumura (2019) found: writing was the application of skills and strategies. The teaching aim was to enable students to demonstrate their ability to understand and apply specific reading and writing skills. In terms of teaching attention locus, the surface conceptions which were found in the current research focus on the content rather than the meaning and concept. Students' observation in ISV-VR was to enrich their composition. Similarly, Wang and Matsumura (2019) pointed out that literary works were regarded as objects of practice writing, rather than sources of knowledge and ideas. From the aspect of teachers' understanding of writing, teachers who held the surface conceptions in the current research regarded writing as a cognitive activity. In the same way, Boulton-Lewis et al. (2001) founded that the superficial conceptions of teaching were considered to impart information or skills, and teachers and content were the focus. These surface conceptions focus more on the technology and means-end aspects of writing, taking ISV-VR acting as a tool to achieve a better writing performance. Such conception may be limiting in that it may result in students viewing and experiencing the writing curriculum as merely a school work to be completed.

On the other hand, the more sophisticated conception categories focus more on deepening students' understanding of the world and improving students' thinking. From the perspective of the conception's orientation, the more sophisticated conceptions summarized in the current research were community-oriented. In category 3, the teacher believed that the application of ISV-VR can promote students' learning participation and motivation. Similarly, Lin (2015) pointed out that teachers were applying ICT to let students become "more focused" and "more engaged" in the writing class. In category 4, teachers regarded VRDW as transforming learning from teacher-centric to student-centric. It is also consistent with the sophisticated conception that Wang and Matsumura (2019) noted: teachers tend to support students in exploring writing topics rather than direct instruction. It also shares the same orientation with the sophisticated conception of "transformation" that Boulton-Lewis et al. (2001) summarized: the teacher organized the situation to provide the stimulus for students to take action, and then faded into the background. In terms of the teachers' understanding of writing, in sophisticated conception categories, teachers considered writing as a social activity. The fifth category is enhancing collaborative learning among students. Lin (2015) also found that teachers considered the application of ICT as engaging students in discussion and increasing interactions between teachers and their students. From the perspective of the teaching attention locus, in the more sophisticated conceptions in this study, teachers focused more on meaning and concept. It is similar to the sophisticated conception that Wang and Matsumura (2019) found: reading and writing were integrated through emphasizing a deeper understanding of the text. This process led the students to a "learning through thinking" model (Lin, 2015) which rely much on social negotiation that makes the thinking visible. Teachers possessing such conceptions can create enriching 21st century learning experiences with technology but it may at times miss the key purpose of developing writers.

Finally, we found that some teachers with strong sophisticated conceptions would connect writing with culture, aiming to develop students' writer identities. This finding sheds new light on our understanding of teachers' conceptions in writing education. Teachers who held the last two categories of conception believed that VRDW was an approach to reinforce students' self-identities as "writers." Although developing students' writer identities is a great challenge (Freedman and Ball, 2004; Vasudevan et al., 2010), the teachers apparently exploit the immersive situational learning environment to stimulate students' inner voice supplemented by literary works from established writers and teacher's guidance. Teachers with sophisticated conceptions would guide students in developing empathy about other people's emotions and help inspire students' inner voices. Therefore, teachers' sophisticated conceptions shape how they configure the learning environment for students' writing development. These teachers would work together with students to establish their personal meaning by stimulating students' unique opinions and feelings about the communities and landscapes when students experiencing communities and landscapes closely in immersive VR learning environments. Bower and Jong (2020) reckoned that much of the success of IVR learning depends on the pedagogical approaches that occur outside the IVR environment. Teachers who hold sophisticated conception leverage the pedagogical affordances of VR by incorporating guided reading of literature and embodied experience in the VR learning environment. The coordination between text and sensory experience allows teachers with this conception to pose questions and offer guidance about text interpretation with multi-dimensional observation. This help students to develop refined understanding about the aspects of words, rhetoric, structure, and narration. These teachers also employ traditional means such as using worksheets for students to record the details their findings of close reading. The teachers encouraged students to think from multiple perspectives and deepen their understanding and feelings of communities. Our findings appear to be well supported by the arguments held by writing educators. From the research on the literature of writing education, we find the basis that supports teacher's conception of applying VR to develop student's writer identity. As suggested by Vasudevan et al. (2010), students could draw on their cultural resources to establish "literate identities" in the classroom by employing multimodal digital tools for composing. Identity creation involves a self-conscious selection from the varied cultural resources at one's disposal. This characterization is also consistent with Bakhtin's theory on how people understand and incorporate the voices of others. Namely, different cultural resources offer different roles (Freedman and Ball, 2004). From the aspect of the conception's orientation, this conception is identity-oriented. Teachers who held

this conception focused on the cultural context and regarded writing as a cultural activity. This result not only shows the high level of teachers' conceptions of writing and VRDW but also increases our knowledge of helping students constructing their writer identities. The proposed structural relationships of the categories provide insights into the nature and variations of teachers' conceptions of VRDW. Meanwhile, they also help to have a more complete understanding of teachers' writing conceptions. From the distribution of individual experiences of teachers, it was found that only few teachers had such sophisticated conceptions, and most teachers didn't realize the role of VR in cultivating students' cultural consciousness and writer identities. This may lie in the fact that the teachers' conceptions of VRDW are closely related to their conceptions of writing, and the teachers' conceptions of writing may not reach higher levels. Our study suggests that in future studies, researchers may need to develop teachers' sophisticated conception in the professional development program to better promote the integration of technology into school curriculum (Hsieh and Tsai, 2017).

We are aware that our research may have two limitations. First, the leaders of the participated school in this study are enthusiastic about integrating ICT into the curriculum, and they have rich attainments in writing education. This may affect teachers' conceptions toward VRDW. The conceptions obtained in this study may be more positive than that of other schools. Therefore, more schools need to be involved in the follow-up research to examine the variations in teachers' conceptions of VRDW in a different school environment. Second, the results of this study mainly come from the interview data of teachers. Therefore, we need to pay attention that there may be a certain inconsistency between the teachers' discourse and their teaching practice. However, we do not try to generalize the results of this study. Instead, readers of this study need to decide to what extent they transfer the results of this article to their own context. Also, the teachers' conceptions obtained in this study can be applied as the basis for formulating the assessment criteria of teachers' attitudes toward ISV-VR, or even other technology-supported writing learning, or as a framework for formulating the curriculum of language teacher education.

## CONCLUSION

The current study investigated variations of Chinese teachers' conceptions of VRDW. It revealed seven conception categories, the distribution of the categories in teachers' conceptions, and the hierarchical relationships of the categories. The findings of this study have some important implications for teacher educators, policymakers, and school administrators in schools, especially for those who are concerned with VR-supported language learning. This research not only corroborates previous results about teachers' conceptions of writing education but also expands our understanding of teachers' experience of applying VR technology in L1 writing. From the perspective of teacher education, the results of this study can provide references for language teachers' professional development to promote VR-based immersion language learning in the school curriculum. If teacher educators want to make VR devices play more of their potential in writing education, it is necessary to cultivate more sophisticated conceptions of VRDW among teachers. Applying the results of this study, teachers can reflect on their teaching practice of adopting VR technology to promote students' learning of writing in their native language, to expand and enhance their understanding of immersive writing learning supported by VR.

It is also necessary to realize that teaching practice may vary depending on the teacher's educational background and environment. Therefore, future research can explore the conceptions of VRDW held by teachers in different educational environments, so as to understand teachers' conceptions and implementation of VRDW more comprehensively. Moreover, future research can explore the consistency of teachers' teaching implementation and their conceptions in this context, to find and explain how the conceptions affect their implementation. Also, the conceptions of VRDW held by students also need to be explored in future research, to find out the connections and inconsistency of teachers' and students' conceptions, so as to find and solve the potential gap.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Local Ethics Committee of the Chinese University of Hong Kong. The participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

MC and C-SC contributed to the conceptualization, research design, data collection, and analysis. MJo contributed to the project's principal investigator, funding acquisition, research and development of the ISV-VR system, and overseeing the project. MC, C-SC, MJo, and MJi contributed to the manuscript writing. All authors approved the submitted version.

## FUNDING

The work described in this article was substantially supported by The Hong Kong Jockey Club Charities Trust (Project Title: Jockey Club VR Project for Enhancing Chinese Language Literacy, project S/N ref: 2019/0013).

## REFERENCES


Åkerlind, G. S. (2005). Variation and commonality in phenomenographic research methods. *Higher Educ. Res. Dev.* 24, 321–334. doi: 10.1080/07294360500284672

Åkerlind, G. S., Bowden, J. A., and Green, P. (2005). "Learning to do phenomenography: a reflective discussion," in *Doing Developmental Phenomenography*, eds J. A. Bowden and P. Green (Melbourne: RMIT University Press), 74–100.

Akincilar, V. (2010). *The Effect of "PLEASE" Strategy Training Through the Self-Regulated Strategy Development (SRSD) Model on Fifth Grade EFL Students' Descriptive Writing: Strategy Training on Planning*. Çankaya: Middle East Technical University.

Alshahrani, K., and Al-Shehri, S. (2012). Conceptions and responses to e-learning: the case of EFL teachers and students in a Saudi Arabian university. *Monash Univ. Linguist. Pap.* 8, 21–31.

Babayigit, S., and Stainthorp, R. (2011). Modelling the relationships between cognitive–linguistic skills and literacy skills: new insights from a transparent orthography. *J. Educ. Psychol.* 103, 169–189. doi: 10.1037/a0021671

Bazerman, C. (ed.) (2009). *Handbook of Research on Writing: History, Society, School, Individual, Text*. London: Routledge, doi: 10.4324/9781410616470

Boulton-Lewis, G. M., Smith, D. J. H., McCrindle, A. R., Burnett, P. C., and Campbell, K. J. (2001). Secondary teachers' conceptions of teaching and learning. *Learn. Instr.* 11, 35–51. doi: 10.1016/s0959-4752(00)00014-1

Bower, M., and Jong, M. S. Y. (2020). Immersive virtual reality in education, *Br. J. Educ. Technol.* 51, 1981–1990. doi: 10.1111/bjet.13038

Carter, K. (2015). *Teaching Descriptive Writing Through Visualization and the Five Senses*. Avaliable at: http://americanenglish.state.gov/english-teaching-forum. (accessed January 9, 2020).

Chai, C. S., Liang, J. C., Tsai, C. C., and Dong, Y. (2019). Surveying and modelling China high school students' experience of and preferences for twenty-first-century learning and their academic and knowledge creation efficacy. *Educ. Stud.* 46, 658–675. doi: 10.1080/03055698.2019.1627662

Chan, C. K., Tsang, T. W., Chong, Y. M., Jong, S. Y. M., and Chen, M. Y. (2019). "運用虛擬實境提升初中生的中文寫作能力與人文素養 [Leveraging Virtual Reality to Promote Junior Secondary Students' Chinese Writing Ability and Humanistic Literacy]," in *2019 International Conference on Game-Based Learning (ICGBL)*, (Hangzhou), 249–261. doi: 10.13140/RG.2.2.17922.09925

Chang, S. C., Hsu, T. C., and Jong, M. S. Y. (2020). Integration of the peer assessment approach with a virtual reality design system for learning earth science. *Comput. Educ.* 146:103758. doi: 10.1016/j.compedu.2019.103758

Chao, C. N. G., Jong, S. Y. M., Chai, C. S., and Luk, T. H. (2020). "Development of the motivation and engagement in virtual reality chinese language learning questionnaire (MEVRCLQ)," in *2020 International Symposium on Educational Technology (ISET)*, Bangkok, 67–72. doi: 10.1109/ISET49818.2020.00024

Chen, Y., Smith, T. J., York, C. S., and Mayall, H. J. (2019). Google earth virtual reality and expository writing for young english learners from a funds of knowledge perspective. *Comput. Assist. Lang. Learn.* 33, 1–25. doi: 10.1080/09588221.2018.1544151

Chien, S., Hwang, G., and Jong, M. S. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Comput. Educ.* 146:103751. doi: 10.1016/j.compedu.2019.103751

Choi, T. H. (2016). "Glocalisation of English language education: comparison of three contexts in East Asia," in *Sociological and Philosophical Perspectives on Education in the Asia-Pacific Region*, (Singapore: Springer), 147–164. doi: 10.1007/978-981-287-940-0_10

Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 2nd Edn. Thousand Oaks, CA: Corwin, doi: 10.7748/nr.12.1.82.s2

Cummins, J. (2000). *Immersion Education for the Millennium: What we Have Learned From 30 Years of Research on Second Language Immersion*. Avaliable at: http://www.researchgate.net/profile/Jim_Cummins5/publication/255638397_Immersion_Education_for_the_Millennium_What_We_Have_Learned_from_30_Years_of_Research_on_Second_Language_Immersion/links/00b7d53cb6f02f3b09000000.pdf (accessed November, 30, 2020).

Curriculum Development Council of Hong Kong (2017). *Curriculum Guide for Chinese Language Education and Learning (Primary 1 to Secondary 6)*. Avaliable at: https://www.edb.gov.hk/attachment/tc/curriculum-development/kla/chi-edu/curriculum-documents/CLEKLAG_2017_for_upload_final_R77.pdf (accessed November 27, 2020).

Earnshaw, S. (2007). *The Handbook of Creative Writing*. Edinburgh: Edinburgh University Press.

Freedman, S. W., and Ball, A. F. (2004). "Ideological becoming: bakhtinian concepts to guide the study of language, literacy, and learning," in *Bakhtinian Perspectives on Language, Literacy, and Learning*, eds A. F. Ball and S. W. Freedman (Cambridge: Cambridge University Press), 3–33. doi: 10.1177/0022487109336181

García-Martín, J., and García-Sánchez, J. N. (2018). The instructional effectiveness of two virtual approaches: processes and product. *Rev. Psicod.* 23, 117–127. doi: 10.1016/j.psicoe.2018.02.003

Geng, J., Chai, C. S., Jong, M. S. Y., and Luk, T. H. (2019). Understanding the pedagogical potential of interactive spherical video-based virtual reality from the teachers' perspective through the ACE framework. *Interact. Learn. Environ.* 1–16. doi: 10.1080/10494820.2019.1593200

Holliway, D. R. (2004). Through the eyes of my reader: a strategy for improving audience perspective in children's descriptive writing. *J. Res. Childh. Educ.* 18, 334–349. doi: 10.1080/02568540409595045

Hong Kong Examination and Assessment Authority (2018). *Hong Kong Diploma of Secondary Education (HKDSE) Examiners' Report*. Avaliable at: http://www.hkeaa.edu.hk/tc/HKDSE/assessment/subject_information/category_a_subjects/chi_lang/or/bs2018.html (accessed January 9, 2020).

How, K., and Larkin, R. F. (2013). Descriptive writing in primary school: how useful are linguistic predictors of reading? *J. Educ. Res.* 106, 360–371. doi: 10.1080/00220671.2012.753864

Hsieh, W. M., and Tsai, C. C. (2017). Taiwanese high school teachers' conceptions of mobile learning. *Comput. Educ.* 115, 82–95. doi: 10.1016/j.compedu.2017.07.013

Huang, H. L., Hwang, G. J., and Chang, C. Y. (2020). Learning to be a writer: a spherical video-based virtual reality approach to supporting descriptive article writing in high school Chinese courses. *Br. J. Educ. Technol.* 51, 1386–1405. doi: 10.1111/bjet.12893

Hyland, K. (2002). Authority and invisibility: authorial identity in academic writing. *J. Pragmat.* 34, 1091–1112. doi: 10.1016/S0378-2166(02)00035-8

Jong, M. S. Y. (2019). Sustaining the adoption of gamified outdoor social enquiry learning in high schools through addressing teachers' emerging concerns: a three-year study. *Br. J. Educ. Technol.* 50, 1275–1293. doi: 10.1111/bjet.12767

Jong, M. S. Y., Tsai, C. C., Xie, H., and Wong, F. K. K. (2020). Integrating interactive learner-immersed video-based virtual reality into learning and teaching of physical geography. *Br. J. Educ. Technol.* 51, 2064–2079. doi: 10.1111/bjet.12947

Kelly, and Beth, L. (2017). Preservice teachers' developing conceptions of teaching English learners. *Tesol. Q.* 52, 110–136. doi: 10.1002/tesq.375

Kong, Y. (2018). TEFL teachers' conceptions of writing: a case of China. *Theory Pract. Lang. Stud.* 8, 285–294. doi: 10.17507/tpls.0803.03

Kozar, O. (2016). Perceptions of webcam use by experienced online teachers and learners: a seeming disconnect between research and practice. *Comput. Assis. Lang. Learn.* 29, 779–789. doi: 10.1080/09588221.2015.1061021

Lameras, P., Levy, P., Paraskakis, I., and Webber, S. (2012). Blended university teaching using virtual learning environments: conceptions and approaches. *Instr. Sci.* 40, 141–157. doi: 10.1007/s11251-011-9170-9

Lin, H. M. (2015). Effects of corpus-aided language learning in the EFL grammar classroom: a case study of students' learning attitudes and teachers' perceptions in Taiwan. *Tesol. Q.* 50, 871–893. doi: 10.1002/tesq.250

Lin, S. T. (2016). Changes in teachers' beliefs after a professional development project for teaching writing: two Chinese cases. *J. Educ. Teach.* 42, 106–109. doi: 10.1080/02607476.2015.1135228

Lin, T. J., and Lan, Y. J. (2015). Language learning in virtual reality environments: past, present, and future. *Educ. Techno. Soc.* 18, 486–497. Available online at: http://www.jstor.org/stable/jeductechsoci.18.4.486 (accessed January 21, 2021).

Lin, H. C. S., Yu, S. J., Sun, J. C. Y., and Jong, M. S. Y. (2019). Engaging university students in a library guide through wearable spherical video-based virtual reality: effects on situational interest and cognitive load. *Interact. Learn. Environ.* 1–16. doi: 10.1080/10494820.2019.1624579

Lipson, M. Y., Mosenthal, J., Daniels, P., and Woodside-Jiron, H. (2000). Process writing in the classrooms of eleven fifth-grade teachers with different

orientations to teaching and learning. *Elemen. Sch. J.* 101, 209–231. doi: 10.1086/499665

Liu, H., Lin, C. H., and Zhang, D. (2017). Pedagogical beliefs and attitudes toward information and communication technology: a survey of teachers of English as a foreign language in china. *Comput. Assist. Lang. Learn.* 30, 745–765. doi: 10.1080/09588221.2017.1347572

Macalister, J., and Nation, I. S. P. (2019). *Language Curriculum Design*. London: Routledge, doi: 10.4324/9780429203763-1

MacArthur, C. A., Graham, S., and Fitzgerald, J. (eds) (2016). *Handbook of Writing Research*. New York, NY: Guilford Press.

Mantziou, O., Papachristos, N. M., and Mikropoulos, T. A. (2018). Learning activities as enactments of learning affordances in MUVEs: a review-based classification. *Educ. Informat. Technol.* 23, 1737–1765. doi: 10.1007/s10639-018-9690-x

Marsh, H., Hau, K. T., and Kong, C. K. (2000). Late immersion and language of instruction in Hong Kong high schools: achievement growth in language and non-language subjects. *Harvard Educ. Rev.* 70, 302–347. doi: 10.17763/haer.70.3.gm047588386655k5

Marton, F., and Pong, W. Y. (2005). On the unit of description in phenomenography. *High. Educ. Res. Dev.* 24, 335–348. doi: 10.1080/07294360500284706

McCarthey, S. J., and Mkhize, D. (2013). Teachers' orientations towards writing. *J. Writ. Res.* 5, 1–33. doi: 10.17239/jowr-2013.05.01.1

Mickan, P., and Wallace, I. (eds) (2019). *The Routledge Handbook of Language Education Curriculum Design*. London: Routledge, doi: 10.4324/9781315661032

Mohiuddin, S., Roshan, D., and Knorpp, H. (2016). Utilization of immersive 360 degree spherical videos and google cardboard in medical training and simulation: a novel and multi-dimensional way of learning. *Anesth. Anal.* 122.

Moustakas, C. (1994). *Phenomenological Research Methods*. Thousand Oaks, CA: Sage, doi: 10.1007/978-1-4615-6989-3_3

Nair, S. M., and Sanai, M. (2018). Effects of utilizing the STAD method (Cooperative Learning Approach) in enhancing students' descriptive writing skills. *Int. J. Educ. Pract.* 6, 239–252. doi: 10.18488/journal.61.2018.64.239.252

Newell, G. E., VanDerHeide, J., and Olsen, A. W. (2014). High school English language arts teachers' argumentative epistemologies for teaching writing. *Res. Teach. Eng.* 49, 95–119.

Newton, L., and Beverton, S. (2012). Pre-service teachers' conceptions of creativity in elementary school English. *Think. Skills Creat.* 7, 165–176. doi: 10.1016/j.tsc.2012.02.002

Norton, B. (1997). Language, identity, and the ownership of English. *Tesol. Q.* 31, 409–429. doi: 10.2307/3587831

Ozkeskin, E. E., and Tunc, T. (2010). Spherical video recording and possible interactive educational uses. *Int. J. New Trends Educ. Their. Implicat.* 1, 69–79.

Paakkari, L., Tynjala, P., and Kannas, L. (2011). Critical aspects of student teachers' conceptions of learning. *Learn. Instr.* 21, 705–714. doi: 10.1016/j.learninstruc.2011.03.003

Pratt, D. D. (1992). Conceptions of teaching. *Adult Educ. Q.* 42, 203–220. doi: 10.1177/074171369204200401

Repetto, C., Germagnoli, S., Triberti, S., and Riva, G. (2018). "Learning into the wild: a protocol for the use of 360 video for foreign language learning," in *International Symposium on Pervasive Computing Paradigms for Mental Health*, eds P. Cipresso, S. Serino, Y. Ostrovsky, and J. Baker (Cham: Springer), 56–63. doi: 10.1007/978-3-030-01093-5_8

Steele, J. L., Slater, R. O., Li, J., Zamarro, G., Miller, T., and Bacon, M. (2018). Dual-language immersion education at scale: an analysis of program costs, mechanisms, and moderators. *Educ. Eval. Policy Anal.* 40, 420–445. doi: 10.3102/0162373718779457

Swain, M., and Lapkin, S. (2005). The evolving sociopolitical context of immersion education in Canada: some implications for program development 1. *Int. J. Appl. Linguist.* 15, 169–186. doi: 10.1111/j.1473-4192.2005.00086.x

Taghizadeh, M., and Yourdshahi, Z. H. (2019). Integrating technology into young learners' classes: language teachers' perceptions. *Comput. Assist. Lang. Learn.* 33, 982–1006. doi: 10.1080/09588221.2019.1618876

Teale, W., and Yokota, J. (2000). "Beginning reading and writing: perspectives on instruction," in *Beginning Reading and Writing*, eds D. S. Strickland and L. M. Morrow (New York, NY: Teachers College Press and International Reading Association), 3–21.

Teng, F. (2016). Immediate and delayed effects of embedded metacognitive instruction on Chinese EFL students' English writing and regulation of cognition. *Think. Skills Creat.* 22, 289–302. doi: 10.1016/j.tsc.2016.06.005

Thompson, A. G. (1992). "Teachers' beliefs and conceptions: a synthesis of the research," in *Hand Book of Research on Mathematics Teaching and Learning*, ed. D. A. Grouws (New York, NY: Macmillan), 127–146.

Vasudevan, L., Schultz, K., and Bateman, J. (2010). Rethinking composing in a digital age: authoring literate identities through multimodal storytelling. *Writt. Commun.* 27, 442–468. doi: 10.1177/0741088310378217

Vermunt, J. D., and Vermetten, Y. J. (2004). Patterns in student learning: relationships between learning strategies, conceptions of learning, and learning orientations. *Educ. Psychol. Rev.* 16, 359–384. doi: 10.1007/s10648-004-0005-y

Wang, E. L., and Matsumura, L. C. (2019). Text-based writing in elementary classrooms: teachers' conceptions and practice. *Read. Writ.* 32, 405–438. doi: 10.1007/s11145-018-9860-7

Wilhelm, J. D. (2008). *"You Gotta BE the Book": Teaching Engaged and Reflective Reading with Adolescents*. New York: Teachers College Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/BF03196322

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# Advantages of publishing in Frontiers

**OPEN ACCESS**
Articles are free to read for greatest visibility and readership

**FAST PUBLICATION**
Around 90 days from submission to decision

**HIGH QUALITY PEER-REVIEW**
Rigorous, collaborative, and constructive peer-review

**TRANSPARENT PEER-REVIEW**
Editors and reviewers acknowledged by name on published articles

**Frontiers**
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

**Visit us:** www.frontiersin.org
**Contact us:** frontiersin.org/about/contact

**REPRODUCIBILITY OF RESEARCH**
Support open data and methods to enhance research reproducibility

**DIGITAL PUBLISHING**
Articles designed for optimal readership across devices

**FOLLOW US**
@frontiersin

**IMPACT METRICS**
Advanced article metrics track visibility across digital media

**EXTENSIVE PROMOTION**
Marketing and promotion of impactful research

**LOOP RESEARCH NETWORK**
Our network increases your article's readership