Research on Preschool and Primary Education
https://ojs.luminescience.cn/RPPE

LUMINESCIENCE
https://luminescience.cn

Review article

Optimal delivery methods in primary school safety education: A scoping review across multiple programs

Chenruisi Xu

Department of Psychology, Durham University, UK

Correspondence to: Xu Chenruisi, Email: chenruisi.xu@durham.ac.uk

Abstract: This review evaluates the efficacy of two delivering methods for primary school students: computerized behavioral-skills training and non-computerized behavioral-skills training across fire safety, road safety, cyber safety, and abduction prevention. Evaluating forty relevant studies through sieve (design, scale, attrition, and outcome measure quality) this review reveals mixed findings. Despite prevalent limitations in many studies, such as inadequate comparators and small sample sizes, computerized delivery appears more effective for fire and road safety programs, while non-computerized methods show higher efficacy in cyber safety and abduction prevention programs. The results highlight the need for tailored delivery methods in specific safety education programs, this review underscores the importance for educators, researchers, and policymakers to discern the most suitable approaches for each program, enhancing the overall quality of safety education initiatives.

Keywords: Safety education, Scoping review, Behavioural skills training & Computerized-behavioural skills training

Introduction

Children often lack knowledge of self-protection in dangerous situations (Beck & Miltenberger, 2009). Safety education programs, BST (Behavioural Skills Training) and CBST (Computerised Behavioural Skills Training), are effective methods for teaching primary school children safety skills to prevent them from getting hurt (Vanselow & Hanley, 2014). BST involves direct instruction, modeling, rehearsal, and feedback (Beck & Miltenberger, 2009), while CBST (Jostad & Miltenberger, 2004) applies multimedia, such as video recordings, to modify children's behavior without direct teacher instruction. While both methods are effective, previous reviews of BST and CBST mostly focus on the effectiveness of a specific teaching method within a certain program (e.g., the effectiveness of BST for fire safety education) rather than comparing these two methods within the same program (e.g., Lazarinis et al., 2020; Raftery & Wundersitz, 2011). This study aims to evaluate the effectiveness of BST and CBST in teaching primary school safety education programs worldwide, including FSE (Fire Safety Education), RSE (Road Safety Education), CSE (Cyber Safety Education), and AP (Abduction Prevention).

Hwang et al., (2006) conducted a randomized controlled trial in which the short-term effectiveness of the BST method in teaching fire safety knowledge was investigated. The experimental group (BST) showed improvements in both behavioral and knowledge tests compared to the control group (national leaflet), suggesting that BST is more effective than the national leaflet in adjusting children's fire misuse behavior in the short term. Lambie and Randell

Received: Oct.28, 2024; Revised: Mar.6, 2025; Accepted: Apr.14, 2025; Published: Apr.30, 2025

DOI: https://doi.org/10.55976/rppe.320251317145-156

(2010) found that even six months after BST training, none of the children with a history of fire misuse displayed fire-setting behaviors or intentions. While a serious history of fire misuse may hinder continuous improvement, BST still demonstrated significant short-term adjustments to unsafe behavior.

For RSE, Fraser and Sentsho (2012) observed persistent unsafe crossing behaviors despite children's knowledge of traffic rules, but BST intervention led to over 80% of children adopting safer habits. Previous RSE studies (Ikeda et al., 2020; Fraser & Sentsho, 2012) indicated BST's effectiveness in enhancing road safety knowledge and rectifying unsafe conduct in the short term. However, they also found that most children were already aware of traffic rules before training. It could be deduced that these unsafe behaviors are not caused by a lack of knowledge, and BST could help children adjust to proper behavior in the short term. For long-term effects, Alonso et al., (2018) found that after three months, more than half of the children resumed unsafe habits, such as crossing on red lights. Despite BST's evident short-term and knowledge impacts, its long-term influence on road safety behaviors remains uncertain.

Given that over 8 million children go missing annually during unsupervised play or separation in public areas (Safeatlast, 2022), abduction prevention (AP) programs for children have started to gain attention. Previous studies (Carver et al., 2000; Tarasenko et al., 2010) evaluated BST's efficacy in AP lessons for 7-year-olds and found positive behavioral test results. Subsequent research combined in-situ training with BST, using real strangers in classrooms, which enhanced efficacy compared to BST alone and maintained its impact over time. The study by Kulkofsky (2021) supported these findings, showing that the BST & in-situ group sustained correct behavior after six months, outperforming the control group. In conclusion, combining BST with in-situ training effectively imparts AP skills, surpassing BST alone, with both variations demonstrating significantly prolonged effects compared to other programs. The CBST method has gained popularity in teaching RSE programs, showcasing both short-term and long-term effects. Vankov and Jankovszky's (2021) research shows that during the post-VR session, 83.5% of children demonstrated a significant improvement in traffic-related knowledge and behavior compared to the pre-test. In addition, Rawi et al.,(2015) show that applying VR in road safety modules can generate a long-term positive impact on primary school children. On the other hand, CBST demonstrates a longer duration of the positive impact compared to BST. For example, Purcell and Romijn (2017) observed that VR users sustained correct road-safe behavior longer than BST participants after six months. This may be attributed to VR's realistic traffic scenarios compared to fire safety scenarios, underpinning the lasting impact of VR in RSE programs.

CBST is widely used in CSE (Grissom et al., 2003). Kritzinger (2015) developed CBST games targeting cyber safety knowledge and preventing internet addiction. These games guide appropriate online behavior, assessing actions like avoiding suspicious links. Tested in five local primary schools, the program received positive feedback, with over half of the parents requesting parental monitoring features. Addressing concerns, Zinkus et al. (2019) introduced "Fakesbook," offering parental monitoring of online activities. Despite controversy over privacy invasion and risks, Fakesbook significantly reduced children's online time and unsafe behaviors in beta testing. Both CBST methods positively influence children's CSE in the short term, indicating a notable impact on cyber safety behavior and knowledge. However, the long-term impact of CBST on CSE remains debatable. The results of Varcoe et al. (2020) showed consistently positive feedback, but the sample in this study consisted of adult women rather than children, and there appears to be a lack of empirical evidence for CBST's prolonged effectiveness in CSE for primary school children.

However, few CBST apps focus on educating children about stranger danger. Most programs are designed for children with autism spectrum disorder (ASD) and have shown notable short-term and long-term effectiveness. For example, the study by Ledbetter-Cho (2021) evaluated a game-based VR platform for teaching AP skills to three children with ASD. The results showed that two of the children continued to exhibit correct responses in abduction scenarios after 16 weeks, without reporting fear or difficulty in interacting with unfamiliar individuals, unlike the potential side effects associated with BST. Both short-term and long-term effects were evident in these CBST studies, but there seems to be a lack of validity for typical children.

In the Vanselow and Hanley's review (2014), two safety education programs at elementary school were evaluated using the BST and CBST methods. The results show that postitive feedback woth a medium to large effect size was reported for the AP program in nine out of eleven studies. Additionally, four longitudinal studies showcased a more substantial long-term effect with the BST method, indicating its higher effectiveness in teaching AP skills to primary school children. Their analysis of FSE suggested that CBST demonstrates higher effectiveness in delivering fire safety knowledge, supported by previous evidence indicating a larger long-term impact of CBST on children's FSE knowledge and behavior.

The review by Giannakakos et al., (2021) evaluated three safety education programs (FSE, RSE, and AP) delivered through BST and CBST methods. For FSE, the overall results suggested that CBST might have a greater long-term effect compared to BST. Vanselow and Hanley (2014) found similar results in their evaluation of studies related to RSE, indicating that CBST, especially with VR for behavioral modification, showed a larger effect size. However, the evaluation results of the AP program indicated that BST might be more effective than CBST in teaching primary school children (Vanselow & Hanley, 2014).

Previous studies show three limitations. The articles

reviewed by Giannakakos et al. (2021) mostly focused on studies with relatively small sample sizes, making it challenging to draw comparisons between the effectiveness of teaching methods. Another limitation is the constrained research scope; as cyber safety issues, such as cyberbullying, have gained prominence in today's society, evaluating the effectiveness of CSE methods has become increasingly crucial (Zhang-Kennedy & Chiasson, 2020). Lastly, there seems to be a lack of research standards for evaluating research quality. These limitations impact the depth and specificity of the conclusions drawn from the reviewed studies.

To address these three limitations and generate more universal suggestions for safety education programs, this study will conduct a scoping review using published standards to evaluate research quality. This study will adapt the Sieve framework (Gorard, 2021) to estimate the reliability of the chosen studies. Additionally, CSE programs will be included in this study. Finally, the sample sizes of the selected studies will be carefully considered to evaluate whether they are sufficient to demonstrate the effectiveness of BST or CBST.

Research questions

RQ1: Which method could be more effective in teaching primary school children different safety education programs?
RQ2: Why these teaching methods are more suitable?
RQ3: Are teaching outcomes affected by electronic device use proficiency?

Methodology

Sieve evaluated a study from four perspectives (design, scale, missing data, and measurement quality). If this study was rated 4, this study has a strong design for the research question (e.g. RCT), large scale number, minimal missing data, and standardized, independent, reasonably accurate measurement method (Gorard, 2021). The detail evaluation criteria of Sieve are in Table 1.

Table 1. A 'Sieve' to assist in the estimation of trustworthiness (Gorard, p.94, 2021)

Design	Scale	Missing data	Measurement quality	Rating
Strong design for research question	Large number of cases (per comparison group)	Minimal missing data, no impact on findings	Standardized, independent, reasonably accurate	4
Good design for research question	Medium number of cases (per comparison group)	Some missing data, possible impact on findings	Standardized, independent, some errors	3
Weak design for research question	Small number of cases (per comparison group)	Moderate missing data, likely impact on findings	Not Standardized, independent, major possible errors	2
Very weak design for research question	Very small number of cases (per comparison group)	High level of missing data, clear impact on findings	Weak measures, high level of error or many outcomes	1
No consideration of design	A trivial scale of study	Huge amount of missing data, or not reported	Very weak measures	0

Inclusion criteria

1. Primary school children from 6 to 12 years old
2. CBST include VR, AR, mobile APP and computer games
3. Children were instructed in school by teachers

Exclusion criteria

1.Non-English or Non-Chinese articles
2.Articles published before 2000

Review steps FSE

1. Searching key words in Google Scholar, Web of Science, and ECOSB. include "fire safety education, primary school, elementary school, children, BST, computer, VR/AR".
2. Selecting experimental studies according to inclusion and exclusion criteria

RSE

1. Searching key words in Google Scholar, Web of Science, and ECOSB include "road safety education, primary school, elementary school, children, BST, computer, VR/AR".
2. Selecting experimental studies according to inclusion and exclusion criteria.

CSE

1. Searching key words in Google Scholar, Web of Science,

and ECOSB include "cyber/online safety education, primary school, elementary school, children, BST, computer, VR/AR".
2. Selecting experimental studies according to inclusion and exclusion criteria.

AP

1. Searching key words in Google Scholar, Web of Science, and ECOSB. include "Abduction prevention, primary school, elementary school, children, BST, computer, VR/AR".
2. Selecting experimental studies according to inclusion and exclusion criteria.

Result

FSE

The initial three BST studies investigated community-based fire safety presentations using pre- and post-tests, questionnaires or behavioral tests (e.g., activating fire alarms). The study by Charez et al. (2013) indicated the positive effect of the Danger Ranger talk on fire safety knowledge; however, tests of behavior change related to fire misuse appear to be lacking. In contrast, Giannakakos et al. (2018) applied behavioral tests and showed a significant increase in emergency response after the lectures, but the sample size appears too small to draw generalizable conclusions. Two experimental designs highlighted the positive impact of community-based lectures and their long-term effects. Lehna and Erika (2014) observed positive outcomes in both the group with the national guide and the group with the community-based lecture groups, with slightly better results in the experimental group. Kolko (2001) similarly noted long-lasting effects from community-based lectures, attributing their effectiveness to adaptability and student engagement.

In contrast, three CBST studies employing pre- and post-baseline designs show the significant influence of the intervention. Cakiroglu and Gokoglu (2019) and Houvouras and Harvey (2014) indicated the significant impact of VR technology and computer programs on fire safety education, demonstrating both immediate and sustained effects. Similar findings were also reported by Oliva et al., (2019). Smith and Ericson's (2019) randomized control study favored VR over community-based teaching for fire safety, showcasing shorter response times to fire alarms in the VR group. Bickel (2017) observed significantly higher knowledge increments in the digital app group compared to the paper card group, while Morrongielli et al., (2012) echoed the superiority of VR over community-based lectures, revealing sustained effectiveness even after six months and suggesting VR's superior long-term impact. The details of the study design and results of the "Sieve" evaluation can be found in Table 2.

Table 2. Evaluation for fire safety education programs

	Design	Sample size	Test method	Effectiveness	Sieve rating
Chavez, et.al., (2013) (P)	Pre- and post- baseline design (3)	166 (4,4)	Questionnaires (2)	77.1%	2
Giannakakos, et.al., (2018) (P)	Pre- and post- baseline design (3)	6 (1,1)	Behavioural test (observation) (1)	83.3%	1
Lehna & Erika (2014) (P)	Experimental design (4)	57 (3,3)	Behavioural test (observation) (2)	96.6%	2
Kolko (2001) (P)	Experimental design (4)	30 (2,2)	Behavioural test (observation) (2)	100%	2
Houvouras & Harvey (2014) (P)	Pre- and post- baseline design (2)	3 (1,1)	Behavioural test (observation) (1)	66.7%	1
Cakiroglu & Gokogln (2019) (P)	Pre- and post- baseline design (3)	10 (1,1)	Behavioural test (observation) (1)	100%	1
Oliva, et.al., (2019) (P)	Pre- and post- baseline design (3)	169 (3,3)	Behavioural test (RTs) (3)	80.5%	3
Bickel (2017) (P)	Experimental design (4)	20 (2,2)	Behavioural test (RTs) (2)	100%	2
Smith & Ericson (2009) (P)	Experimental design (4)	22 (2,2)	22 (2,2)	95.5%	2
Mornongiello, et.al., (2012) (P)	Experimental design (4)	76 (3,3)	76 (3,3)	94.7%	3

Note. Numbers in the brackets represent the ratings of sieve items, and sample size include the ratings of both scale and missing data. P represent positive effect; Neu represent neutral effect and Neg represent negative effect.

RSE

Both Morrongiello et al. (2015) and Purcell and Romijin (2020) applied pre-post designs to assess road safety education programs. Morrongiello et al.,'s (2015) VR-based lessons showed a significant improvement (96.8%) in both knowledge and behavior, whereas Purcell and Romijin's app-based intervention showed an increase in knowledge enhancement (83%). However, due to the lack of behavioral evaluation and a high dropout rate (82 children), the validity of this study was diminished. In contrast, Cerasi et al., (2021) and McComas et al., (2002) compared VR and BST methods, revealing VR's higher effectiveness in both knowledge and behavior. Unlike the previous studies that recruited children as participants, Salfarina et al., (2021) designed an iPad game for road safety but relied on the ratings of experts, 75.9% of whom found the game useful. Although it seems useful for teaching children the RSE program, it would be better to test it with actual children.

In BST studies, Trevino-Siller et al., (2016) and Nkuruho et al., (2021) found that children exhibited both positive knowledge and behavior changes after the BST intervention. Additionally, Alonso et al., (2018) observed significant knowledge improvement but minor behavior changes in a larger sample, suggesting ambiguous efficacy regarding behavior change. Zeedyk et al., (2001) demonstrated a substantial knowledge increase but limited behavior change (30%), attributed to a significant dropout rate (73 children), which impacted the validity of the results. Bojesen and Rayce (2019) compared textbook and BST learning, demonstrating BST's superior knowledge acquisition (large effect size: 0.18) but found no significant behavioral difference. The details of the study design and results of the "Sieve" evaluation can be found in Table 3.

Table 3. Evaluation for road safety education programs

	Design	Sample size	Test method	Effectiveness	Sieve rating
Morrongiello et al., (2015) (P)	Pre- and post-baseline design (2)	95 (2,2)	Questionnaires & Behavioural test (2)	96.8%	2
Salfarina et al., (2021) (P)	Pre- and post-baseline design (2)	120 (2,2)	Expert evaluation (0)	75.9%	0
Purcell & Romijin (2020) (P)	Pre- and post-baseline design (2)	218 (2,2)	Questionnaire (1)	83%	1
Cerasi et al., (2021) (P)	Experimental design (4)	61 (2,2)	Behavioural test & Questionnaire (2)	48.3%	2
McComas et al., (2002) (P)	Experimental design (4)	95 (2,2)	Behavioural test & Questionnaire (0)	Not reported	0
Trevino – Siller et al., (2016) (P)	Pre- and post-baseline design (2)	219 (2,2)	Behavioural test & Questionnaire (1)	69.505%	1
Nkuruho et al., (2021) (P)	Pre- and post-baseline design (2)	100 Primary Schools (2,2)	Questionnaire (1)	83%	1
Alonso et al., (2018) (P)	Pre- and post-baseline design (2)	1930 (2,2)	Behavioural test & Questionnaire (1)	80.5% Knowledge 25% Behaviour	1
Zeedyk et al., (2001) (P)	Experimental design + pre- and post-test (4)	120 (2,0)	Behavioural test & Questionnaire (0)	80.4% Knowledge 30% Behaviour	0
Bojensen & Rayce (2019) (P)	Experimental design (4)	3536 (4,4)	Behavioural test & Questionnaire (4)	80.4% Knowledge 30% Behaviour	4

Note. Numbers in the brackets represent the ratings of sieve items, and sample size include the ratings of both scale and missing data. P represent positive effect; Neu represent neutral effect and Neg represent negative effect.

CSE

Studies employing CBST to educate elementary students about cyber safety showed mixed results. Zahed et al. (2019) and Zinkus et al., (2019) explored the impact of computer games on cyber safety knowledge, indicating a positive effect on knowledge and a minor effect on behavior change (e.g., avoiding unsafe links). However, the observational results showed no substantial change in unsafe behavior. According to the interview results of Lazarinis et al., (2020) regarding the effectiveness of a mobile teaching app, children reported that they tended to

focus on experience rather than on learning knowledge and adjusting behavior. For this reason, Mishna et al., (2011) emphasized the importance of evaluating teaching apps for their impact on behavior and knowledge rather than user experience, which is crucial given children's extensive online time and insufficient cyber safety knowledge (Macauley et al., 2019). These findings suggest that while CBST enhances awareness and knowledge, altering unsafe behaviors remains challenging.

To test the effectiveness of BST, Ford-Gilboe et al., (2020) conducted a randomized controlled trial (RCT) with 450 children, comparing cyber safety education via computer games and textbooks. The experimental group exhibited a larger effect size than the control group, suggesting the superiority of game-based learning. Similarly, Asain et al. (2019) noted the effectiveness of computer games in enhancing cyber safety knowledge but lacked behavioral assessments. Studies by Zinicola (2021), Gordillo et al., (2021), and Nicolaidou and Venizalou (2020) utilizing BST showcased positive outcomes in both cyber safety knowledge and behavior change. Nicolaidou and Venizalou's study particularly highlighted BST's ability to modify unsafe cyber behaviors compared to CBST, despite the limited sample sizes in these studies. The details of the study design and results of the "Sieve" evaluation can be found in Table 4.

Table 4. Evaluation for cyber safety education programs

	Design	Sample size	Test method	Effectiveness	Sieve rating
Asain et al., (2019) (P)	Pre- and post & Experimental design (3)	151 (2,1)	Questionnaires (1)	93.4%	1
Zahed et al., (2019) (P)	Pre- and post- baseline design (1)	49 (1,1)	Behavioural test & Questionnaire (1)	24.5% (B) 73.5% (Q)	1
Lazarinis et al., (2020) (P)	Pre- and post- qualitative design (0)	4 (0,0)	Interview (0)	100%	0
Ford – Gilboe et al., (2020) (P)	Experimental design (3)	450 (3,3)	Behavioural test & Questionnaire (3)	92.5%	3
Zinkus et al., (2019) (P)	Pre- and post- baseline design (1)	55 (1,1)	Questionnaire (1)	80.0%	1
Zinicola (2021) (P)	Pre- and post- baseline design (2)	10 (0,0)	Questionnaire (0)	100%	0
Nicolaidou & Venizelou (2020) (P)	Pre- and post & Experimental design (3)	48 (3,2)	Questionnaire & Behavioural test (1)	80.7% (Q) 73.1% (B)	1
Gordillo et al., (2021) (P)	Pre- and post- baseline design (1)	179(1,0)	Behavioural test (0)	80.7% (Q) 73.1% (B)	0

Note. Numbers in the brackets represent the ratings of sieve items, and sample size include the ratings of both scale and missing data. P represents positive effect; Neu represent neutral effect and Neg represent negative effect. Q and B represent the results of effectiveness of questionnaire and behavioural test separately.

AP

Muller et al., (2014) applied an experimental control design to assess online behavioral training's effectiveness in teaching abduction prevention (AP) skills to elementary students. Results favored the online game over textbooks, indicating higher effectiveness in both questionnaire and behavioral tests. This suggests that CBST positively impacts AP knowledge and behaviors. Similarly, the study of Jones and Pozzebon (2010) and Jones et al., (2020) showed significant effectiveness through CBST, with the latter combining the game with textbooks, resulting in higher effectiveness. Different from quantitative research, Badillo-Urquiola et al., (2019) conducted qualitative research on a mobile teaching APP for AP skills and received positive feedback from interviews. However, the studies above lacked knowledge and behavioral tests and involved only a small sample size, which hinders broader applicability in schools.

Five BST studies indicate significant impacts on AP knowledge and behavior. Weatherley et al., (2012) demonstrated notable improvements in the experimental group, similar to the findings in White et al., (2018a), which emphasized behavioral changes. Johnson et al., (2006) and Berube et al., (2020) incorporated in situ training, reporting notable improvements, albeit with small sample sizes. In contrast, White et al., (2018b) failed to show significant behavioral changes post-BST, despite enhanced knowledge. This study lacked a control group and had a medium-small sample size, limiting its conclusive findings. Building on these results, White et al., (2018c) improved their approach with a 6-month BST session involving 611 children. The study revealed increased AP knowledge in both groups, but significant behavioral improvements were observed only

in the experimental group. This suggests that while both textbooks and BST enhance AP knowledge, long-term BST programs are necessary to effectively modify AP behaviors. The details of the study design and results of the "Sieve" evaluation can be found in Table 5.

Discussion

FSE & RSE

The results regarding more effective delivery methods for FSE and RSE were consistent with previous reviews (Vanselow & Hanley, 2014; Giannakakos et al., 2020) indicating that CBST (computerized skills training) can be more effective in teaching primary school children. The consistent results in FSE and RSE programs may be attributed to modern technology, such as VR, which could be more engaging for children compared to textbooks or PowerPoint presentations when learning and practicing safety skills in a near-real situation in the form of a game (McComas, Pivik, & Laflamme, 1998). From the perspective of learning outcomes, Himle and Wright (2014) suggested that the learning results of FSE and RSE programs for primary school children could be enhanced through VR practice.

Table 5. Evaluation for abduction prevention programs

	Design	Sample size	Test method	Effectiveness	Sieve rating
Badillo-Urquiola et al., (2019) (P)	Qualitative design (0)	6 (0,0)	Questionnaires & Behavioural test (0)	96.8%	0
Muller et al., (2014) (P)	Experimental design (3)	286 (3,3)	Expert evaluation (3)	75.9%	3
Jones & Pozzebon (2010) (P)	Pre- and post + Experimental design (3)	64 (2,2)	Questionnaire (2)	83%	2
Jones et al., (2020) (P)	Pre- and post + Experimental design (4)	126 (3,3)	Behavioural test & Questionnaire (2)	48.3%	2
Weatherley et al., (2012) (P)	Pre- and post + Experimental design (3)	261 (3,3)	Behavioural test & Questionnaire (1)	Not reported	1
White et al., (2018a) (P)	Pre- and post + Experimental design (3)	140 (2,2)	Behavioural test & Questionnaire (2)	69.505%	2
White et al., (2018b) (P)	Pre- and post + Experimental design (2)	118 (2,1)	Questionnaire (1)	83%	1
White et al., (2018c) (P)	Pre- and post + Experimental design (3)	611 (3,3)	Behavioural test & Questionnaire (3)	80.5% Knowledge 25% Behaviour	3
Johnson et al., (2006)	Experimental design + pre- and post-test (4)	46 (2,2)	Behavioural test & Questionnaire (1)	80.4% Knowledge 30% Behaviour	1

Note. Numbers in the brackets represent the ratings of sieve items, and sample size include the ratings of both scale and missing data. P represent positive effect; Neu represent neutral effect and Neg represent negative effect.

CSE & AP

The results indicating that BST is more suitable for teaching children the CSE program contrast with the findings of Giannakakos et al., (2020). This discrepancy may be due to differences in the sample sizes of the chosen articles, as this study includes more studies with medium or large sample sizes rather than focusing solely on small sample sizes. Larger sample sizes may provide more valid conclusions about which delivery method is more suitable for teaching children CSE. Therefore, we suggest that BST is still more suitable for teaching children CSE knowledge.

The results of the AP program evaluation were consistent with the studies of Giannakakos et al., (2020) and Vanselow and Hanley (2014), which showed that BST was more effective in teaching primary school children AP programs. Additionally, empirical evidence from Baruni and Miltenberger (2022) supports this result, demonstrating that children exhibited more long-term positive feedback in AP knowledge and behavior after BST lessons compared to CBST. Therefore, it can be concluded that BST has higher effectiveness in teaching primary school children AP skills.

BST vs CBST

VR technology is more attractive and engaging compared to traditional BST for delivering FSE and RSE, especially for curious children (Freina & Ott, 2015). According to Schwebel et al., (2014), children can achieve higher learning efficiency when they learn in a safe, controlled environment. Therefore, from the perspective of children's learning experience, using VR for FSE and RSE programs would be a beneficial option.

However, the study by Smith and Ericson (2009) compared the learning outcomes of children from BST and CBST groups in the area of fire safety and found no significant difference between the learning outcomes of the two methods. Nevertheless, further self-reported results indicated that children were more engaged in VR compared to traditional classroom sessions. The results of this study suggest that although BST and CBST methods are similarly effective in teaching children fire safety knowledge and behavior, VR may provide a more enjoyable learning experience. Even though this study indicated similar effectiveness, most comparative studies nevertheless indicate a significant difference between BST and CBST, with CBST demonstrating a larger and longer-lasting impact in teaching children fire safety knowledge (e.g., Romanova et al., 2020). Therefore, the effectiveness of BST and CBST remains a topic of debate.

For CSE, on the other hand, CBST's efficacy was not higher than that of BST in teaching children about cyber safety, reflecting differences in the intent of instructional materials (e.g., Zhang-Kennedy & Chiassan, 2020). Zhang-Kennedy and Chiassan (2020) used computer games that emphasized entertainment and user experience (Lazarinis et al., 2020) rather than focusing on cyber safety learning. Although these apps showed positive results in user training and cybersecurity awareness, their primary purpose may require refinement. The outcomes of the AP (abduction prevention) program were consistent with prior research (e.g., Vanselow & Hanley, 2014), demonstrating BST's heightened effectiveness in teaching AP skills compared to CBST. BST's success may stem from its focus on realistic scenarios, such as involving parents as strangers in classrooms or playgrounds (Gunby et al., 2010). This practical approach fosters a deeper understanding and connection to real-life situations compared to VR games. Unlike fire or traffic simulations in FSE and RSE programs, AP skills can be safely trained with actual people. Consequently, VR technology or CBST appears less effective than BST for teaching AP skills due to their focus on controlled, realistic scenarios (Rossi et al., 2017).

The result that BST was more effective in CSE and AP suggests that safety knowledge requiring more real-person involvement is better taught through BST. Poche et al. (1988) indicated that BST can provide children with more detailed feedback about safety behaviors, facilitating effective learning. Baruni and Miltenberger (2022) found that BST shows greater effectiveness when combined with in-situ training. In contrast, the combination of BST and VR yielded mixed results rather than a consistently larger effectiveness. Similar findings were demonstrated by Johnson et al. (2006), which showed that BST plus in-situ training was significantly more effective than BST plus VR in AP programs. These results imply that situations that children can easily relate to or experience in their daily lives are better taught using the BST method (Miltenberger et al., 2013).

Moreover, two review articles (Smith & Ericson, 2009; McComas et al., 2002) mentioned that CBST learning outcomes may be related to the familiarity with technology, which could be influenced by children's location. Smith and Ericson (2009) pointed out that the similar level of improvement observed in children after the VR intervention might be due to the fact that this intervention was conducted in only one school, where the children shared a similar level of basic knowledge about both traffic rules and familiarity with VR. Additionally, most interventions took place within the same school, as is the case with many studies included in this review. Children in the same school often come from similar family backgrounds, making it likely that they possess similar foundational knowledge about modern technology.

It could be argued that schools from different areas may show significant differences in learning outcomes from CBST programs. The results of McComas et al. (2002) showed that children from rural areas achieved poorer learning outcomes compared to children from urban areas. The pre-test results suggested that children from both urban and rural areas had the same level of basic knowledge about pedestrian safety (this study did not assess familiarity with technology). However, after the CBST intervention (VR), urban children demonstrated significant improvement in both traffic rules and unsafe behaviors, while rural children did not. They showed a better understanding of traffic rules and slight changes in unsafe behavior, but the changes did not seem to last as long compared to those from urban schools. This difference may not only stem from varying levels of familiarity with VR but also from differences in law enforcement and basic infrastructure between urban and rural areas, which could influence the duration of children's learning outcomes (Barrett et al., 2019).

In general, these additional results suggest that children from different locations showed varying levels of improvement after the CBST intervention, highlighting the effectiveness of CBST. However, it is still debatable whether familiarity with technology is related to the results of the CBST intervention in delivering safety education programs, as other factors, such as law enforcement and facility complexity, could also play a role. This presents an interesting research question for future studies. Furthermore, the discussion of differences between rural and urban areas also implies that in developing countries, VR or modern technology is not as prevalent as in developed nations. Consequently, many developing countries rely on storytelling or traditional BST for teaching, highlighting a

disparity in educational resources between developed and developing regions (Figueiredo et al., 2018).

Limitation

However, this study has five limitations. The first limitation is the restricted range of investigated safety programs for children. According to Hwang et al., (2015), children can also be harmed by other sources, such as poison and bullying. Future reviews of safety education programs could include these relevant topics to provide a more comprehensive understanding of this field.

The second limitation is that only fully functional children aged 6 to 12 years were studied in this review. There is considerable research on safety education for children with Autism Spectrum Disorder (ASD) or disabilities, as these children are often more vulnerable than their typically developing peers (Klas et al., 2015). Additionally, many studies have indicated that adolescents and the elderly also require safety education programs to acquire essential safety knowledge. For instance, developmental psychologists (Ellis et al., 2012) point out that it is important to modify teenagers' risky behaviors by educating them to reduce the motivation for engaging in dangerous activities (e.g., drug use). The elderly, on the other hand, need to learn how to protect themselves effectively due to their lower cognitive abilities (Duchossois et al., 2009).

The third limitation concerns the potential bias in the "sieve" evaluation, as the ratings were agreed upon by only one researcher, which can lead to subjective bias. For example, if one researcher gave a study a score of 4 while other reviewers agreed on a grade of 3, this discrepancy indicates the need for a more collaborative approach. Involving multiple researchers in the rating process and employing a double-blind review could help mitigate potential biases.The fourth limitation is that most selected studies were conducted in developed countries. Future research would benefit from including findings from more developing countries, as children in those regions often experience riskier living conditions.

Finally, the fifth limitation is publication bias. Studies reporting low or neutral effects of certain instructional methods are less likely to be published, leading to potentially misleading findings that may appear as "false positives." However, preventing publication bias, especially in the social sciences, remains a significant challenge (Franco et al., 2014). The implementation of preventive measures, such as the inoculation theory proposed by Lewandowsky and Linden (2021), involves introducing accurate definitions and application methods for BST and CBST. Lilienfeld et al. (2012) suggest involving those with misconceptions in scientific research to help distinguish accurate theories from false beliefs, although the empirical effectiveness of these prevention methods remains questionable.

Future suggestions

For government

The first one is to apply CBST for FSE and RSE programs and BST for CSE and AP programs. The second is to balanc educational resources in different areas and try to develop more effective methods to teach students safety knowledge. The last one is to test and apply the prevention methods (Lewandowsky & Linden, 2021; Lilienfeld et al., 2012) to avoid the misuse of BST and CBST in research and teaching. In this way, primary school children can learn more complete safety knowledge with suitable methods by well-trained teachers.

For researchers

To generate higher quality empirical studies about this topic, Oberauer and Lewandowsky (2019) stress the importance of third-party Randomized Controlled Trials (RCTs) to validate suggestions before real-world implementation, thereby avoiding publication bias. Misuse and misunderstanding regarding educational theories, exemplified by studies like Charez et al. (2013) and Giannakakos et al. (2013), emphasize the need for precise testing aligned with theoretical concepts to prevent confusion about the application of BST and CBST in safety education.

To address the limitations of this study, the first recommendation is to include a wider range of safety education programs in the reviews. The second is to encompass a broader age range and not focus solely on a fully functional sample, to discover better methods for learning safety knowledge. The third is to aim for a larger sample size and comparison groups when testing the teaching methods (Vanselow & Hanley, 2014). Lastly, it is essential to involve more reviewers when applying the sieve evaluation to obtain fewer subjective results and include more unpublished results. By doing so, future studies in the safety education area will yield more objective results and generate more universal findings and suggestions that can contribute to safety education.

Conflict of interest

There is no conflict of Interest.

Acknowledgements

I would like to pay my sincere gratitude to my supervisor Prof. Stephen Gorard for supporting me finishing this article.

References

Alonso, F., Gonzalez-Marin, A., Esteban, C., & Useche, S. A. (2020). Behavioral health at school: do three competences in road safety education impact the protective road behaviors of Spanish children?. *International Journal of Environmental Research and Public Health, 17*(3), 935.

Alotaibi, F., Furnell, S., Stengel, I., & Papadaki, M. (2016). A review of using gaming technology for cyber-security awareness. *Int. J. Inf. Secur. Res.(IJISR), 6*(2), 660-666.

Asain, M., Ayi, B. N., & Wodi, I. (2019). Effect of guided e-learning on students 'academic achievement in Bayelsa state. *Journal of the Nigerian Council of Educational Psychologists, 14*(1).

Barrett, P., Treves, A., Shmis, T., & Ambasz, D. (2019). The impact of school infrastructure on learning: A synthesis of the evidence (pp. 33-36).

Badillo-Urquiola, K., Smriti, D., McNally, B., Golub, E., Bonsignore, E., & Wisniewski, P. J. (2019). Stranger danger! social media app features codesigned with children to keep them safe online. In Proceedings of the 18th ACM International Conference on Interaction Design and Children (pp. 394-406).

Baruni, R. R., & Miltenberger, R. G. (2022). Teaching safety skills to children: A discussion of critical features and practice recommendations. *Behavior analysis in practice, 15*(3), 938-950.

Beck, K. V., & Miltenberger, R. G. (2009). Evaluation of a commercially available program and in situ training by parents to teach abduction-prevention skills to children. *Journal of applied behavior analysis, 42*(4), 761-772.

Bickel, E. (2017). The effect of a mobile application on fire safety education. *Journal of the South Carolina Academy of Science, 15*(2), 10.

Bojesen, A. B., & Rayce, S. B. (2020). Effectiveness of a school-based road safety educational program for lower secondary school students in Denmark: A cluster-randomized controlled trial. *Accident Analysis & Prevention, 147*, 105773.

Çakiroğlu, Ü., & Gökoğlu, S. (2019). Development of fire safety behavioral skills via virtual reality. *Computers & Education, 133*, 56-68.

Carver, A., Timperio, A., & Crawford, D. (2008). Playing it safe: The influence of neighbourhood safety on children's physical activity—A review. *Health & place, 14*(2), 217-227.

Cerasi, I. R., Moe, D., Skjermo, J., & Wigum, J. P. (2021). Innovative Road Safety Education Program.

Chavez, A. A., Duzinski, S. V., Wheeler, T. C., & Lawson, K. A. (2014). Teaching safety at a summer camp: Evaluation of a fire safety curriculum in an urban community setting. *Burns, 40*(6), 1172-1178.

Dixon, D. R., Bergstrom, R., Smith, M. N., & Tarbox, J. (2010). A review of research on procedures for teaching safety skills to persons with developmental disabilities. *Research in developmental disabilities, 31*(5), 985-994.

Figueiredo, M., Mafalda, R., & Kamensky, A. (2018). Virtual reality as an educational tool for elementary school. Interdisciplinary Conference on Innovation, Desgin, Entrepreneurship, And Sustainable Systems (pp. 261-267).

Ford-Gilboe, M., Varcoe, C., Scott-Storey, K., Perrin, N., Wuest, J., Wathen, C. N., ... & Glass, N. (2020). Longitudinal impacts of an online safety and health intervention for women experiencing intimate partner violence: randomized controlled trial. *BMC public health, 20*(1), 1-17.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science, 345*(6203), 1502-1505.

Freina, L., & Ott, M. (2015, April). A literature review on immersive virtual reality in education: state of the art and perspectives. In *The international scientific conference elearning and software for education, 1*(133), 10-1007.

Giannakakos, A. R., Vladescu, J. C., Reeve, K. F., Kisamore, A. N., Fienup, D. M., & Carrow, J. N. (2021). Using behavioral skills training and equivalence-based instruction to teach children safe responding to dangerous stimuli: A proof of concept. *The Psychological Record, 71*(1), 119-131.

Gorard, S. (2021). How to make sense of statistics. p. 94.

Gordillo, A., Barra, E., López-Pernas, S., & Quemada, J. (2021). Development of Teacher Digital Competence in the Area of E-Safety through Educational Video Games. *Sustainability 2021*(13), 8485.

Grissom, J. A., & Loeb, S. (2011). Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. *American Educational Research Journal, 48*(5), 1091-1123.

Gunby, K. V., Carr, J. E., & Leblanc, L. A. (2010). Teaching abduction-prevention skills to children with autism. *Journal of Applied Behavior Analysis, 43*(1), 107-112.

Houvouras IV, A. J., & Harvey, M. T. (2014). Establishing fire safety skills using behavioral skills training. *Journal of applied behavior analysis, 47*(2), 420-424.

Hwang, J. I. (2015). What are hospital nurses' strengths and weaknesses in patient safety competence? Findings from three Korean hospitals. *International Journal for Quality in Health Care, 27*(3), 232-238

Ikeda, E., Mavoa, S., Cavadino, A., Carroll, P., Hinckson, E., Witten, K., & Smith, M. (2020). Keeping kids safe for active travel to school: A mixed method examination of school policies and practices and children's school travel behaviour. *Travel behaviour and society, 21*, 57-68.

Islam, S. (2022). Ensuring risk-free digital-banking in us-economy: Application of akim's model.

Jones, C. M., & Pozzebon, K. (2010). Being Safety Smart:

Social issue game for child protective behaviour training.

Jones, C., Scholes, L., Rolfe, B., & Stieler-Hunt, C. (2020). A serious-game for child sexual abuse prevention: an evaluation of orbit. Child Abuse & Neglect, 107, 104569. Jostad, C. M., Miltenberger, R. G., Kelso, P., & Knudson, P. (2008). Peer tutoring to prevent firearm play: Acquisition, generalization, and long-term maintenance of safety skills. *Journal of Applied Behavior Analysis, 41*(1), 117-123.

Joubert, I. J., Fraser, W. J., & Sentsho, M. L. (2013). Road Safety Education: A paradoxical state for children in a rural primary school in South Africa. *Journal of Asian and African studies, 48*(2), 209-228.

Kardefelt-Winther, D., Heeren, A., Schimmenti, A., van Rooij, A., Maurage, P., Carras, M & Billieux, J. (2017). How can we conceptualize behavioural addiction without pathologizing common behaviours? *Addiction, 112*(10), 1709-1715.

Kolko, D. J. (2001). Efficacy of cognitive-behavioral treatment and fire safety education for children who set fires: Initial and follow-up outcomes. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(3), 359-369.

Kritzinger, E. (2015). Enhancing cyber safety awareness among school children in South Africa through gaming. In 2015 Science and Information Conference (SAI) (pp. 1243-1248). IEEE.

Kulkofsky, S., & Perez, C. O. (2021). Factors affecting the reliability of children's forensic reports. Current Issues in Memory: *Memory Research in the Public Interest, 104*.

Lambie, I., & Randell, I. (2011). Creating a firestorm: A review of children who deliberately light fires. *Clinical Psychology Review, 31*(3), 307-327.

Lazarinis, F., Alexandri, K., Panagiotakopoulos, C., & Verykios, V. S. (2020). Sensitizing young children on internet addiction and online safety risks through storytelling in a mobile application. *Education and Information Technologies, 25*(1), 163-174.

Ledbetter-Cho, K., Lang, R., Lee, A., Murphy, C., Davenport, K., Kirkpatrick, M & O'Reilly, M. (2021). Teaching children with autism abduction-prevention skills may result in overgeneralization of the target response. *Behavior Modification, 45*(3), 438-461.

Lehna, C., Janes, E. G., Rengers, S., Graviss, J., Scrivener, D., Knabel, T & Myers, J. (2014). Community partnership to promote home fire safety in children with special needs. *Burns, 40*(6), 1179-1184.

Lewandowsky, S., & Van Der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology, 32*(2), 348-384.

McComas, J., MacKay, M., & Pivik, J. (2002). Effectiveness of virtual reality for teaching pedestrian safety. *Cyberpsychology & Behavior, 5*(3), 185-190.

Morrongiello, B. A., Sandomierski, M., Zdzieborski, D., & McCollam, H. (2012). A randomized controlled trial evaluating the impact of the Supervising for Home Safety program on parent appraisals of injury risk and need to actively supervise. *Health psychology, 31*(5), 601.

Müller, A. R., Röder, M., & Fingerle, M. (2014). Child sexual abuse prevention goes online: Introducing "Cool and Safe" and its effects. Computers & Education, 78, 6065.

Nicolaidou, I., & Venizelou, A. (2020). Improving children's E-safety skills through an interactive learning environment: A quasi-experimental study. *Multimodal Technologies and Interaction, 4*(2), 10.

Norkhairani, A. R., Azilawati, R., Maizan, M. A., & Wan Malini, W. I. (2014). Teaching Dzikir Through 2D Games.

Nkuruho, T. F., Isingoma, C., & Senserrick, T. (2021). School road safety education in Uganda: progress and lessons learned. *Journal of road safety, 32*(1), 45-51.

Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic bulletin & review, 26*(5), 1596-1618.

Oliva, D., Somerkoski, B., Tarkkanen, K., Lehto, A., & Luimula, M. (2019, April). Virtual reality as a communication tool for fire safety-Experiences from the VirPa project. In GamiFIN (pp. 241-252).

Poche, C., Yoder, P., & Miltenberger, R. (1988). Teaching self-protection to children using television techniques. Journal of Applied Behavior Analysis, 21(3), 253-261. Purcell, C., & Romijn, A. R. (2020). Teaching Children Road Safety Using a Simulated Environment. *Journal of Education and Educational Development, 7*(1), 44-54.

Raftery, S. J., & Wundersitz, L. N. (2011). The efficacy of road safety education in schools: A review of current approaches. *Criminology, 50*, 88-100.

Rawi, N. A., Mamat, A. R., Deris, M. S. M., Amin, M. M., & Rahim, N. (2015). A Novel multimedia interactive application to support road safety education among Primary School children in Malaysia. *Jurnal Teknologi, 77*(19).

Rossi, M. R., Vladescu, J. C., Reeve, K. F., & Gross, A. C. (2017). Teaching safety responding to children with autism spectrum disorder. *Education and Treatment of Children, 40*(2), 187-208.

Romanova, I. N., Morkovkin, D. E., Nezamaikin, V. N., Gibadullin, A. A., & Ivanova, M. A. (2020). Formation of a policy to ensure environmental safety in modern economic conditions. In IOP conference series: materials science and engineering (Vol. 734, No. 1, p. 012166). IOP Publishing.

Salfarina, A. (2021). Cultivating the Ideal Teamwork through Real Software Development Project Experience. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(3), 1923-1929.

Sisson, L. A., & Van Hasselt, V. B. (1984). Emergency fire-

safety skills for blind children and adolescents: Group training and generalization. *Behavior modification, 8*(2), 267-286.

Smith, S., & Ericson, E. (2009). Using immersive game-based virtual reality to teach fire-safety skills to children. *Virtual reality, 13*(2), 87-99.

Tarasenko, M. A., Miltenberger, R. G., Brower-Breitwieser, C., & Bosch, A. (2010). Evaluation of peer training for teaching abduction prevention skills. *Child & Family Behavior Therapy, 32*(3), 219-230.

Treviño-Siller, S., Pacheco-Magaña, L. E., Bonilla-Fernández, P., Rueda-Neria, C., & Arenas-Monreal, L. (2017). An educational intervention in road safety among children and teenagers in Mexico. *Traffic injury prevention, 18*(2), 164-170.

Vankov, D., & Jankovszky, D. (2021). Effects of using headset-delivered virtual reality in road safety research: A systematic review of empirical studies. *Virtual Reality & Intelligent Hardware, 3*(5), 351-368.

Vanselow, N. R., & Hanley, G. P. (2014). An evaluation of computerized behavioral skills training to teach safety skills to young children. *Journal of Applied Behavior Analysis, 47*(1), 51-69.

Varcoe, C., Scott-Storey, K., Perrin, N., Wuest, J., Wathen, C. N., ... & Glass, N. (2020). Longitudinal impacts of an online safety and health intervention for women experiencing intimate partner violence: randomized controlled trial. *BMC public health, 20*(1), 1-17.

Weatherley, R., Hajar, A. S., Noralina, O., John, M., Preusser, N., & Yong, M. (2012). Evaluation of a school-based sexual abuse prevention curriculum in Malaysia. *Children and Youth Services Review, 34*(1), 119-125.

White, C., Shanley, D. C., Zimmer-Gembeck, M. J., Walsh, K., Hawkins, R., Lines, K., & Webb, H. (2018a). Promoting young children's interpersonal safety knowledge, intentions, confidence, and protective behavior skills: Outcomes of a randomized controlled trial. *Child Abuse & Neglect, 82*, 144-155.

White, C., Shanley, D. C., Zimmer-Gembeck, M. J., Walsh, K., Hawkins, R., & Lines, K. (2018b). "Tell, tell, tell again": The prevalence and correlates of young children's response to and disclosure of an in-vivo lure from a stranger. *Child Abuse & Neglect, 82*, 134-143.

White, C., Shanley, D. C., Zimmer-Gembeck, M. J., Walsh, K., Hawkins, R., & Lines, K. (2018c). Outcomes of in situ training for disclosure as a standalone and a booster to a child protective behaviors education program. *Child maltreatment, 24*(2), 193-202.

Wollscheid, S., Sjaastad, J., & Tømte, C. (2016). The impact of digital devices vs. Pen (cil) and paper on primary school students' writing skills–A research review. *Computers & education, 95*, 19-35.

Wright, C. Y., Albers, P. N., Oosthuizen, M. A., & Phala, N. (2014). Self-reported sun-related knowledge, attitudes and behaviours among schoolchildren attending S outh A frican primary schools. *Photodermatology, Photoimmunology & Photomedicine, 30*(5), 266-276.

Zahed, A. H., Hussain, M., Farooq, M., Riaz, U., & Alam, T. M. (2019, November). A journey of WEB and Blockchain towards the Industry 4.0: An Overview. In 2019 International Conference on Innovative Computing (ICIC) (pp. 1-7). IEEE.

Zhang-Kennedy, L., & Chiasson, S. (2021). A systematic review of multimedia tools for cybersecurity awareness and education. *ACM Computing Surveys (CSUR), 54*(1), 1-39.

Zeedyk, M. S., Wallace, L., Carcary, B., Jones, K., & Larter, K. (2001). Children and road safety: Increasing knowledge does not improve behaviour. *British journal of educational psychology, 71*(4), 573-594.

Zinicola, J. (2021). Behavioral Skills Training to Teach Online Safety Responses to Youth with Autism Spectrum Disorder.

Zinkus, M., Curry, O., Moore, M., Peterson, Z., & Wood, Z. J. (2019). Fakesbook: A social networking platform for teaching security and privacy concepts to secondary school students. In Proceedings of the 50th ACM Technical Symposium on Computer Science Education (pp. 892-898).