Smartphones demonstrate growing popularity as a delivery mechanism for pediatric interventions for multiple reasons. First, smartphone ownership is growing exponentially in the United States (Pew Research Center, 2018) and internationally (Poushter, 2016). Many now rely on smartphones as their sole source of Internet access (Smith, 2015), particularly users who are younger, minorities, and of lower socioeconomic status (Pew Research Center, 2018). Smartphones therefore promote access to large numbers of diverse groups, and do not require users to have costly broadband and/or computer access nor to take the time to physically sit down in front of a computer to complete intervention tasks. Second, smartphones tend to be carried throughout the day, promoting in-the-moment and ongoing opportunities for assessment and intervention (Proudfoot, 2013). Smartphones therefore also represent a delivery mechanism that is already highly used in daily life, promoting the likelihood of higher intervention use, and potentially overcome barriers to traditional delivery mechanisms, such as forgetfulness (Mohr et al., 2010). However, while smartphones are growing in popularity as daily tools and as a delivery mechanism for evidence-based pediatric interventions (Fedele, Cushing, Fritz, Amaro, & Ortega, 2017), concerns are present in the mobile health (mHealth) literature around their efficacy for specific pediatric populations (Grist, Porter, & Stallard, 2017).

The present work by Schwebel et al. (2017) used a nonrandomized trial methodology to evaluate a smartphone-based virtual reality (VR) intervention to increase pedestrian safety in 56 schoolchildren in China (Schwebel et al., 2017). This work contributes to the literature on smartphones as a delivery mechanism and provides an opportunity to consider current practices in mHealth evaluations, including (1) strengths, (2) limitations, and (3) usability testing as a means to increase generalizability.

This study has multiple strengths. First, the methodology is described clearly, with enough detail provided around the design to allow for replicability. Unfortunately, likely to comply with journal constraints in manuscript length, studies often sacrifice clear descriptions of the design. This gap in describing a critical methodological aspect of mHealth interventions typically limits generalizability and insights into the potential mechanisms impacting outcomes. Second, the authors used multimethod outcomes, such that pre/post comparisons could be made with both coded behavioral observations and child self-report. This is unfortunately only applied to roughly half of the sample (25 and 24 children’s actual pedestrian behavior was captured preintervention and post-intervention, respectively; it is unclear if these children differ in any way from those not captured) but bolsters the significant findings for changes in self-reported data. Finally, the authors highlight that their intervention has minimal cultural specificity. While slight design modifications would be necessary for deployment in other countries and cultures (e.g., language translation), the majority of the described design appears to allow for evaluation and implementation for schoolchildren outside of China. This is an important
strength, both for this study and mHealth interventions, from the perspective of delivering global health interventions to pediatric populations, particularly in resource challenged areas.

While it is important to note that this study is a pilot intervention, there are also several limitations of the current work that highlight targets for improvement. The authors note: (1) the cost of equipment (i.e., $160 US per unit); (2) the possible introduction of bias (i.e., participants wore identifiable buttons and bright colors when crossing the street); (3) protocol gaps (i.e., managing participant excitement and attention when introducing the intervention); (4) lack of a control group; and (5) no long-term follow-up data (Schwebel et al., 2017). However, other issues central to mHealth interventions also warrant discussion. These include the clinical significance of findings (e.g., at postintervention, ~80% of children still selected “Maybe” or “No” to a prompt assessing if they feel safe crossing a busy street) and the usability of the system.

Usability indicates how intuitive and easy to use a technology may be; usability testing is a systematic, formative, or summative evaluation to improve design (Tullis & Albert, 2008). Given that the current study was a pilot, the authors may have been constrained in their ability to measure usability or to report usability given their focus on the intervention. Previous usability testing of the VR system occurred with 68 adults, who rated the realism and their level of motion sickness (Schwebel, Severson, & He, 2017). However, usability testing guidelines specific to children have been developed and implemented (Hanna, Risden, & Alexander, 1997), and typically target five specific attributes: learnability, efficiency, memorability, errors, and satisfaction (Nielsen, 1993). Given the authors’ aim to implement learned skills for successful pedestrian behaviors, opportunities to evaluate usability metrics would improve generalizability. Indeed, usability testing of attributes such as learnability (i.e., to evaluate how well users may learn an intervention skill through use of an app) have begun to be increasingly implemented before trials or public deployment with adults (Sarkar, et al., 2016, Stiles-Shields et al., 2017, Vilardaga et al., 2018). Satisfaction is another critical item to evaluate. If users are not satisfied with a technology, they are unlikely to use it. The authors identified that the children found the system to be “fun” and “educational.” However, determining satisfaction from the perspective of the children, teachers, and parents would better speak to the generalizability of this intervention to schools or families. Future trials limited in their ability to conduct formal usability testing may also consider evaluating metrics of usability via a variety of validated self-report questionnaires (e.g., System Usability Scale; Brooke, 1996).

The assessment of usability is a critical element of interpreting how well a smartphone-delivered intervention performs during interactions with intended users. Therefore, usability testing improves the claims of feasibility, efficacy, and generalizability of mHealth findings.

The present work stands as a positive representation of evaluating a smartphone-delivered intervention to improve a crucial public health target for global pediatric populations. However, it also highlights opportunities for improving generalizability. For example, just as a review of the training, competency, and expertise of therapists delivering a face-to-face intervention are commonplace in interpreting the merit of a traditionally delivered intervention, usability findings should be provided in all reports in future mHealth literature.

Funding

This research was supported in part by grants from the National Institute of Nursing Research and the Office of Behavioral and Social Sciences Research (grant number R01 NR016235), National Institute of Child Health and Human Development (grant number R01 HD048629), and the March of Dimes Birth Defects Foundation (grant number 12-FY13-271).

Conflicts of interest: None declared.

References


