Behavioural Public Policy (2025), 1–20
doi:10.1017/bpp.2024.29

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Preventing risks by experiencing them: effects on knowledge and behaviour

Rob Bulterman[1], Thijs Verwijmeren[1], Rob W. Holland[1] and Harm Veling[1,2]

[1]Behavioral Science Institute, Radboud University, Nijmegen, The Netherlands and [2]Consumption and Healthy Lifestyles, Wageningen University & Research, Wageningen, The Netherlands
Corresponding author: Thijs Verwijmeren, email: thijs.verwijmeren@ru.nl.

(Received 1 December 2023; revised 13 May 2024; accepted 21 May 2024)

Abstract
Safety villages are interventions that aim to boost children's knowledge and behaviour regarding risk-taking behaviours and their consequences via an experiential learning approach. In safety villages, children experience scenarios involving risks that resemble real-life situations. We investigated the extent to which desirable learning outcomes from a single-session safety village visit are visible outside the safety village context. In a well-powered quasi-experimental preregistered field study, we compared students (aged 11–13) who received experiential safety education to a control group of students who had not yet received the education on three important learning outcomes: Knowledge-application, risk-taking behaviour and general risk-taking tendencies. Data were collected outside of the safety village environment, before or after the visit, and without explicit reminders of the visit. Results show students who received experiential safety education outperformed those who did not yet receive experiential education on knowledge-application and reduced risk-taking behaviours. We found no differences on general risk-taking tendencies. These results show a single visit to a safety village visit can reduce risk-taking of risks that were experienced in the village, but not general risk-taking tendencies. Theoretical and policy implications are discussed.

Keywords: experiential learning; risk-taking behaviour; safety education; adolescents; field study

Introduction

Children aged 12–18 have long been associated with increased risk-taking and its negative consequences (Jessor, 1998; Finucane *et al.*, 2000; Crone *et al.*, 2016). Risk-taking behaviours that are especially prevalent among adolescents are cyber-related intimidation violations, aggression and property crime (Willoughby *et al.*, 2021). In the Netherlands, risk-taking while partaking in traffic is very common and is especially prevalent among younger people. For instance, in 2021, 77% of children aged 12–14 occasionally used their phone while cycling in traffic, and this percentage declined across older age groups (van der Kint and Mons, 2021). Educating children on safety before this period thus seems an important opportunity towards preventing risk-related behaviours. However, risk education is challenging. First, informing people, and children in particular, about risk behaviours is often insufficient to prevent these behaviours (Cook and Bellis, 2001; Zeedyk *et al.*, 2001). Second, safety and health do not always have a structural place in school systems. For example, in most countries in Europe, children learn about basic traffic rules at elementary school (Mütze and De Dobbeleer, 2019), but there is no clear embedded teaching about important topics such as social safety, cyber safety or mental health (Eurydice, n.d.).

To address these challenges, a relatively recent development in public policy is to boost children's knowledge and skills in experiential education programmes called safety villages (Thomson, 2006). These programmes provide children with interactive scenarios in which they experience real-life risky situations in a controlled environment. For example, to learn potential risks when using the internet, children engage in a variety of online tasks like creating a social media profile or reacting to people's online content and experience the (potentially negative) consequences of their actions. Learning is further facilitated by reflecting on the experience, together with peers and a guide. Through this reflection, students gain insight into what went well and what did not, and learn about the knowledge and skills needed to adequately address the situation in real life. Therefore, safety villages can be considered a boosting approach to influencing students' behaviour (Hertwig and Grüne-Yanoff, 2017), as it aims to improve students' decision-making in potentially risky situations in real life.

The theoretical foundation for safety villages comes from experiential learning theory (Kolb, 1984), which states that learning is 'the process whereby knowledge is created through the transformation of experience' (p. 38). Going through and reflecting on concrete hands-on experiences is theorised to help learners connect the presented information to real-world situations (for a theoretical background on experiential learning, see Kolb (2015) and Morris (2019)). The main critique on experiential learning, however, is the lack of empirical evidence (Bergsteiner *et al.*, 2010; Morris, 2019). This critique seems especially relevant for safety villages. That is, even though policymakers invest in experiential (safety) education due to high face validity and positive reactions by students and teachers alike (Thomson, 2006; James and Williams, 2017), we only found four published empirical articles on the effectiveness of safety villages in a span of nearly three decades of literature (Gielen *et al.*, 1996; Lamb *et al.*, 2006; Teyhan *et al.*, 2016; Boam and Pulford, 2019; for an unpublished government report, see Oxford Evaluation Team (2003)).

The modest amount of existing studies do offer some valuable first insights on the effectiveness of experiential safety education. Boam and Pulford (2019) showed that safety villages have the potential to increase risk-perception. Students rated potentially dangerous behaviours that were discussed in the safety village as more risky after a safety village visit, compared to before. Gielen *et al.* (1996) showed that students, compared to a pre-test, scored better on knowledge tests about the specific information addressed in the safety village, both directly after a visit and after a delayed period of several months. Lamb *et al.* (2006) replicated these within-participant findings and added a control group (who had not received safety village education), showing that students who received safety village education also outperformed students who had not received safety village education in terms of knowledge recollection. In addition, they measured safety behaviours in hypothetical re-enactments inside the safety village. Results showed that, both directly after a visit and in a follow-up visit after 3 months, students who visited the safety village showed less risky behaviours compared with the control group that had not received safety village education.

Despite existing research showing promising results, it is still unclear how well experiential safety education outcomes generalise to other situations. We assume that public policymakers who implement these programmes aim to help students deal with risky situations in real life, outside of the safety village context. Therefore, it is imperative that the potential outcomes of such programmes generalise to other situations and times, particularly since safety villages typically rely on one visit from the target group. However, the aforementioned results are found in situations specific to the safety village context. Concretely, either the exact knowledge addressed in the safety village was tested, or behaviour was measured inside the safety village. According to research on memory, seeing the same content or revisiting a location can serve as a cue to remind students of what they have learned before (Smith, 2013; Robin and Moscovitch, 2014). This can even happen without a deliberate attempt to remember what they experienced (Congleton *et al.*, 2021). As such, coming back to the safety village for measurements or seeing the exact same items might have given students an explicit reminder about their safety village education, possibly influencing some of the found effects. This raises questions about whether students can demonstrate the correct knowledge and behaviour in situations where they are not explicitly reminded of the safety village education. To contribute to a better understanding of the potential of safety villages in boosting knowledge and behaviour, it is thus necessary to investigate potential outcomes in situations without explicit reminders of the safety village material and context.

Previous research that is suggestive of the possibility for generalisation to other situations is a study Teyhan and colleagues conducted in 2016. Data from the Avon Longitudinal Study of Parents and Children (ALSPAC; Boyd *et al.*, 2013) was used to compare high school students on a number of self-reported risk behaviours (e.g., related to road safety, substance use and hospital attendance). These items were not designed to measure the specific content of the safety village education. Although the study did not find that effects were different for students who visited a safety village for most behaviours, it did show that students who visited a safety village during elementary school reported safer use of pedestrian crossings and less recent smoking compared to students who did not visit a safety village. This study points to the possibility that a visit to a safety village may indeed impact risk behaviours, but the study has two important limitations. First, the self-reported measurements used in the study may have suffered from difficulties with accurate retrospective memory or from social desirability, which especially occurs in research on safety behaviour (Wåhlberg *et al.*, 2010; Keiser and Payne, 2019). Second, the control group consisted of students from schools who did not schedule a visit to the safety village when these students were in elementary school. Schools who schedule a visit might find safety education more important than the schools who do not

https://doi.org/10.1017/bpp.2024.29 Published online by Cambridge University Press

visit, making it difficult to assess whether the found differences were indeed due to generalisation of the experiential education or due to differences between schools that visit or do not visit the villages (e.g., other actions schools might have taken in terms of safety education).

In the present study, we aim to investigate the generalisation of effects of a single visit to a safety village on three outcomes: Knowledge-application, behaviour and general risk-taking tendencies. Furthermore, in contrast with previous research (Gielen *et al.*, 1996; Lamb *et al.*, 2006), we aim to examine to what extent lessons learned in the safety village, generalise to a broader context. First, we tested these outcomes outside of the safety village context (i.e., at their own schools, without mention of the safety village). Second, we included knowledge and behavioural items that were not learned explicitly in the safety village. Even more so, the general risk-taking measure was completely new to our participants.

In a collaboration with safety village *Risk Factory*,[1] school classes (both urban and rural) were recruited to participate in the current study. This safety village is aimed at students age 11–13, who are in their final year of Dutch primary school. A visit consists of a set of 6–8 scenarios in which students learn about safety and health in different domains. A typical visit takes around 3 h. We recruited classes that were either scheduled to visit this safety village and classes that had already visited the village. This way, we aimed to ensure participating classes were relatively similar in how much safety education is on their agenda, something that was uncertain in the previously mentioned study on generalisation of safety village education (Teyhan *et al.*, 2016). As such, participating students had either recently received experiential education on safety and health at Risk Factory, or would receive it in the near future. Measurements took place in class in the form of computer assignments, outside of the safety village context, without explicit reminders of their (upcoming) visit to Risk Factory.

We formulated three preregistered hypotheses.[2] First, we expected that students who visited Risk Factory would show better knowledge-application than students who did not visit Risk Factory, indicated by better performance on two assignments. One of these assignments measured how well students assessed the riskiness of various requests by strangers (hypothesis 1a). The other assignment measured how well students could apply their knowledge of Positive Health[3] (hypothesis 1b). Second, we expected students to show safer behaviour outside the Risk Factory environment when they had visited the safety village compared to when they had not (hypothesis 2). We measured two specific safety behaviours: the number of unnecessary personal details shared during an online sign-up (hypothesis 2a) and the number of pictures with personal identifiers shared without permission (hypothesis 2b).

[1]For an impression, see https://www.riskfactorylimburgnoord.nl/ (website in Dutch).

[2]The order in which we present the hypotheses in this article differs from the order of the hypotheses in the preregistration, with the purpose of benefitting the structure of the article. Therefore, the numbers used in this article to indicate the hypotheses (e.g., hypothesis 1a, 1b, etc.) do not match the numbers used in the preregistration.

[3]Positive Health is a concept of health students learn about during a Risk Factory scenario. Students learn that health has more aspects than just a physical aspect. For a detailed description of the concept, see Huber *et al.* (2016).

Third, we expected that students who visited Risk Factory would show less risk-taking tendencies in general, represented by less risk-taking during the Balloon Analog Risk Task (BART; Lejuez *et al.*, 2002), compared to students who had not visited Risk Factory (hypothesis 3).

Method

Design and participants

As classes visited Risk Factory in their final year of Dutch elementary school, participants were aged 11–13. Sample size depended on the number of classes that visited Risk Factory during April, May and June 2023. As such, 325 students[4] from 16 classes that visited Risk Factory during this period were recruited for the experimental group, for which data collection took place an average of 16.25 days after a visit (range = 11–28 days). Classes that had not yet visited Risk Factory, but would do so later that month, were recruited until the amount of participants from the experimental group was roughly matched. In the end, the control condition consisted of 329 students from 13 classes, resulting in a total sample size of 654 students (339 females, 304 males, 11 non-binary; AR vs BR: 162 vs 177 females, 155 vs 149 males, and 8 vs 3 non-binary).

Because the sample size was based on resource constraints, we developed measurements with the goal to have multiple datapoints (i.e., measurement items) for each hypothesis within each participant to increase reliability of the measurements and hence statistical power. The lowest number of datapoints for a hypothesis test was 4,324 (616 participants * 7 measurement points, for hypothesis 2a), which according to Westfall *et al.* (2014) should give us sufficient power (>0.90) to detect even very small effect sizes ($d = 0.1$). Table 1 shows the number of datapoints for each hypothesis test.

Measurements

Students participated in five assignments. These assignments were pilot-tested, first with one student and subsequently with a class of 25 students. During this pilot test, we measured whether students understood the instructions and whether students evaluations of the used stimuli were as intended (e.g., whether pictures we estimated were interesting for elementary students were indeed rated as interesting). Students who participated in the pilot had not yet visited Risk Factory and did not participate in the final study. Please note that the development and specific content of the experiential scenarios at Risk Factory is beyond the scope of this article. The focus of this article will be on the measurements that were developed to test the outcomes of these scenarios. For further descriptions of what students experience in the respective scenarios when visiting Risk Factory, see the Supplementary material.

[4]We preregistered 295 students during this period. The eventual number of students is larger, because we got a positive response from another school afterwards and we thought it best to include them to increase power.

Table 1. Total amount of datapoints tested per hypothesis, as well as per condition (before Risk Factory [BR] vs after Risk Factory [AR])

	N BR	*N* AR	Total *N*
Hypothesis 1: Knowledge-application			
a) AR rates target requests as more risky than BR	2,296 requests	2,275 requests	4,571 requests
b) AR rates target questions as more important for health than BR	2,289 questions	2,275 questions	4,564 questions
Hypothesis 2: Topic-specific behaviour			
a) AR shares less details than BR	2,155 details	2,169 details	4,324 details
b) AR shares less target pictures than BR	2,624 pictures	2,600 pictures	5,224 pictures
Hypothesis 3: General risk-taking behaviour			
AR pumps less on non-burst balloons than BR	4,244 balloons1	5,395 balloons	9,639 balloons

1The difference in the amount of balloons between conditions is due to the BART not working on one school (three classes) during data collection due to technical issues. Because of planning (summer holiday for schools), we opted not to collect extra data as this would result in a delay of several months. We expect to have sufficient power still due to the high amount of balloons per condition still present.

Knowledge-application

To test the knowledge-application hypothesis, we developed two assignments based on two separate Risk Factory scenarios. The first measured how adequately students assess potentially risky requests. At Risk Factory, students learn that criminals use a strategy to recruit youth into their criminal network by letting them perform illegal acts unknowingly. In this, seemingly harmless people (often peers) make a seemingly harmless request, like dropping off a package, and offer a reward for doing so. However, such a package typically contains illegal products and thus delivering it is an illegal act. Delivering the package is used as blackmail to pressure children in performing more criminal acts. Any credible behavioural measure for this scenario (actually letting students partake in illegal acts) would have severe ethical downsides. Therefore, we asked participants to imagine themselves in a hypothetical situation and measured whether students were able to apply the knowledge learned in the scenario to this hypothetical situation.

In the hypothetical situation, participants met two children, aged 14, called Noah and Jaimy (two common names for Dutch children of that age) who ask them for help. In total, participants would see 21 requests that these children could ask them. These included seven target items that seemed not risky, but could be used by criminals to lure teenagers into criminal activities. At Risk Factory, students only learn about delivering a package, but the seven target requests contained other requests used by criminals as well (e.g., posting an envelope or handing off a sealed bag to someone). This way, we aimed to assess whether knowledge about packages would not only be recollected, but also applied to similar requests. The

https://doi.org/10.1017/bpp.2024.29 Published online by Cambridge University Press

other 14 items were filler items, 7 of which were pilot-tested as risky (e.g., stealing a bike) and 7 that were pilot-tested as not risky (e.g., lending out a pen). The goal of the filler items was to rule out that students simply rated everything as more risky, instead of adequately applying the knowledge learned in the safety village to relevant situations. Ideally, we would only find a difference between groups for the target items, but not for clearly risky and non-risky items. The order of the requests was randomised per participant. For each request, students rated how risky it was to engage in it on a scale from 1 (not risky at all) to 10 (very risky).

The second measurement for knowledge-application was about positive health: A definition of health that states that health has not only a physical dimension but consists of a total of six dimensions (for an overview, see Huber *et al.* (2016)). At Risk Factory, students learn about this broader concept of health by a series of interactive assignments and sharing of experiences regarding all six dimensions of positive health. The goal of the scenario is to let students realise that, apart from your physical state, things like 'being yourself' or 'having meaningful friends' can also be very important for someone's health.

Participants were instructed to rate a series of 21 questions. Their task was not to answer these questions as if someone was asking them, but to indicate whether the questions were important to ask someone else if they wanted to know more about that person's health. They indicated this by rating the questions on a scale from 1 (not important at all) to 10 (very important). This way, we aimed to test if students were able to apply the knowledge they learned at Risk Factory (that health is more than physical) to their current perception of health. The seven target items were questions related to one of the five non-physical dimensions of positive health, such as 'Do you feel like you can be yourself most of the time?'. These questions were not specifically asked at Risk Factory, but were related to the content addressed. In addition, there were 14 filler items. Seven of these were questions that were important to ask if you want to know about the physical aspects of someone's health (e.g., 'Do you exercise a lot?'). The other seven were questions that were not important to ask if you want to know something about someone's health (e.g., 'What is your favourite T-shirt?'). This way, we aimed to explore if students correctly applied their knowledge instead of finding every question important for health, even when they were not. The order of the questions was randomised per participant.

Behaviour

We developed two assignments to test the second hypothesis, about the effects of experiential safety education on behaviour. The first assignment measured whether students unnecessarily shared their personal details. This was briefly addressed at Risk Factory during a scenario on online safety. More specifically, we were interested in whether students would share personal details in an online sign-up situation outside the safety village environment. This measurement was not framed as an assignment, but as a prerequisite to start working on the assignments. Instructions told participants to 'Please sign up so you can continue with the assignments'. They saw 11 text fields, each with a header asking for a specific (personal) detail. Four of these details were non-personal and mandatory to fill in, to obtain anonymous identifiers for all participants: school, class, gender and a participant number received at the start of the assignments. The seven target items were personal details that were not mandatory to fill in: first name, last name, age, place of residence, address (two fields, one for street and one for home number) and postal code. Text field headers either had a *-sign at the end or not and the sentence 'Fields with a * are mandatory' indicated which details were mandatory.

We measured whether participants shared a target detail (yes/no) by saving the amount of characters entered in the respective text field. We did not collect nor save actual answers for privacy reasons. If the amount of characters was zero, the score for that item was 'no'. If the amount of characters was larger than zero, it counted as 'yes'. We argued that some students might fill in non-sensical answers if they would rather not share a detail, instead of leaving the field blank. Therefore, to gain some extra control over the validity of the measurement, we added a measurement-specific exclusion criterium. This criterium was based on two fields that should always have a fixed amount of characters if you fill them in seriously, namely age (two characters, as age for this target group should be 10–13 years old) and postal code (always six or seven characters in the Netherlands). If one of these fields was not filled in correctly, we excluded all seven answers from that participant during data analysis on this specific measurement.

The second behavioural measurement revolved around the sharing of pictures. During a scenario at Risk Factory, students learn to not take pictures of situations requiring emergency services that show recognisable people or other identifiers, such as a face or a licence plate. As it is difficult to measure taking pictures in real life, we decided to measure whether students instead share pictures of such situations that contain objective identifiers. Importantly, using this slightly different behavioural outcome also gives us more insights in the extend of behavioural generalisation. In other words, when students have experienced that taking pictures could have negative consequences, do they understand that sharing them should be discouraged as well and act accordingly?

In this assignment, participants created their own 'feed' on which they could share pictures, similar to some social media. Participants saw 24 pictures sequentially and could indicate whether they wanted to share the picture on their feed (yes/no). The pictures were pilot-tested on how interesting they were. The target items consisted of eight interesting pictures of emergency situations (e.g., traffic accident or fire) containing objective identifiers, which should not be shared.[5] The other items were eight non-safety-related fillers that were interesting (e.g., amusement parks or Minecraft) and eight that were not interesting (e.g., a pile of bricks or a stapler). The goal of the filler items was to mask the goal of the assignment and to check if students would behave logically (i.e., share interesting pictures and not share uninteresting ones). The order of the pictures was randomised between participants.

[5]The images were evaluated by professionals from the safety village in terms of whether they indeed should not be shared, and would be suitable to show to children (e.g., an ambulance with sirens). Images that contained aversive graphical content, like visible injuries or violence, were not included. See the Supplementary material for a description of the pictures. The authors can be contacted for further details about the scenarios.

To make the sharing of pictures consequential, instructions informed students that one of the created feeds would be shown in class and all students could then indicate how interesting they found the feed. This way, we tried to create a situation where a student would behave in a way that roughly resembled a real-life social media environment, in which participants had the motivation to create an interesting feed.

General risk-taking tendencies

To test the hypothesis on general risk-taking tendencies, participants performed a version of the Balloon Analogue Risk Task (BART; Lejuez *et al.*, 2002). This task is widely used in psychological research to measure risk-taking tendencies (Lauriola *et al.*, 2013) with high reliability (White, Lejuez and de Wit, 2008). In terms of validity, research has demonstrated that the BART is positively related to a variety of risk behaviours during adolescence (Lejuez *et al.*, 2003, 2007; Aklin *et al.*, 2005; Hopko *et al.*, 2006; MacLean *et al.*, 2018), so changes in risk-taking tendencies on the BART due to safety village education could be an important indicator for generalisation to risk behaviours in real life.

In this task, 30 balloons were presented sequentially. Participants pumped up these balloons using the spacebar. Every pump slightly increased the size of the balloon and the value of the balloon with 5 cents. However, with every pump, the balloon could burst resulting in losing the money saved up for that balloon. Participants did not know when a balloon would burst, so every extra pump increases the risk of bursting. A balloon would burst after a number of pumps based on a randomly generated sequence, where balloon 1 would burst after X pumps, balloon 2 after Y pumps, etc. This sequence was kept the same for all participants (balloon 1 always burst after X and balloon 2 always after Y) to decrease error variance between participants due to different procedures (such as a first balloon bursting very quickly for one participant, reducing their subsequent risk-taking). Participants were informed that the person with the most money would win a small prize, without mentioning what the prize was (a bag of modelling balloons). Risk-taking was measured by the average number of pumps on non-burst balloons (in line with Lejuez *et al.*, 2002).

Procedure

To collect data, a researcher paid a visit to participating classes at school. To prevent explicit reminders of students' former or future experiences at Risk Factory, teachers informed students that the researcher's main interest was developing assignments for Dutch grade 8 elementary students, but made no explicit link with Risk Factory. At the start of the visit, the researcher would confer with the teacher in private to check if the visit was indeed not linked to Risk Factory beforehand. This was the case for all classes. Before data collection, the researcher asked students to do the assignments honestly and with their full attention. After this, students individually took the assignments.

To further dissociate the assignments from Risk Factory, students received the BART first as this task is not at all used at Risk Factory. The behavioural measurements were second, and the knowledge-application measurements were last. More specifically, once all participants finished the BART, they were directed to the second part of assignments and were presented with the online sign-up. After the sign-up, participants took the picture-sharing task, because we noticed during pilot-testing that students were quite engaged in this task. We assumed that by doing this task early, engagement would remain relatively high throughout the subsequent tasks. After the picture-sharing task, students continued with the assessment of risky requests and finally ended with the assignment on positive health.

After the final assignment, the researcher reflected with the students on the assignments plenarily, without discussing individual answers. Only at this point the explicit link with Risk Factory was made. Students were informed about the goal of the assignments, which was either to investigate how much students learned at Risk Factory (experimental condition) or how much they already knew about what they were going to learn at Risk Factory in the future (control condition).

Data analysis

To test the three hypotheses, five models were run to compare the five outcome measurements separately between the two between-participant conditions: before Risk Factory (BR) vs after Risk Factory (AR). Contrasts were set to sum-to-zero, with BR as −1 and AR as +1. For all models, we used the lme4 package (version 1.1.26; Bates *et al.*, 2015) in R (R Core Team, 2019).

To test the knowledge-application hypothesis, we ran two linear mixed models, one with the outcome of risk assessment of a request (score 1–10) and one with the outcome of the importance of a question for health (score 1–10). Both models consisted of a fixed intercept, a fixed effect for condition, and random intercepts for the respective item (e.g., request or question), participant and class.

For the hypothesis on topic-specific observed behaviour, we ran two generalised mixed models, using the logit link function, as both behavioural tasks consisted of multiple binomial measurements. These were either whether participants shared a target detail (yes/no) or whether they shared a target picture (yes/no). Again, these models both consisted of a fixed intercept, a fixed effect for condition, and random intercepts for item (e.g., detail or picture), participant and class.

Lastly, to test the general risk-taking tendency hypothesis, we ran a linear mixed model on the amount of pumps on non-burst balloons from the BART. The model structure consisted of a fixed intercept, a fixed effect for condition, and random intercepts for balloon, participant and class. Table 2 shows the model specifications for each hypothesis.

p values for the linear mixed models were calculated with the Anova function from the package car (version 3.0.10; Fox and Weisberg, 2019; which uses the KRmodcomp function from the package pbkrtest: version 0.5.0.1; Halekoh and Højsgaard, 2014) using conditional *F*-tests (Type III) with Kenward-Roger correction of degrees-of-freedom. For the generalised linear models, p values were calculated with the Anova function from the stats package (version 3.6.2; R Core Team, 2019). The confint function of lme4 was used to determine confidence intervals, using bootstrapping with 1,000 simulations.

Results

The following section describes the results of the analyses per hypothesis. For a visual overview of the results on all measurements at once, see Figure 1.

Table 2. Model specifications for every confirmatory hypothesis, comparing the two conditions (before Risk Factory [BR] vs after Risk Factory [AR]) on the respective outcome measures

Hypothesis	Analysis approach	Model
AR pumps less on non-burst balloons than BR	Linear mixed effects model	Pumps on non-burst balloon ~ condition + (1 | item) + (1 | participant) + (1 | class)
AR shares less details than BR	Generalised linear mixed effects model	Detail shared ~ condition + (1 | item) + (1 | participant) + (1 | class)
AR shares less target pictures than BR	Generalised linear mixed effects model	Picture shared ~ condition + (1 | item) + (1 | participant) + (1 | class)
AR rates target requests as more risky than BR	Linear mixed effects model	Risk assessment of request ~ condition + (1 | item) + (1 | participant) + (1 | class)
AR rates target questions as more important for health than BR	Linear mixed effects model	Importance of question for health ~ condition + (1 | item) + (1 | participant) + (1 | class)

Knowledge-application

Like hypothesised, the linear mixed model on knowledge-application about risky requests indicated a significant difference in risk assessment between the two conditions for the target items ($B = 1.52$, $SE = 0.14$, 95% CI [1.24, 1.79], $F(1, 27) = 107.84$, $p < 0.001$). Participants in the AR group rated target requests as more risky ($M = 6.82$, $SD = 2.58$) than the BR group ($M = 3.79$, $SD = 2.18$). We explored filler items, which suggested a difference in the assessment of the risky filler items ($B = 0.19$, $SE = 0.09$, 95% CI [0.01, 0.38], $F(1, 26) = 4.27$, $p = 0.049$) in that the AR group assessed these requests as slightly more risky ($M = 8.10$, $SD = 1.89$) than the BR group ($M = 7.73$, $SD = 1.89$). There were no significant differences in the non-risky filler items ($B = 0.18$, $SE = 0.09$, 95% CI [−0.01, 0.36], $F(1, 26) = 4.17$, $p = 0.051$).

Also in line with our hypothesis, the linear mixed model for the application of knowledge on positive health indicated a significant difference between the conditions in how much the target questions were rated as important for health ($B = 0.35$, $SE = 0.13$, 95% CI [0.09, 0.60], $F(1, 26) = 7.22$, $p = 0.012$). Participants in the AR group indicated the questions related to the five non-physical dimensions of positive health were more important for someone's health ($M = 5.20$, $SD = 2.21$) than participants in the BR condition ($M = 4.51$, $SD = 2.21$). There were no differences between the conditions in the rating of the filler items, both for the items related to physical health ($B = 0.03$, $SE = 0.09$, 95% CI [−0.15, 0,19], $F(1,26)$, $p = 0.76$) as the items non-related to health ($B = 0.35$, $SE = 0.13$, 95% CI [−0.29, 0.18], $F(1, 27) = 0.16$, $p = 0.69$).

Behaviour

The generalised mixed model for the sharing of personal details indicated a significant difference between the two groups in the amount of shared personal details ($B = -8.35$, $SE = 0.68$, 95% CI [−17.73, −6.21], χ^2 $(4, N = 4{,}324) = 155.32$,

https://doi.org/10.1017/bpp.2024.29 Published online by Cambridge University Press

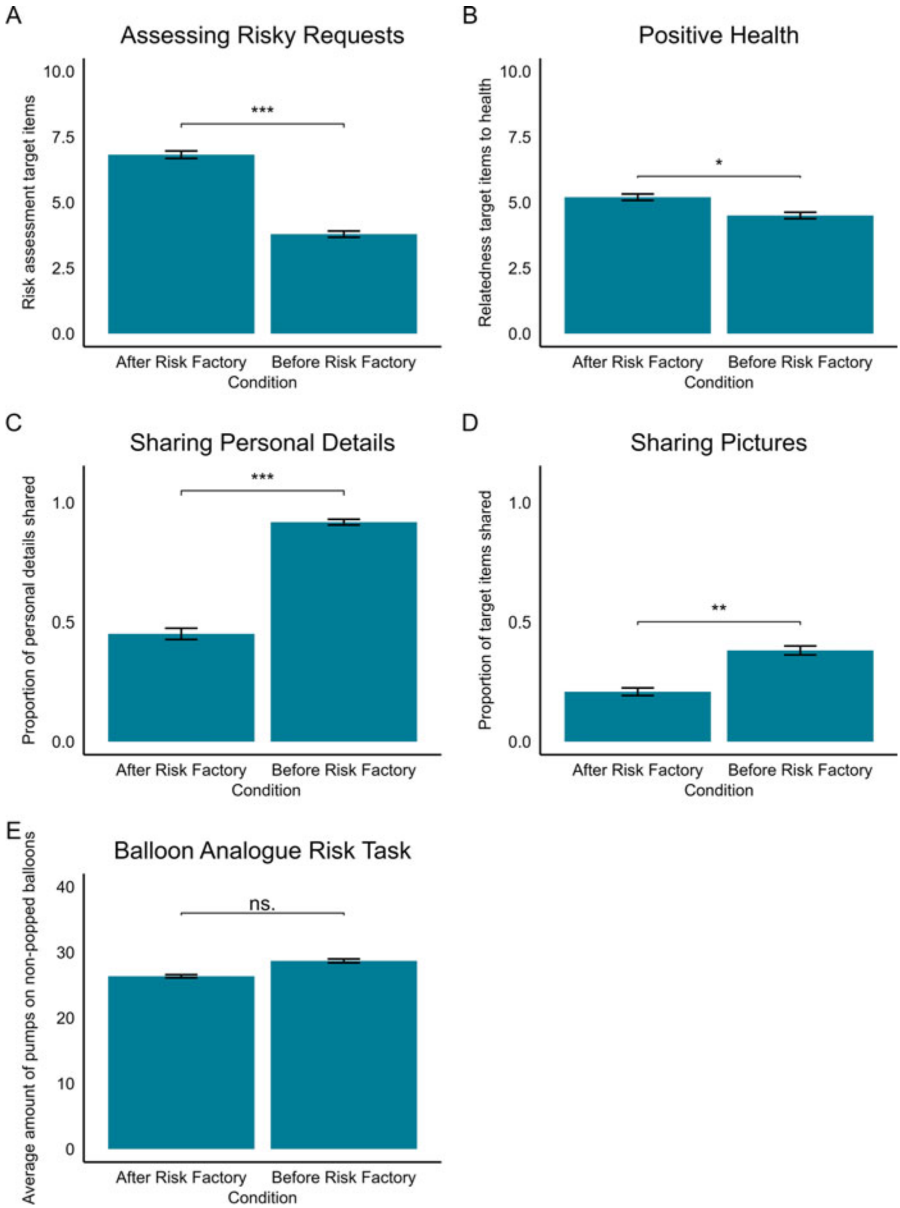

Figure 1. Visual overview of the data for every confirmatory hypothesis

$p < 0.001$). In line with our hypothesis, participants in the BR group were 14.06 times more likely to share a personal detail than participants in the AR group. Out of 2,155 analysed personal details in the BR group, 1,977 were filled in (92%) compared to 980 from 2,169 (45%) in the AR group.

https://doi.org/10.1017/bpp.2024.29 Published online by Cambridge University Press

The generalised mixed model on the picture-sharing behaviour indicated a significant difference between the groups as well ($B = -0.75$, $SE = 0.14$, 95% CI [−1.04, −0.46], χ^2 (4, $N = 5{,}224$) = 18.60, $p < 0.001$). As hypothesised, participants in the BR group were 2.31 times more likely to share a target picture than participants in the AR group. That is, from the 2,624 target pictures analysed in the BR group, 1,004 were shared (38%). For the AR group, 544 out of 2,600 target pictures were shared (21%). The exploratory analyses on the filler items indicated that participants in the AR group shared significantly more filler pictures (both interesting and uninteresting) than the BR group. Participants in the AR group were 1.11 times more likely to share interesting filler items than the BR group ($B = 0.27$, $SE = 0.08$, 95% CI [0.11, 0.44], χ^2 (4, $N = 5{,}224$) = 8.67, $p = 0.003$) and 1.65 times more likely to share uninteresting filler items than the BR group ($B = 0.47$, $SE = 0.17$, 95% CI [0.14, 0.79], χ^2 (4, $N = 5{,}224$) = 7.01, $p = 0.008$).

Descriptives for both topic-specific behavioural measurements (sharing personal details and sharing pictures) can be found in Table 3, including the descriptives for the exploratory analyses.

General risk-taking tendency

Contrary to expectations, the linear mixed model for the BART data showed no differences in the amount of pumps on non-burst balloons between the two conditions ($B = -1.10$, $SE = 0.78$, 95% CI [−2.48, 0.40], $F(1, 22) = 2.07$, $p = 0.164$). Participants in the AR group pumped roughly the same amount on non-burst balloons ($M = 26.36$, $SD = 17.91$) as participants in the BR group ($M = 28.70$, $SD = 18.77$).

Exploratory analyses

Exploratorily, we examined whether the effects of group on the five outcomes were moderated by gender. As such, we included both the main effect of gender and the interaction with group as fixed factors in the previously described models. To run these analyses, we excluded the 11 participants that identified as non-binary, as this group was too small to draw any conclusions from. As such, we were left with 607 participants, of which 320 identified as girls and 287 as boys. Crucially, the main effects of group were similar to the results of the confirmatory analyses on all outcomes. As such, we do not report on those below, but instead focus on the interaction of the group effect with gender.

The linear mixed model on knowledge-application about risky requests showed a significant main effect of gender ($B = 0.19$, $SE = 0.09$, $F(1, 27) = 2.06$, $p = 0.040$). However, this effect was qualified by a group × gender interaction ($B = 0.21$, $SE = 0.09$, $F(1, 27) = 2.28$, $p = 0.022$). The effect of group was stronger for girls ($B = 3.46$, $SE = 0.35$, 95% CI [2.77, 4.15], $p < 0.001$) than boys ($B = 2.62$, $SE = 0.36$, 95% CI [1.92, 3.33], $p < 0.001$). Note that in both genders, the effect of group is still significant in that both boys and girls are less prone to agree to risky requests in the AR than in the BR group.

The linear mixed model for the application of knowledge on positive health also indicated a significant interaction between group and gender ($B = 0.21$, $SE = 0.09$, 95% CI [0.04, 0.38], $F(1, 26) = 6.19$, $p = 0.013$). Again, the effect was stronger for

https://doi.org/10.1017/bpp.2024.29 Published online by Cambridge University Press

Table 3. Descriptive statistics of both behavioural measures: amount of personal details shared and amount of pictures shared per condition

	Confirmatory analyses				Exploratory analyses			
Measurement	Personal details (%)		Target pictures (%)		Interesting filler pictures (%)		Uninteresting filler pictures (%)	
Condition	AR	BR	AR	BR	AR	BR	AR	BR
Shared	980 (45%)	1,977 (92%)	544 (21%)	1,004 (38%)	1,960 (75%)	1,786 (68%)	630 (24%)	285 (11%)
Not shared	1,189 (55%)	178 (8%)	2,056 (79%)	1,620 (62%)	640 (25%)	838 (32%)	1,970 (76%)	2,239 (89%)
Total	2,169	2,155	2,600	2,624	2,600	2,624	2,600	2,624

girls ($B = 1.08$, $SE = 0.31$, 95% CI [0.47, 1.68], $p < 0.001$) compared to boys ($B = 0.23$, $SE = 0.032$, 95% CI [−0.40, 0.85], $p = 0.481$). In this case, the effect of group was only significant for girls.

The generalised mixed model for the sharing of personal details showed a significant effect of gender ($B = -1.48$, $SE = 0.74$, χ^2 (6, $N = 4{,}247$) = 7.74, $p = 0.045$). This main effect was qualified by an interaction of group and gender ($B = -1.55$, $SE =$ 0.74, χ^2 (6, $N = 4{,}247$) = 8.51, $p = 0.003$). Even though the effect of group is stronger in girls compared to males, contrasts showed that in both genders the effect of group was significant (girls: $B = -19.7$, $SE = 2.10$, 95% CI [−23.8, −15.5], $p < 0.001$; boys: $B = -13.5$, $SE = 2.11$, 95% CI [−17.6, −9.35], $p < 0.001$).

The generalised mixed model on the picture-sharing behaviour indicated a significant main effect of gender ($B = -0.25$, $SE = 0.11$, 95% CI [−0.46, −0.04], χ^2 (6, $N = 5{,}136$) = 5.56, $p = 0.018$). Boys were 2.20 times more likely to share target pictures than girls.

Even though the confirmatory analysis did not show an effect of group, the linear mixed model for the BART data did show a significant group × gender interaction ($B = -0.94$, $SE = 0.45$, 95% CI [−1.82, −0.06], $F(1, 22) = 4.41$, $p = 0.036$). The effect of group on general risk-taking tendency was significant in girls ($B = -4.41$, $SE = 1.79$, 95% CI [−7.91, −0.91], $p = 0.014$), but not in boys ($B = -0.65$, $SE = 1.84$, 95% CI [−4.24, 2.95], $p = 0.725$). Girls in the AR group scored lower on risk-taking tendency ($M = 21.8$, $SE = 2.31$) than girls in the BR group ($M = 26.2$, $SE = 2.43$).

Discussion

The results show that a single safety village visit reduces risk-taking of risks that were experienced in the safety village, but it does not reduce general risk-taking tendencies. Moreover, students showed better application of knowledge of the riskiness of a variety of requests, and understanding of positive health. The most important finding is that our results show that an experiential programme such as Risk Factory can change key behaviours related risk, outside of the safety village environment and without explicit reminders of their (future) safety village visit. Specifically, students who visited Risk Factory shared considerably less personal details in an online sign-up situation and shared notably less pictures of recognisable people in emergency situations when doing a consequential picture-sharing task. This is the first preregistered study on safety villages that investigates actual risk behaviour outside of the safety village as an outcome. It is noteworthy that a number of the effects of the visit were stronger for girls than for boys. There are many possible reasons for this unexpected effect (e.g., nature of the scenarios, gender differences in attention or behaviour during the visit), and because these effects were exploratory we refrain from drawing conclusions on this observed difference. Further work is needed to examine whether such gender effects of experienced risk are robust and how they can be explained.

Contrary to the third hypothesis, the visit to Risk Factory did not influence general risk-taking tendencies. One explanation could be that the task-goal of gaining as much money as possible to gain a reward could have overwritten students' learned risk-aversive tendencies. Research shows that, especially when risky behaviour is perceived as rewarding, adolescents' (adequate) risk-perception is insufficient to prevent

https://doi.org/10.1017/bpp.2024.29 Published online by Cambridge University Press

risky behaviour (Reyna and Farley, 2006). Alternatively, it might be argued that the absence of an effect on the BART suggests that, although safety village education seems to be effective in changing students' behaviour within the topics addressed in the education, this learning does not generalise to more trait-like risk-taking tendencies. Research on learning and generalisation to more trait-like constructs indicates that this broader step of generalisation is often hard to accomplish. For example, many people belief that learning the Latin language is useful beyond learning Latin, as learning this language may contribute to cognitive skills such as logical thinking and better learning of other languages (Gerhards *et al.*, 2019). However, a recent review of almost a century of data did not find convincing evidence for such generalisation (Bracke and Bradshaw, 2020). Similarly, a recent meta-analysis (Kassai *et al.*, 2019) showed that interventions training inhibitory control are successful in training the specific components addressed in the trainings, but people rarely transfer the learned skills to untrained components outside of the training (see also Veling *et al.* (2022)).

The present work also suggests that generalisation within experiential safety education occurs up to the level of the specific topics learned. That is, students successfully applied the content from Risk Factory on assessments related to the experienced content areas even when the exact measurements were not directly experienced during the safety village visit. However, effects did not generalise to a measure tapping into more general, decontextualised, risk-taking tendencies. These results are consistent with embodied or situated approaches to cognition proposing that tangible previous experiences can form the basis for later decision making via partial cognitive re-enactment of the previous experiences in relevant situations (e.g., Papies *et al.*, 2022). An important implication of this reasoning is that in order to have societal impact with safety villages on reducing problems with actual risk-taking, it is crucial to select risks that have high prevalence and cause severe problems.

The focus of the present work was to test whether a visit to the safety village would influence risk-taking outside the village after some time and without explicit reminders. There are a number of limitations of the present work that can be addressed in future studies. First, although we did not provide explicit reminders about Risk Factory, we cannot rule out our material did remind students of their visit. For example, in the assignment on risky requests, one item about delivering a sealed package could have reminded students about the scenario in which they specifically learn about the potential dangers of delivering a sealed package. Although this would not influence the BART and both behavioural measurements that came before this assignment, it could have influenced the subsequent measurement on knowledge-application of Positive Health. In addition, we cannot exclude that some of the content of our assignments came up during the reflectional parts of the scenarios at Risk Factory. We had no control over what was discussed during these reflections. It could be that a student addressed a situation during reflection that is completely the same as the one we presented students with in the assignments, and as such explicitly reminded the students that were part of that reflection of their experiences at Risk Factory. This point raises the question whether and how being reminded about the visit is a possible mechanism to understand how the visit can influence risk behaviours even in everyday life. Future work may examine such a mechanistic account.

https://doi.org/10.1017/bpp.2024.29 Published online by Cambridge University Press

Second, it is still unclear what specific strategies within the education caused the effects, as we worked with pre-existing scenarios and did not manipulate anything in these scenarios. In order to gain a better understanding of which concrete strategies cause the most positive outcomes, future research could differentiate the strategies used in an experiential scenario and compare these strategies on their respective outcomes.

Third, we tested the effects of a safety village after an average of 16 days after the visit. This is a fairly short interval, making it hard to draw conclusions about long-term effects of such a visit. The main goal of the current study was to find out whether effects of a single visit on knowledge, skills and behaviour would remain visible outside of the safety village context. Therefore, we decided to examine this after a relatively short 2-week interval. It would be interesting to examine in future research whether the improvements would still be observed after a longer period of time (e.g., 3 months or a year later).

Fourth, since the classes that participated in the control condition would still visit Risk Factory later in the year, we could only use a control condition that received no education. That is, giving any information about Risk Factory's content would diminish students' experiences during the upcoming visit. It thus remains unclear how experiential safety education compares to more traditional forms of presenting information (e.g., learning about risks and safety in the classroom). In the current literature, experiential and traditional methods often seem to be contrasted (James and Williams, 2017), so further research could investigate how experiential safety education compares to more traditional methods.

Finally, risk-taking behaviours related to the experienced scenarios substantially decreased in the context of the present study, but were not completely eliminated. Therefore, an important question for future work is to examine how the safety village approach can contribute to reducing the prevalence of actual risk-taking problems in areas where such villages are implemented. Moreover, an important question is to what degree safety villages reduce risks prevalent among the target group, and are closely related but different from the risks experienced. For instance, the prevalence of cyber-related intimidations is relatively high among younger people (Willoughby *et al.*, 2021), and a reduction in sharing personal details may be helpful to reduce this problematic behaviour by providing less opportunity for young people to become the victim of being intimidated. It is important to examine such indirect effects of this kind of intervention.

Despite the above-mentioned considerations, we argue that this research has demonstrated the value of experiential education in behavioural public policy. As mentioned before, target groups often visit these programmes only once and beneficial outcomes should therefore generalise to situations outside of the initial learning environment and transcend the specific content addressed. In a well-powered preregistered study, we find that students who receive experiential safety education not only recollect information, but apply their knowledge to other contexts and content as well. In addition, we find that students who received experiential safety education behave more safely on two concrete risk-taking behaviours outside the safety village context. These results were found after an average period of roughly 2.5 weeks in a context that used no explicit reminders about Risk Factory. As such, the current paper provides initial support for the type of generalisation-effects experiential education should have in order to properly contribute to the health and safety of society.

Supplementary Material. To view supplementary material for this article, please visit https://doi.org/10.1017/bpp.2024.29.

Funding statement. This research was funded by the Dutch Province of Limburg (Province Limburg). The Province had no role in the writing of this study, nor in the design, execution, analysis and interpretation of the data.

Competing interests. The authors declare none.

Data availability statement. The data that support the findings of this study, as well as the preregistration and supplemental material are openly available in *Preventing risks by experiencing them: Effects on knowledge and behaviour* at https://osf.io/9tcfv/.

Ethical statement. This study was approved by the Ethics Committee of Social Sciences from Radboud University Nijmegen on the protection of the participants (Ethical approval number: 23N.002945).

References

Aklin, W. M., C. W. Lejuez, M. J. Zvolensky, C. W. Kahler and M. Gwadz (2005), 'Evaluation of behavioral measures of risk taking propensity with inner city adolescents', *Behaviour Research and Therapy*, **43**(2): 215–228. https://doi.org/10.1016/j.brat.2003.12.007.

Bates, D., M. Mächler, B. Bolker and S. Walker (2015), 'Fitting linear mixed-effects models using lme4', *Journal of Statistical Software*, **67**(1): 1–48. https://doi.org/10.18637/jss.v067.i01.

Bergsteiner, H., G. C. Avery and R. Neumann (2010), 'Kolb's experiential learning model: critique from a modelling perspective', *Studies in Continuing Education*, **32**(1): 29–46. https://doi.org/10.1080/01580370903534355.

Boam, S. and B. D. Pulford (2019), 'Experiencing risk: the effect of the experiential life-skills centre "Warning Zone" on children's risk perception', *Journal of Risk Research*, **22**(2): 177–190. https://doi.org/10.1080/13669877.2017.1351481.

Boyd, A., J. Golding, J. Macleod, D. A. Lawlor, A. Fraser, J. Henderson, L. Molloy, A. Ness, S. Ring and G. D. Smith (2013), 'Cohort profile: the "children of the 90s"—the index offspring of the Avon Longitudinal Study of Parents and Children', *International Journal of Epidemiology*, **42**(1): 111–127. https://doi.org/10.1093/ije/dys064.

Bracke, E. and C. Bradshaw (2020), 'The impact of learning Latin on school pupils: a review of existing data', *The Language Learning Journal*, **48**(2): 226–236. https://doi.org/10.1080/09571736.2017.1400578.

Congleton, A. R., N. P. Nielsen and D. Berntsen (2021), 'Through the gateway of the senses: investigating the influence of sensory modality-specific retrieval cues on involuntary episodic memory', *Psychological Research*, **85**: 1292–1306. https://doi.org/10.1007/s00426-020-01304-5.

Cook, P. A. and M. A. Bellis (2001), 'Knowing the risk: relationships between risk behaviour and health knowledge', *Public Health*, **115**(1): 54–61. https://doi.org/10.1038/sj.ph.1900728.

Crone, E. A., A. C. van Duijvenvoorde and J. S. Peper (2016), 'Annual Research Review: Neural contributions to risk-taking in adolescence–developmental changes and individual differences', *Journal of Child Psychology and Psychiatry*, **57**(3): 353–368. https://doi.org/10.1111/jcpp.12502.

Eurydice (n.d.). *National Education Systems*, accessed 15 September 2023.

Finucane, M. L., A. Alhakami, P. Slovic and S. M. Johnson (2000), 'The affect heuristic in judgments of risks and benefits', *Journal of Behavioral Decision Making*, **13**(1): 1–17. https://doi.org/10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S.

Fox, J. and S. Weisberg (2019), *An R Companion to Applied Regression*, 3rd edn. Thousand Oaks: Sage.

Gerhards, J., T. Sawert and U. Kohler (2019), 'The emperor's old clothes: fiction and reality of the benefits of knowledge of Latin', *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, **71**: 309–326. https://doi.org/10.1007/s11577-019-00624-8.

Gielen, A. C., A. L. Dannenberg, N. Ashburn and J. Kou (1996), 'Teaching safety: evaluation of a children's village in Maryland', *Injury Prevention*, **2**(1): 26–31. http://dx.doi.org/10.1136/ip.2.1.26.

Halekoh, U. and S. Højsgaard (2014), 'A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models – the R package pbkrtest', *Journal of Statistical Software*, **59**(9): 1–30. https://www.jstatsoft.org/v59/i09/

Hertwig, R. and T. Grüne-Yanoff (2017), 'Nudging and boosting: steering or empowering good decisions', *Perspectives on Psychological Science*, **12**(6): 973–986. https://doi.org/10.1177/1745691617702496.

Hopko, D. R., C. W. Lejuez, S. B. Daughters, W. M. Aklin, A. Osborne, B. L. Simmons and D. R. Strong (2006), 'Construct validity of the Balloon Analogue Risk Task (BART): relationship with MDMA use by inner-city drug users in residential treatment', *Journal of Psychopathology and Behavioral Assessment*, **28**: 95–101. https://doi.org/10.1007/s10862-006-7487-5.

Huber, M., M. van Vliet, M. Giezenberg, B. Winkens, Y. Heerkens, P. C. Dagnelie and J. A. Knottnerus (2016), 'Towards a 'patient-centred' operationalisation of the new dynamic concept of health: a mixed methods study', *BMJ Open*, **6**(1): e010091. http://dx.doi.org/10.1136/bmjopen-2015-010091.

James, J. K. and T. Williams (2017), 'School-based experiential outdoor education: a neglected necessity', *Journal of Experiential Education*, **40**(1): 58–71. https://doi.org/10.1177/1053825916676190.

Jessor, R. E. (1998), *New Perspectives on Adolescent Risk Behavior*. Cambridge: Cambridge University Press.

Kassai, R., J. Futo, Z. Demetrovics and Z. K. Takacs (2019), 'A meta-analysis of the experimental evidence on the near-and far-transfer effects among children's executive function skills', *Psychological Bulletin*, **145**(2): 165. https://doi.org/10.1037/bul0000180.

Keiser, N. L. and S. C. Payne (2019), 'Are employee surveys biased? Impression management as a response bias in workplace safety constructs', *Safety Science*, **118**: 453–465.

Kolb, D. A. (1984), *Experiential Learning: Experience as the Source of Learning and Development*. Upper Saddle River: Prentice Hall.

Kolb, D. A. (2015), *Experiential Learning: Experience as the Source of Learning and Development*. London: Pearson.

Lamb, R., M. S. Joshi, W. Carter, G. Cowburn and A. Matthews (2006), 'Children's acquisition and retention of safety skills: the Lifeskills program', *Injury Prevention*, **12**(3): 161–165. https://doi.org/10.1136/ip.2005.010769.

Lauriola, M., A. Panno, I. P. Levin and C. W. Lejuez (2013), 'Individual differences in risky decision making: a meta-analysis of sensation seeking and impulsivity with the balloon analogue risk task', *Journal of Behavioral Decision Making*, **27**(1): 20–36. https://doi.org/10.1002/bdm.1784.

Lejuez, C. W., J. P. Read, C. W. Kahler, J. B. Richards, S. E. Ramsey, G. L. Stuart, D. R. Strong and R. A. Brown (2002), 'Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART)', *Journal of Experimental Psychology: Applied*, **8**(2): 75–84. https://doi.org/10.1037/1076-898X.8.2.75.

Lejuez, C. W., W. M. Aklin, M. J. Zvolensky and C. M. Pedulla (2003), 'Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours', *Journal of Adolescence*, **26**(4): 475–479. https://doi.org/10.1016/S0140-1971(03)00036-8.

Lejuez, C. W, W. Aklin, S. Daughters, M. Zvolensky, C. Kahler and M. Gwadz (2007), 'Reliability and validity of the youth version of the Balloon Analogue Risk Task (BART–Y) in the assessment of risk-taking behavior among inner-city adolescents', *Journal of Clinical Child & Adolescent Psychology*, **36**(1): 106–111. https://doi.org/10.1080/15374410709336573.

MacLean, R. R., A. L. Pincus, J. M. Smyth, C. F. Geier and S. J. Wilson (2018), 'Extending the Balloon Analogue Risk Task to assess naturalistic risk taking via a mobile platform', *Journal of Psychopathology and Behavioral Assessment*, **40**(1): 107–116. https://doi.org/10.1007/s10862-017-9628-4.

Morris, T. H. (2019), 'Experiential learning – a systematic review and revision of Kolb's model', *Interactive Learning Environments*, **28**(8): 1–14. https://doi.org/10.1080/10494820.2019.1570279.

Mütze, F. and W. De Dobbeleer (2019), 'The status of traffic safety and mobility education in Europe', *European Transport Safety Council*. accessed 15 September 2023.

Oxford Evaluation Team (2003), *An evaluation of the Lifeskills: learning for living program (HSE Research Report 187)*. Norwich: HMSO. https://www.hse.gov.uk/research/rrhtm/rr187.htm

Papies, E. K., L. W. Barsalou, M. A. Claassen, T. Davis, S. T. Farrar, E. Gauthier, A. Rodger, B. Tatar, L. H. Wehbe and J. Werner (2022), 'Grounding Motivation for Behavior Change', in B. Gawronski (ed.) , *Advances in Experimental Social Psychology*, Vol. **66**. Academic Press, 107–189. https://doi.org/10.1016/bs.aesp.2022.04.002

R Core Team (2019), *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Reyna, V. F. and F. Farley (2006), 'Risk and rationality in adolescent decision making: implications for theory, practice and public policy', *Psychological Science in the Public Interest*, **7**(1): 1–44. https://doi.org/10.1111/j.1529-1006.2006.00026.x.

Robin, J. and M. Moscovitch (2014), 'The effects of spatial contextual familiarity on remembered scenes, episodic memories, and imagined future events', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **40**(2): 459–475. https://doi.org/10.1037/a0034886

Smith, S. M. (2013), 'Effects of Environmental Context on Human Memory', in D. S. Lindsay and T. J. Perfect (eds), *Sage Handbook of Applied Memory*, London: SAGE Publications Ltd.

Teyhan, A., R. Cornish, J. Macleod, A. Boyd, R. Doerner and M. S. Joshi (2016), 'An evaluation of the impact of "Lifeskills" training on road safety, substance use and hospital attendance in adolescence', *Accident Analysis & Prevention*, **86**: 108–113. https://doi.org/10.1016/j.aap.2015.10.017.

Thomson, J. A. (2006), 'Issues in safety education interventions', *Injury Prevention: Journal of the International Society for Child and Adolescent Injury Prevention*, **12**(3): 138–139. https://doi.org/10.1136/ip.2006.012625.

van der Kint, S. and C. Mons (2021), *Interpolis Barometer 2021*. Den Haag, The Netherlands: Interpolis.

Veling, H., D. Becker, H. Liu, J. Quandt and R. W. Holland (2022), 'How go/no-go training changes behavior: a value-based decision-making perspective', *Current Opinion in Behavioral Sciences*, **47**: 101206.

Wåhlberg, A. E., L. Dorn and T. Kline (2010), 'The effect of social desirability on self reported and recorded road traffic accidents', *Transportation Research Part F: Traffic Psychology and Behaviour*, **13**(2): 106–114. https://doi.org/10.1016/j.trf.2009.11.004.

Westfall, J., D. A. Kenny and C. M. Judd (2014), 'Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli', *Journal of Experimental Psychology: General*, **143**(5): 2020–2045. https://doi.org/10.1037/xge0000014.

White, T. L., C. W. Lejuez and H. de Wit (2008), 'Test-retest characteristics of the Balloon Analogue Risk Task (BART)', *Experimental and Clinical Psychopharmacology*, **16**(6): 565–570. https://doi.org/10.1037/a0014083.

Willoughby, T., T. Heffer, M. Good and C. Magnacca (2021), 'Is adolescence a time of heightened risk taking? An overview of types of risk-taking behaviors across age groups', *Developmental Review*, **61**: 100980. https://doi.org/10.1016/j.dr.2021.100980.

Zeedyk, M. S., L. Wallace, B. Carcary, K. Jones and K. Larter (2001), 'Children and road safety: increasing knowledge does not improve behaviour', *British Journal of Educational Psychology*, **71**(4): 573–594. https://doi.org/10.1348/000709901158686.

Cite this article: Bulterman, R., T. Verwijmeren, R.W. Holland and H. Veling (2025), 'Preventing risks by experiencing them: effects on knowledge and behaviour', *Behavioural Public Policy*, 1–20. https://doi.org/10.1017/bpp.2024.29