Virtual reality training for police officers: a comparison of training responses in VR and real-life training

Kleygrewe, Lisanne; Hutter, Vana; Koedijk, Matthijs; Oudejans, Raôul

published in
Police Practice and Research
2023

Police Practice and Research
An International Journal

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gppr20

Virtual reality training for police officers: a comparison of training responses in VR and real-life training

Lisanne Kleygrewe, R. I. Vana Hutter, Matthijs Koedijk & Raôul R. D. Oudejans

To cite this article: Lisanne Kleygrewe, R. I. Vana Hutter, Matthijs Koedijk & Raôul R. D. Oudejans (2023): Virtual reality training for police officers: a comparison of training responses in VR and real-life training, Police Practice and Research, DOI: 10.1080/15614263.2023.2176307

To link to this article: https://doi.org/10.1080/15614263.2023.2176307

Published online: 09 Feb 2023.

POLICE PRACTICE AND RESEARCH
https://doi.org/10.1080/15614263.2023.2176307

Virtual reality training for police officers: a comparison of training responses in VR and real-life training

Lisanne Kleygrewe[a,b,c], R. I. Vana Hutter[a,b,c,d], Matthijs Koedijk[a,b,c] and Raôul R. D. Oudejans[a,b,c,e]

[a]Department of Human Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands; [b]Institute of Brain and Behaviour Amsterdam, Amsterdam, the Netherlands; [c]Amsterdam Movement Sciences, Amsterdam, the Netherlands; [d]Netherlands Institute for the Study of Crime and Law Enforcement (Nederlands Studiecentrum Criminaliteit en Rechtshandhaving; NSCR), Amsterdam, the Netherlands; [e]Faculty of Sports and Nutrition, Amsterdam University of Applied Sciences, Amsterdam, the Netherlands

ABSTRACT

In policing, Virtual Reality scenario-based training (VR SBT) is being explored to complement real-life scenario-based training (RL SBT). We investigated whether relevant training responses in VR SBT, namely heart rate (HR), level of physical activity, mental effort, and perceived stress, resemble those in RL SBT. Utilizing a within-subject study design, we investigated the training responses of 237 police officers of the Dutch National Police. We found that the maximum HR and average level of physical activity were significantly higher in RL SBT, whereas invested mental effort was significantly higher in VR SBT. No significant differences were found in average HR and perceived stress. We also found that perceived stress in VR was predicted by participants' VR experiences such as engagement with VR and experience of negative effects, but not by participant characteristics. Participants' mental effort in VR was predicted by their VR experiences and participant characteristics, particularly gaming frequency. In conclusion, VR SBT can elicit perceived stress, mental effort, and average HR that resemble or exceed responses in RL SBT, providing a promising tool to complement police training.

ARTICLE HISTORY
Received 3 October 2022
Accepted 30 January 2023

KEYWORDS
Police training; training responses; virtual reality; mental effort; perceived stress

1 Introduction

Police officers respond to complex on-duty situations on a regular basis (Anderson et al., 2002; Gershon et al., 2009). For instance, when engaging in a domestic violence dispute, police officers are faced with a variety of demands: they have to establish the context of the situation, evaluate the risk of harm to themselves and others, communicate with dispatch, the possible perpetrator, the victim, and each other, and make decisions regarding the appropriate interventions (e.g., the need for use of force, arrest, medical services, etc.). While most routine police tasks, such as administrative duties or traffic respondence are likely to be non-violent (Famega, 2005), police officers also have to be prepared to respond to high-risk incidents such as escalating situations with armed perpetrators (MacDonald et al., 2003). Consequently, police officers have to be prepared for numerous situations ranging from regulated and routine tasks to complex, ambiguous, and potentially life-threatening encounters.

CONTACT Lisanne Kleygrewe l.kleygrewe@vu.nl Department of Human Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

Through training, police agencies equip their officers with the knowledge and skills to respond to the wide range of situations encountered on duty. An established method to prepare officers for a variety of situations in training and replicate realistic and life-like on-duty incidents is scenario-based training (Baldwin et al., 2021; DiNota & Huhta, 2019; Renden et al., 2015). Real-life scenario-based training (RL SBT) aims to simulate on-duty incidents by constructing training environments that closely resemble on-duty situations. It incorporates immersive role-plays (e.g., actors that play the perpetrator and/or victims) and other props (e.g., a suitable training location mimicking a living room, sounds such as loud music) to replicate the on-duty situation as realistically as possible. By engaging in interactive and dynamic training scenarios, trainees learn to execute verbal, cognitive, and physical skills concurrently under representative conditions and can explore a variety of behavioral strategies to resolve complex situations in training (DiNota & Huhta, 2019; Rajakaruna et al., 2017). Implementing RL SBT that repeatedly exposes police officers to the stress they experience during complex on-duty tasks helps officers to gain familiarity with stress and may enhance performance under those conditions (Nieuwenhuys & Oudejans, 2011).

Although RL SBT enhances officers' performance of on-duty tasks under stressful circumstances (Andersen & Gustafsberg, 2016; Andersen et al., 2016; Baldwin et al., 2021), several issues with the implementation of realistic RL SBT into practice remain, such as inadequate delivery and resource- and labor-intensiveness (Cushion, 2020, 2022; Rajakaruna et al., 2017). For example, police instructors require suitable locations that fit the training scenarios (e.g., an apartment building for a domestic violence scenario), actors, experienced instructors, or well-instructed trainees that play the role-player (e.g., to play the victims or perpetrator), functional gear that resembles the operational gear (e.g., FX systems that replicate the operational weapon but use non-lethal munition), and sufficient training time to set-up the scenario and complete it with multiple trainees. Thus, limitations in delivery and resource availability make the efficacious application and implementation of RL SBT in police training challenging (Kleygrewe et al., 2022).

The use of virtual reality training – scenario-based training using a virtual simulator (VR SBT) — offers new ways to complement RL SBT (Haskins et al., 2020; Murtinger et al., 2021). Immersive VR training systems provide a 3D environment in which trainees can freely move about and interact with simulated surroundings (Scarfe & Glennerster, 2019). VR SBT provides the advantage that it can be performed largely independent of a training location as it relies only on sufficient space to set up a VR system. Additionally, police instructors can control the content of the scenarios and create a wide variety of simulated environments without needing additional resources such as props or different training locations. Most VR systems offer an after-action review in which instructors and trainees can play back the scenario from various perspectives while also providing a variety of performance data (e.g., shots fired, targets hit, bystanders flagged). Having these after-action review options from the VR system may enhance the way instructors provide and trainees receive feedback, as the objective and visual information provided during the review is less abstract compared to verbal feedback after RL SBT. Using VR SBT may thus provide the opportunity to negate some of the challenges that RL SBT poses (Giessing, 2021).

In recent years, researchers have investigated the efficacy of VR as a training tool in practice. Groer et al. (2010) demonstrated that simulated environments (i.e., a video projected onto a wall) can elicit similarly high levels of stress reactivity to lethal use of force as those experienced during real-life incidents. With evolvement in technology, Bertram et al. (2015) showed that the training transfer (measured by a complex task in reality) is similar in VR SBT and RL SBT indicating that training in VR may lead to similar performance outcomes as training in real-life. More recently, VR has been used to explore police officers' psychophysiological responses (i.e., electroencephalographic and heart rate variability parameters) to firearm shooting tasks (Muñoz et al., 2020). Moreover, Binsch et al. (2022) found that preparation for police surveillance tasks was more effective for performance and recovery when done in VR compared to real-life. Taken together, these findings highlight that VR can be an effective tool for training cognitive-perceptual skills (Harris et al., 2021) and retaining and applying police-specific knowledge (Saunders et al., 2019).

However, little is known about how police officers respond to dynamic, interactive VR SBT physically and psychologically.

On-duty, police officers engage in physically and psychologically demanding tasks (Andersen et al., 2016). In training, these task demands should be reflected in the responses of police officers to prepare them for duty (DiNota & Huhta, 2019; Kleygrewe et al., 2022). Training modalities such as RL SBT and VR SBT should therefore elicit physical and psychological training responses that expose officers to experiences they may have in the field. Physical training responses such as heart rate and level of physical activity give insights into the cardiovascular demands and amount of movement required during the training. Eliciting physical training responses is relevant because police officers are required to perform under high physical strain (Andersen et al., 2016; Baldwin et al., 2019) and benefit from experiencing this type of strain in training. Similarly, psychological training responses such as mental effort and perceived stress (Houtman & Bakker, 1989; Zijlstra, 1993) provide insights into the cognitive processes that trainees experience during the training. For instance, according to attentional control theory (Eysenck et al., 2007), investing mental effort is a compensatory strategy to negate the influence of stress and anxiety on performance and has been shown to provide a potentially effective strategy for police officers after training in high-anxiety conditions (Nieuwenhuys & Oudejans, 2011). Therefore, training modalities such as RL SBT and VR SBT should provoke psychological responses to familiarize police officers with the influence of perceived stress and mental effort in training. While research has extensively investigated RL SBT in police (e.g., Andersen et al., 2016; Baldwin et al., 2021; Cushion, 2022; Jenkins et al., 2021), insights into physical and psychological training responses to VR SBT are sparse. Understanding these training responses to VR SBT enhances the use of VR in police and provides information on how VR SBT differs from RL SBT to determine where the two modalities could complement each other to improve current training practices.

As VR is a fairly new training tool in policing, police officers are likely to encounter VR as a training modality for the first time. Some officers may find it easy to be immersed and engaged from the start, others may have difficulties adjusting to the virtual environment and require additional training time in VR to become sufficiently immersed and engaged. It is likely that overall knowledge about and frequent use of technology, as well as prior experience with VR, may make it easier for police officers to engage with VR as a training tool (Pletz, 2021). Other factors such as being prone to cybersickness or experiencing other adverse effects may negatively influence the way in which police officers experience and engage with VR (Weech et al., 2019). Yet, little is currently known about how police officers engage with VR as a training technology and whether specific protocols should be in place to make VR SBT more efficient and effective for police officers. Investigating factors such as sense of presence (i.e., the feeling of 'being there' in the virtual environment; Lessiter et al., 2001), cybersickness, and user characteristics related to technology affinity may specify how VR training can be adjusted to engage police officers and consequently improve the application of VR in police.

The current study aims to close the research gaps surrounding VR for police, particularly regarding police officers' training responses to VR SBT and their experiences with VR as a training tool. First, we investigate physical (i.e., heart rate, level of physical activity) and psychological (invested mental effort and perceived stress) training responses of police officers to VR SBT and RL SBT. Inherent differences between RL SBT and VR SBT exist (see for instance Giessing, 2021); therefore, we investigate the differences of the two training modalities as they are currently used in police practice (e.g., short sequences of VR scenarios; long, extensive reality-based scenarios). Examining the physical and psychological responses to the two types of training may highlight the strengths and weaknesses of each and how they can be used to complement each other. Second, we investigate factors (i.e., participant characteristics and VR-specific experiences) that we hypothesize to influence the psychological training responses (i.e., invested mental effort and perceived stress) of police officers. Understanding how participant characteristics (such as age or experience with technology) and VR-specific experiences (such as sense of presence or the experience of negative effects) relate to or influence psychological training responses may provide initial guidelines on how VR SBT can be tailored to be integrated into existing training frameworks. Thus, the current study takes a first step at investigating the application of VR for dynamic, interactive SBT and how police officers (physically and psychologically) respond to and experience VR as a training technology.

2 Methods

This study was conducted in collaboration with the Dutch National Police. We conducted our experiment around the annual training days of the special intervention unit with a primary focus on close protection tasks (e.g., ensuring safety of clients, assessing security, and providing surveillance). By conducting our experiments as part of the annual training days, we were able to test and assess the training modalities RL SBT and VR SBT as they are applied in practice. This means that we utilized a large-scale scenario set-up with various role-players and props in the RL SBT, followed by a verbal debrief with the instructor. Comparatively, for VR SBT we utilized three short scenarios (to highlight the possibility of training a variety of adaptable scenarios in a short time) in combination with the possibility to receive an objective after-action review after each scenario. Hence, this study was not set up to compare identical trainings in real-life and VR environments but to compare the physical and psychological training responses to RL SBT and VR SBT as applied in practice.

2.1 Participants

In total, 237 street patrol officers of the Dutch National Police with additional tasks in the Dutch special intervention unit (227 male, 8 female, and 2 other; *M* age = 39.39, *SD* = 7.82) participated in this study. The participants' experience on the job ranged from 3 to 42 years (*M* years = 15.16 years, *SD* = 6.78). Participants provided informed consent before the start of the experiment. Ethical approval was obtained from the Social and Societal Ethics Committee of the Katholieke Universiteit Leuven as part of the SHOTPROS project which is funded by the European Union's Horizon 2020 Research and Innovation Programme (Grant number: 833672).

2.2 Design

We utilized a within-subject study design. All participants completed RL SBT and the VR SBT in their respective training groups. To counterbalance the training order, half of the training groups in this study completed the RL SBT followed by the VR SBT; the other half of the training groups completed the VR SBT followed by RL SBT. The training groups consisted of 12 to 16 participants depending on the size of the operational unit. For the RL SBT, the training group of 12 to 16 participants executed one large training scenario. The training scenario took on average 40 minutes. For the VR SBT, the operational unit was split into two training groups consisting of six to eight participants (depending on the initial size of the operational unit). This was done to allow sufficient computing power in the VR. Each group of six to eight participants completed a sequence of three VR SBT scenarios which on average took six minutes each (total VR SBT time on average 18 minutes).

In this study, we assessed three main measures: physical training responses (HR, physical activity), psychological training responses (perceived stress, mental effort), and VR experiences (participant characteristics, sense of presence). Due to the training schedule of the Dutch police, the set-up of the study was such that some participants were unable to complete all three study measures (without causing a delay in the training schedule). In addition, due to limited availability of measurement equipment (i.e., Zephyr Bioharness devices), we could not monitor the physical activity and heart rate of all participants. Thus, we utilized three distinct sub-samples for the three measures we recorded in this study. A total of 210 participants completed the VR experience measures, a total of 114 participants completed the measures of psychological response measures, and a total of 54 participants took part in the measures of physical responses. The overlap of participants taking part in the measures can be found in Figure A1 in the Appendix; descriptive statistics of the sub-samples can be found in Table A1 in the Appendix. There are no notable differences in the demographic distribution of the sub-samples (see Table A1).

2.3 Training set-up

All training took place at a training facility consisting of multiple empty buildings rented specifically for the purpose of the training. The VR SBT took place in a large hangar. The RL SBT took place on the outside and inside of an empty office-like building. As this training was conducted for police officers in the Dutch special intervention unit with tasks as close protection officers, the overall training objective focused on protection tasks; particularly on patrolling to ensure safe surroundings, protecting the entrances of buildings, and spotting suspicious behavior. Training scenarios in VR and RL were adjusted to accommodate for the training objectives while also remaining suitable for the particular training modality. For instance, to use VR effectively, we selected three shorter training scenarios that combined building protection, patrolling, and spotting of suspicious behavior, using different virtual environments in each of the scenarios. For RL SBT, we utilized one large-scale cohesive scenario that combined the same training objectives as VR.

2.3.1 Real-life scenario-based training (RL SBT)

RL SBT was designed together with experienced training instructors of the Dutch National Police. The scenario was staged around an empty one-story building consisting of various rooms and one main entrance. The front side of the building faced a street. Participants wore their regular patrol gear, protective helmets, vests, earplugs, and protective glasses for use with FX cartridges. Participants were equipped with FX weapons (rifle-type) and FX non-marking cartridges, non-lethal training ammunition. Props used in the scenario included smoke bombs, police cars, civilian cars, and scooters. For the scenario, various role-play actors were hired to act in the following roles: a lawyer, a police officer, two suspects (equipped with FX handguns and non-marking cartridges, protection equipment), and four bystanders. The role-players were well rehearsed and adhered to the scenario script developed by the police instructors.

In the training scenario, participants were tasked with protecting a lawyer (role-player) and securing the lawyer's office (the empty one-story building on the training premise). Participants took on various tasks according to their roles in their operational protection unit. Some participants were positioned outside the building to spot suspicious behavior of people or suspicious vehicles approaching the building. Another group of participants was positioned inside the building to provide personal protection for the lawyer. Additional participants were positioned at the main entrance to only give cleared people (role-players) access to the building. In the first 25 minutes of the scenarios, various people (role-players) would approach the building, attempting to get access. After 25 minutes passed, a police car arrived with a police officer as a role-player who exited the car. Meanwhile, two civilian cars approached the building and stopped in the middle of the road. The two civilians (role-players) left the car and threw smoke bombs and started shooting at the building to gain access. The police officer (role-player) who left the police car appeared to be shot and acted seriously injured. The participants' task was to control the situation (e.g., attend to the injured officer and protect the lawyer). The training instructors who oversaw the training session were in charge of ending the scenario, typically as soon as the suspects were arrested or the threat was stopped.

2.3.2 Virtual reality scenario-based training (VR SBT)

The VR system used in this experiment was provided by RE-liON (www.re-lion.com). Participants were equipped with RE-liONs' Blacksuit consisting of a binocular head-mounted display (including

Figure 1. RE-liOn's Blacksuit – VR Equipment. The VR equipment was provided by RE-liON.

microphone, sound effect, radio chatter), a smart vest with full-body tracking, a computing box (backpack style), and replica rifle. Figure 1 shows the VR equipment used in this study. In addition to the VR equipment, there was one VR system operator and two police instructors guiding the training. To ensure the safety of the participants and avoid falls, collisions, and limit the occurrence of cybersickness, the participants were instructed not to run during the VR SBT.

The training group of 12 to 16 participants was split into two smaller groups of six to eight participants to train the VR scenarios in small units ensuring sufficient computing power. The two small groups underwent calibration of the VR suits and equipment together and then completed a 5-minute instruction and familiarization scenario in the system. Next, the first small unit (in the following referred to as Group 1) started the first VR scenario while the other small unit (in the following referred to as Group 2) had the option to view the scenario from the outside on a large screen. Once Group 1 finished the first VR sequence, the second group started their first VR sequence. While Group 2 prepared for the first training scenario, Group 1 received an after-action review with a police instructor on a large screen next to the VR station. Once Group 2 started the training scenario, Group 1 had the option to view the scenario from the outside on a large screen. Once Group 2 finished their VR SBT sequence, the groups would swap again, and Group 2 received the after-action review while Group 1 prepared for the next training scenario. This process was repeated until all participants completed three VR SBT scenarios.

The VR scenarios depicted a large, square, multi-story building surrounded by parking lots and small streets, as can be seen in Figure 2. The participants' task was to patrol around the building to

Figure 2. VR Training Environment. The left picture shows the top-down layout of the VR training environment. The right picture shows the VR environment looking at the main entrance of the building. Pictures were provided by RE-liON.

spot suspicious behavior, engage perpetrators when necessary, and resolve any altercations accordingly. Various non-player characters (NPCs) with predetermined behaviors were placed in the scenarios and controlled by an operator from RE-liON who – under the guidance of an instructor – directed which threat would appear when and where. These on-the-fly scenario variations allowed the adjustment of the level of complexity to the performance of the participants in the scenarios, ensuring that participants completed three slightly different VR scenarios, while keeping the overall structure and objective of the training the same for all participants. On average, one VR scenario lasted approximately six minutes.

2.4 Measures

In this study, we assessed three main measures: physical training responses (average heart rate, maximum heart rate, activity), psychological training responses (mental effort, perceived stress), and experiences in VR (sense of presence). The specific sub-measures are described below.

2.4.1 Physical training responses

2.4.1.1 Heart rate (HR). Average and maximum heart rate (HR) in beats per minute (bpm) were recorded using a Zephyr Bioharness 3.0 device (www.zephyranywhere.com) at a recording frequency of 1 Hz. HR recordings of the active training time in the training scenarios for VR and RL were extracted and analyzed. HR Confidence (degree of validity of HR value, as a %) provided in the Zephyr output were used for data correction: data points at or below 25% HR Confidence were considered invalid and were removed from analysis (see 'Heart Rate Confidence', OmniSense Analysis Help, 2016). Physical inspection of extreme bpm values (outside of a realistic HR range, e.g., 0 bpm) were removed from analysis.

The Zephyr Bioharness device provides valid and reliable HR measurements (Nazari et al., 2018) and is frequently used in police research (e.g., Andersen & Gustafsberg, 2016; Bertilsson et al., 2020). In police research, HR has consistently been used as a common parameter to assess police officers' cardiovascular response to stress (Andersen & Gustafsberg, 2016; Andersen et al., 2016; Anderson et al., 2002, 2019; Baldwin et al., 2019; Bertilsson et al., 2020; Vonk, 2008).

2.4.1.2 Level of physical activity. Level of physical activity was obtained using the Zephyr Bioharness 3.0 device (www.zephyranywhere.com). The Zephyr Bioharness contains a 3-axis accelerometer that records vertical, lateral, and sagittal acceleration magnitudes. The level of physical activity was quantified as velocity magnitude units (VMU), measured in g. VMU = $\sqrt{x^2 + y^2 + z^2}$ where x, y, and z are the averages of the three axial acceleration magnitudes over the previous

1-second epoch, sampled at 100 Hz. A VMU of higher than .2 indicates a walking equivalent activity, and a VMU of higher than .8 indicates a running equivalent activity (OmniSense Analysis Help, 2016).

2.4.2 Psychological training responses

To assess psychological training responses, we utilized visual analogue scales (VAS) to assess mental effort and perceived stress. In police research, VAS for mental effort and perceived stress (originally anxiety) have been frequently used to quantify self-perceived psychophysiological experiences during training and complex police tasks (e.g., Giessing et al., 2019; Nieuwenhuys et al., 2009; R. Oudejans, 2008; Wilson et al., 2007).

2.4.2.1 Mental effort. Subjective ratings of mental effort were obtained using the VAS 'Rating Scale for Mental Effort' (RSME; Zijlstra, 1993) at the end of the RL and VR SBT (once after all VR scenarios were completed). The RSME was assessed on a VAS from 1 to 150. According to Zijlstra (1993), the RSME has adequate test-retest reliability with correlation coefficients between 0.78 in work settings and 0.88 in laboratory settings.

2.4.2.2 Perceived stress. Subjective ratings of perceived stress were obtained using the VAS for anxiety (adjusted to 'stress' instead of anxiety; Houtman & Bakker, 1989) at the end of the RL SBT and VR SBT (after all VR scenarios were completed). Perceived stress refers to the participants' subjective experience of stress during the training scenarios. The stress scale was assessed on a VAS from 1 to 100. The original anxiety scale has a fair validity and test-retest reliability with correlation coefficients ranging between 0.60 and 0.78 (Houtman & Bakker, 1989).

2.4.3 Experiences in VR

2.4.3.1 Sense of presence. The experiences in the virtual environment were assessed using the ITC-Sense of Presence Inventory (ITC-SOPI; Lessiter et al., 2001). Each item of the inventory is rated on a Likert-scale from 1 'strongly disagree' to 5 'strongly agree'. The ITC-SOPI results in four factors. Spatial presence refers to the sense of being part of the virtual environment. Engagement refers to the feeling of involvement with the content and feeling psychologically engaged. Ecological validity or naturalness refers to the tendency of perceiving the virtual environment as life-like and natural. Negative effects refer to the experience of any adverse physiological experiences such as dizziness or headaches (Lessiter et al., 2001).

According to Lessiter et al. (2001), the ITC-SOPI has good internal consistency. The Cronbach alpha coefficient for spatial presence was .94, for engagement .89, for ecological validity .76, and for negative effect .77. In the current study, the Cronbach alpha coefficient for the factor spatial presence was .89, for the factor engagement .80, for the factor ecological validity .74, and for the factor negative effect .83.

In addition to assessing sense of presence factors, the ITC-SOPI contains a section in which the participant's background information is obtained. Specifically, participant characteristics such as age, level of computer experience, gaming frequency, prior experience with VR, and knowledge about VR.

2.5 Statistical analysis

To investigate whether VR and RL SBT elicit differences in training responses in police officers, we conducted five paired-samples t-tests using each of the training response variables (average HR, maximum HR, average activity, invested mental effort, perceived stress). To further examine whether the psychological training responses in VR were influenced by factors that are hypothesized to impact the VR experience (e.g., age, technology experience, cybersickness in VR as obtained with the ITC-SOPI), we performed two separate hierarchical multiple linear regression analyses for invested mental effort in VR and perceived stress in VR. For each of the models, the predictor variables were entered in two pre-determined steps. First, we entered participants characteristics consisting of age, VR knowledge, gaming frequency, and prior VR experience. Second, we entered the VR sense of presence factors from the ITC-SOPI consisting of spatial presence, engagement, ecological validity, and negative effects. Entering the predictors in two separate steps allowed us to investigate the change in explained variance for each block of predictors and evaluate the relative importance of each predictor with each step. For each model, we conducted residual analyses to ensure no violations of assumptions. To check for multicollinearity, we used correlation coefficients above 0.7 among independent variables, Variance Inflation Factor (VIF) above 10, and Tolerance values of less than .10 as indicators for multicollinearity (Miles, 2005). To detect multivariate outliers, we used visual inspection of the scatterplot of the standardized residuals (cut-off for outliers>3.3 or < −3.3), Mahalanobis Distance (critical value of Chi-Square for eight predictor variables = 26.13 at $p = .001$) and Cook's Distance (critical value>1; Tabachnick & Fidell, 2018). P-values of<0.05 were considered statistically significant. Cohen's d was calculated as an estimate for effect size. A value of $d = 0.2$ indicated a small effect size, a value of $d = 0.5$ indicated a medium effect size and a value of $d = 0.8$ indicated a large effect size (Cohen, 1988). All statistical analyses were performed using IBM SPSS, version 27.

3 Results

3.1 Differences in training responses between RL SBT and VR SBT

Paired-samples t-tests were conducted to investigate differences in training responses between VR SBT and RL SBT. No statistically significant difference was found in average HR (bpm). Maximum HR (bpm) was significantly higher in RL ($M = 136.96$, $SD = 16.82$) than in VR ($M = 126.28$, $SD = 16.23$), $t(53) = -4.2$, $p < .001$, $d = 0.65$. Average level of physical activity (VMU) was significantly higher in RL ($M = 0.08$, $SD = 0.02$) than in VR ($M = 0.06$, $SD = 0.01$), $t(53) = -5.71$, $p < .001$, $d = 1.20$. Invested mental effort (RSME) was significantly higher in VR ($M = 52.64$, $SD = 25.57$) than in RL ($M = 46.11$, $SD = 21.17$), $t(113) = 2.68$, $p = .008$, $d = 0.28$. No statistically significant difference was found in perceived stress. Detailed statistics of the paired-samples t-tests can be found in Table 1.

3.2 Variance in mental effort in VR explained by participant characteristics and VR experience factors

Hierarchical multiple regression was used to assess whether sense of presence indicators in VR (spatial presence, engagement, and negative effects) predicted mental effort (VAS RSME) in VR after controlling for the influence of participant characteristics (age, VR knowledge, gaming

Table 1. Paired-samples t-test results.

	VR SBT			RL SBT				95% CI					
	n	*M*	*SD*	*n*	*M*	*SD*	Mean Difference	Lower	Upper	*df*	*t*	*p*	Cohen's *d*
Average HR	54	90.91	12.84	54	88.63	13.22	2.28	−0.92	5.49	53	1.43	.159	0.18
Maximum HR	54	126.28	16.23	54	136.96	16.82	−10.69	−15.78	−5.58	53	−4.2	<.001	0.65
Average Activity	54	0.06	0.01	54	0.08	0.02	−0.02	−0.03	−0.01	53	−5.71	<.001	1.20
Mental Effort	114	52.64	25.57	114	46.11	21.17	6.53	1.71	11.34	113	2.68	.008	0.28
Perceived Stress	114	36.10	20.37	114	38.98	19.90	−2.89	−7.11	1.34	113	−1.35	.179	0.14

Average and maximum HR in bpm. Average (physical) activity in VMUs. Mental effort (RSME) on a visual analogue scale from 1 to 150. Perceived stress on a visual analogue scale from 1 to 100.

Table 2. Hierarchical multiple regression analysis summary predicting mental effort in VR with age, VR knowledge, gaming frequency, prior VR experience, VR spatial presence, VR engagement, and VR negative effects.

Step and predictor variable	*B*	*SE B*	Beta	*sr*	Change in R^2	R^2	Adjusted R^2
Step 1					.10*	.10	.07
Constant	76.48	13.68					
Age	−0.57	0.27	−.18*	−.17			
VR Knowledge	5.44	2.79	.18	.16			
Gaming Frequency	−6.90	2.03	−.31**	−.28			
Prior VR Experience	1.01	4.64	.02	.02			
Step 2					.21**	.31	.28
Constant	−18.90	21.80					
Age	−0.46	0.24	−.14	−.14			
VR Knowledge	3.12	2.53	.11	.09			
Gaming Frequency	−6.26	1.82	−.28**	−.25			
Prior VR Experience	0.27	4.12	.01	.00			
VR Spatial Presence	−11.80	6.01	−.23	−.14			
VR Engagement	30.71	6.14	.61**	.37			
VR Negative Effects	8.31	2.19	.30**	.28			

sr = semipartial correlation coefficient. * $p < .05$. ** $p < .001$.

frequency, prior VR experience). Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity, and homoscedasticity. Due to high correlation coefficients amongst the predictor variables spatial presence and ecological validity (.80) indicating violations of the assumptions of multicollinearity, we omitted ecological validity from the model (resulting in a higher adjusted *R* squared compared to a model retaining ecological validity as a predictor). Age, VR knowledge, gaming frequency, prior VR experience were entered at Step 1, explaining 10% of the variance in mental effort. After entry of VR spatial presence, VR engagement, and VR negative effects at Step 2 the total variance explained by the model as a whole was 31%, F (7, 126) = 8.21, $p < .001$. The sense of presence indicators explained an additional 21% of the variance in mental effort after controlling for participant characteristics, *R* squared change =.21, *F* change (3, 126) = 12.92, $p < .001$. In the final model, only one participant characteristic variable was statistically significant, with gaming frequency recording a semipartial correlation value of −.25 ($p < .001$), indicating that trainees with a high gaming frequency exerted less mental effort in VR. Two of the sense of presence indicators in VR were statistically significant, with engagement recording the highest semipartial correlation value ($sr = .37$, $p < .001$), closely followed by VR negative effects ($sr = .28$, $p < .001$). Therefore, only gaming frequency, VR engagement, and VR negative effects have unique contributions to the invested mental effort (RSME) after statistically removing the overlapping effects of all other variables. Table 2 provides an overview of all predictor variables and the two steps of the hierarchical multiple regression model predicting mental effort in VR. Descriptive statistics and correlation coefficients can be found in the Appendix in Tables A2 and A3, respectively.

3.3 Variance in perceived stress in VR explained by participant characteristics and VR experience factors

Hierarchical multiple regression was used to assess whether sense of presence indicators in VR (spatial presence, engagement, and negative effects) predicted perceived stress in VR after controlling for the influence of participant characteristics (age, VR knowledge, gaming frequency, prior VR experience). Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity, and homoscedasticity. When checking for multicollinearity, we observed high correlation coefficients amongst the predictor variables spatial presence, engagement, and ecological validity (>.70, see Table A5 in Appendix) indicating possible violations of the assumptions of multicollinearity. While VIF and Tolerance values were well below critical threshold, we rebuilt two iterative models omitting ecological validity in the first and engagement in

Table 3. Hierarchical multiple regression analysis summary predicting perceived stress in VR with age, VR knowledge, gaming frequency, prior VR experience, VR spatial presence, VR engagement, and VR negative effects.

Step and predictor variable	*B*	*SE B*	Beta	*sr*	Change in R^2	R^2	Adjusted R^2
Step 1					.03	.03	.00
Constant	25.26	11.78					
Age	.13	.23	.05	.05			
VR Knowledge	4.13	2.43	.17	.15			
Gaming Frequency	−1.27	1.78	−.07	−.06			
Prior VR Experience	.74	4.00	.02	.02			
Step 2					.20**	.23	.17
Constant	−60.99	19.73					
Age	.28	.21	.11	.10			
VR Knowledge	2.62	2.26	.11	.09			
Gaming Frequency	−1.28	1.64	−.07	−.06			
Prior VR Experience	−.24	3.64	−.01	−.01			
VR Spatial Presence	−7.65	6.35	−.17	−.10			
VR Engagement	16.12	5.72	.38*	.22			
VR Ecological Validity	10.09	5.55	.25	.14			
VR Negative Effects	5.31	1.94	.23*	.22			

sr = semipartial correlation coefficient. * $p < .05$. ** $p < .001$.

the second model. Both models yielded a lower adjusted R^2 (.163 and .172, respectively) than the initial model that contained all sense of presence indicators as predictors (.174); thus, we retained the initial model with all four sense of presence indicators. During the check for multivariate outliers, we have identified two cases that exceeded the maximum Mahalanobis distance value (cut-off value of 26.13 for eight predictors, see Tabachnick & Fidell, 2018, Table C.4) and omitted the two cases from further analysis. Age, VR knowledge, gaming frequency, prior VR experience were entered at Step 1, explaining 3% of the variance in perceived stress ($p = .460$). After entry of VR spatial presence, VR engagement, and VR negative effects at Step 2 the total variance explained by the model as a whole was 23%, $F\,(8, 123) = 4.45$, $p < .001$. The sense of presence indicators explained an additional 20% of the variance in perceived stress after controlling for participant characteristics, R squared change =.20, F change (4, 123) = 7.80, $p < .001$. In the final model, none of the participant characteristic variables were statistically significant. However, two of the sense of presence indicators in VR were statistically significant, with engagement recording a slightly higher semipartial correlation value ($sr = .224$, $p = .006$) than VR negative effects ($sr = .218$, $p = .007$). Therefore, only VR engagement and VR negative effects have unique contributions to the perceived stress in VR after statistically removing the overlapping effects of all other variables. Table 3 provides an overview of all predictor variables and the two steps of the hierarchical multiple regression model predicting perceived stress in VR. Descriptive statistics and correlation coefficients can be found in the Appendix in Tables A4 and A5, respectively.

4 Discussion

In the present study, we compared the training responses police officers experienced during RL SBT and VR SBT. Because police officers engage in physically and psychologically demanding tasks in the field, this should also be reflected and implemented in police training (Andersen et al., 2016; DiNota & Huhta, 2019). Thus, we operationalized physical (i.e., average and maximum HR, activity level) and psychological (i.e., mental effort, perceived stress) responses to capture and compare the training responses of police officers to RL SBT and VR SBT. We further investigated what other factors (i.e., participant characteristics, sense of presence in VR) influence psychological responses of police officers to VR SBT.

We found that police officers had higher maximum HRs and higher levels of activity in RL SBT compared to VR SBT. These may, at least partly, be explained by the training instructions, the first-time experience with VR of most officers, and the inherent characteristics of VR systems such as bulky equipment leading to movement constraints (Giessing, 2021). During VR SBT, trainees were asked not to run with the VR equipment to ensure safety. Comparatively, in the RL SBT, participants had to promptly perceive the severity of the situation (i.e., being under attack) and quickly put on their protection gear (running to put on heavy protection plate, helmet, etc.) leading to significantly higher peak levels of activity and maximum HRs. Additionally, first-time VR exposure can induce initial postural instability (da Silva Marinho et al., 2022), thus forcing further inhibition of fast or vigorous movement in VR that is not present in RL.

Further, we found that average HRs did not differ between the two training modalities. In our study, police officers experienced similar average levels of cardiovascular responses to VR and RL SBT. This suggests that, in their current state of application, VR and RL are both capable of producing training situations that induce physiological responses to external stressors. However, compared to the literature reporting HR in police training (e.g., Armstrong et al., 2014; R. R. Oudejans & Pijpers, 2009), the average HRs elicited in our study were fairly low (on average 91 bpm in VR and 89 bpm in RL). This may be explained by the overall training objective of training officers in protecting a building and spotting suspicious behavior rather than placing them in physically demanding situations. This training objective led to training scenarios comprised of largely stationary tasks (i.e., officers positioning themselves to protect a building and look for threats, with occasional patrolling, and encountering attackers) which is also reflected in the low levels of activity during both RL and VR and recorded VMUs that are well below values that would indicate a continuous walking equivalent activity (OmniSense Analysis Help, 2016). Note that largely stationary tasks are not uncommon in police work (see Famega, 2005), which is why stationary task components were implemented in the scenarios of this study.

Overall, the differences in physical responses to the training modalities seem to be dependent on the training objective and thus tasks within the training scenarios, as well as the inherent characteristics of VR systems and the safety requirements that trainees should adhere to when wearing the VR equipment. Therefore, when aiming to train scenarios that require high levels of physical activity (i.e., running), it appears most appropriate to train these tasks using RL SBT.

Regarding psychological responses elicited by RL SBT and VR SBT, we found that police officers experienced similar levels of perceived stress during RL SBT and VR SBT. Akin to the average HRs found in our study, the recorded levels of perceived stress are relatively low compared to studies that looked at perceived anxiety or stress of police officers during low- and high-stress training conditions (e.g., Giessing et al., 2019; R. Oudejans, 2008; Wilson et al., 2007). Nonetheless, we found that police officers perceived VR SBT to be as stressful as RL SBT, closely aligning with existing literature showing that training in VR can elicit significantly elevated psychological and physiological stress responses (van Dammen et al., 2022). These findings suggest that while VR SBT may not be appropriate for training objectives that require high levels of physical activity, VR SBT is suited for the training of mentally demanding tasks that put additional strain on the officer by providing ecologically valid training environments.

Since participants in our study experienced similar levels of perceived stress during VR SBT and RL SBT, we would expect that participants invest similar amounts of mental effort during both training modalities. Because high levels of (perceived) stress can negatively influence the performance of police officers, a compensatory strategy to negate the influence of stress is the investment of extra mental effort (see attentional control theory, Eysenck et al., 2007). Thus, similar levels of perceived stress would require similar levels of mental effort to negate the stress responses. In contrast, we found that participants invested more mental effort during VR SBT compared to RL SBT. A possible explanation for this finding may be that VR SBT places more extraneous cognitive load on trainees than RL SBT does (Mugford et al., 2013). In accordance with the cognitive load theory (CLT; Van Merrienboer & Sweller, 2005), extraneous cognitive load refers to the demands a trainee experiences that are not relevant for and potentially harmful to learning (Clark et al., 2011; Paas et al., 2003). During VR SBT, it may be that participants experience additional extraneous cognitive load and had to invest more mental effort to get used to the newness of the virtual environment, VR equipment, and VR as a training tool. To test this reasoning, we further investigated whether factors such as participants' level of prior use of technology and VR experiences affect invested mental effort in VR.

Indeed, we found that invested mental effort in VR SBT was related to participant characteristics that are suggestive of technological affinity. This relation between participants characteristics such as age, VR knowledge, gaming frequency, or previous VR experience was only found for mental effort and not for perceived stress. We therefore (precautiously) infer that during VR SBT mental effort is invested as a strategy to process information of the virtual environment, rather than for stress mitigation alone, and that experience with VR and gaming makes navigating the VR environment mentally less demanding. This notion is in line with existing research (Rosa et al., 2016), indicating that trainees with a high gaming frequency need to invest little mental effort to get used to and engage with the virtual environment. Similarly, people with more gaming experience have been shown to recover from cybersickness better than people with little gaming experience (da Silva Marinho et al., 2022). Thus, police trainers have to be aware that trainees with differing characteristics (e.g., trainees with much gaming experience vs. trainees with little gaming experience) may have different psychological experiences in the VR SBT.

Next to participant characteristics that influence the psychological responses of trainees during VR SBT, the momentary experiences of the virtual environment during the training also appears to play a vital role in the psychological responses that officers experience during VR SBT. Investigating VR experience-specific factors like those measured by the ITC-SOPI (i.e., spatial presence, engagement, ecological validity, and negative effects) provides insights into the trainees' experience with VR as a training tool and whether the experience with the training tool itself influences the psychological responses of trainees. Of the predictor variables measured in this study, we found that the VR experience variables explained the greatest single proportion of variance in perceived stress and mental effort (20% and 21%, respectively). When police officers experienced more engagement with the virtual environment and more negative effects, they perceived higher levels of stress and invested more mental effort. The relation between experiencing negative effects in VR and the psychological responses provoked by VR indicates that, in order for VR SBT to elicit similar psychological responses to RL SBT, the experience of negative effects in VR should be reduced as much as possible. To this end, any available technical, content, and human-factor solutions should be considered to minimize the risk of cybersickness and other adverse effects (e.g., Chang et al., 2020). Similarly, the relation between engagement and psychological responses to VR suggests that VR SBT requires sufficient levels of engagement with the virtual environment and content. To increase engagement, findings of Pengnate et al. (2020) showed that VR environments that have an interactive, narrative design result in higher engagement compared to VR that did not have a narrative-based environment (for practical recommendations on how to design a VR scenario that promotes engagement, see section 5. 'Concluding remarks').

4.1 Limitations and future directions

Although the study aimed to investigate VR and RL SBT in an operational field-training environment to obtain ecologically valid results, the design and application have limitations. First, the current study aimed to investigate differences in the training modalities of VR and RL SBT as they are currently used in police practice. To take advantage of the inherent differences in the training modalities, a direct comparison between the two training modalities and inferences about their effectiveness compared to each other was not possible as the training set-ups differed in scenario repetition, task, context, duration, and number and intensity of confrontations in the scenarios. Additionally, modality-specific instructions such as the safety instruction to avoid running during VR SBT provided differences in training delivery of the modalities. While the scenarios in VR SBT and RL SBT were based on pre-written scenario-scripts, the developments of the scenarios were based on the actions of the trainees leading to differences in scenario execution and timing of events even within training groups. Thus, if the goal is to achieve a one-to-one comparison of the training modalities in terms of their effectiveness in producing a certain training response (and other measures of interest), future studies may want to standardize the training set-ups as much as possible. In this study, however, we wanted to gain initial insights into (i) the application of the modalities (taking into account that the two types of training are inherently different), (ii) show that VR can elicit training responses that are similar to those elicited in RL SBT, and (iii) derive implications of how and where VR and RL SBT can supplement each other based on the assessed measures (see section 5, 'Concluding remarks').

Second, due to the strict training schedule, we were not able to obtain HR baseline measures. Thus, in this experiment, we did not compare the elicited HRs to the basal levels of the participants and therefore could not examine how responses in VR and RL SBT were similar to either basal levels or the heightened responses in-situ. Due to the difference in scenario length and development between RL SBT and VR SBT, the averaging of HR over the length of the scenario may also limit the concrete comparability of the obtained values. However, using a within-subject design, we were able to make general inferences regarding the differences in HR responses from VR SBT to RL SBT; particularly, as VR SBT and RL SBT took place on the same day and were counterbalanced in sequence to avoid the influence of an order effect. Nonetheless, research in the field of police training has demonstrated that RL SBT is able to elicit HRs up to 150 bpm (e.g., Baldwin et al., 2021). It is yet to be shown whether VR SBT is able to elicit comparable HR levels in order to facilitate training objectives that require extreme physiological stress reactivity experienced on-duty and during RL SBT. Future work on VR SBT in police should obtain valid and reliable baseline measures and consider assessing additional measures such as heart rate variability as an objective method of quantifying acute operational stress (Corrigan et al., 2021). In addition, future studies should, next to physical and psychological training responses, investigate behavioral responses to VR SBT. Behavioral responses may provide valuable insights into the action possibilities of police officers in VR and therefore inform police agencies for which training areas and tasks VR is suitable.

5 Concluding remarks

VR SBT is becoming an increasingly popular topic for police agencies. In this paper, we have taken the first steps to explore how VR SBT compares and fits into existing training practices. To this end, our results revealed that VR SBT is capable of eliciting similar training responses in police officers as RL SBT does, providing initial considerations for the implementation of VR SBT to complement current RL SBT practices. For instance, when a training objective requires officers to exert high levels of physical activity (such as chasing, apprehending, and arresting a suspect), these objectives should be trained using RL SBT as it allows for more flexibility in movement and physical interaction with a role-player. To complement such training, VR can be used to broaden the scope of the objective. For example, while chasing and apprehending a suspect are important skills, police officers also have to be able to spot suspects and suspicious behavior in novel and disorganized environments. For tasks like these, VR offers valuable training opportunities (see, for instance, Harris et al., 2021 for use of VR in police room searching procedures). It allows for the adjustment and variation of virtual environments and the use of multiple NPCs (including groups usually not easily integrated in training such as children) and role-players that can take on various avatars making them less identifiable as suspects (as is oftentimes a problem with RL SBT due to the protective equipment worn by suspect role-players). Based on our findings, VR SBT may even be tailored to specific trainee characteristics to be more effective. For instance, trainees with technology affinity can start VR SBT rather quickly whereas first-time users with little technology experience may need additional VR tutorials and more training time to get used to VR. To ensure that VR SBT elicits psychological responses that are similar to RL SBT, factors such as the occurrence of negative effects and the engagement in VR should be considered and addressed beforehand. For instance, to minimize the risk of cybersickness in VR, particularly for first-time users, the structure of the VR SBT should be adjusted so that the initial exposure duration is brief and becomes gradually and incrementally longer to aid adaptation to VR (da Silva Marinho et al., 2022). Additionally, in setting up a VR training, instructors should ensure sufficient break time (for instance, for performance feedback) before re-immersion into the next VR scenario. To increase engagement with and sense of presence in VR, the training scenarios should take place in an interactive narrative-based virtual environment (Pengnate et al., 2020). For example, when designing training scenarios for VR, police trainers should, just as for RL SBT, develop scenario scripts that specify the task the trainees should perform, include realistic soundscapes (e.g., police car sirens), and interactive features such as role-players and NPCs.

Further systematic investigation of how the practical recommendations provided in this section impact physical, psychological, and behavioral training responses is critical for the effective implementation of VR SBT in police training practices. Investigating and validating how practical recommendations such as the structure, delivery, and scenario design of VR SBT influence training responses (e.g., HR, physical activity, perceived stress, mental effort, behavior) and the VR experience (i.e., sense of presence, occurrence of cybersickness) provides police agencies with clear guidelines on how they can effectively implement and apply VR SBT into current training curricula. Taken together, we foresee that with continuous technological developments and further systematic investigation, VR SBT will play an important role in enhancing current police training practices.

Acknowledgments

The authors would like to thank RE-liON, the Dutch National Police, their police trainers, and the police officers who participated in this experiment for their invaluable contribution to this study. In particular, we would like to thank Jeffrey Sprakel for the organization and oversight of the training.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 833672.The content reflects only authors´ view. Research Executive Agency and European Commission are not liable for any use that may be made of the information contained herein.

Notes on contributors

Lisanne Kleygrewe is a PhD-candidate at the Department of Human Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands. Her interests are training and performance of police officers under pressure. Her PhD research focuses on the use and effectiveness of VR training in police practice.

Dr Vana Hutter is a senior researcher for the research program What works in policing: towards evidence-based policing in the Netherlands at the Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), and assistant professor Sport and Performance Psychology, Department of Human Movement Sciences, Vrije Universiteit Amsterdam. Her interests are in training and assessment, professional development, and in teamwork under pressure.

Dr Matthijs Koedijk is a postdoctoral researcher in Human Movement Sciences, Department of the Human Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands. His interests are training and assessment of public safety organizations to develop and test their staff members' ability to perform job-specific tasks. His PhD research aims to enhance the practice of assessment of professional skills and fitness for duty.

Dr Raôul Oudejans is (Associate) Professor Sport and Performance Psychology, Department of Human Movement Sciences, Vrije Universiteit Amsterdam and Department of Sports and Nutrition, Amsterdam University of Applied Sciences, The Netherlands. His main research and teaching areas are perceiving and moving in sports and other high-pressure contexts such as police work and performing arts with emphasis on the psychological factors involved in performing.

ORCID

Lisanne Kleygrewe http://orcid.org/0000-0001-9685-4198
R. I. Vana Hutter http://orcid.org/0000-0002-3583-9362
Matthijs Koedijk http://orcid.org/0000-0001-5577-8926
Raôul R. D. Oudejans http://orcid.org/0000-0002-1643-0107

References

Andersen, J. P., & Gustafsberg, H. (2016). A training method to improve police use of force decision making: A randomized controlled trial. *SAGE Open*, *6*(2), 215824401663870. https://doi.org/10.1177/2158244016638708

Andersen, J. P., Pitel, M., Weerasinghe, A., & Papazoglou, K. (2016). Highly realistic scenario based training simulates the psychophysiology of real world use of force encounters: Implications for improved police officer performance.

Anderson, G. S., DiNota, P., Metz, G., & Andersen, J. (2019). The impact of acute stress physiology on skilled motor performance: Implications for policing. *Frontiers in Psychology*, *10*, 2501. https://doi.org/10.3389/fpsyg.2019.02501

Anderson, G. S., Litzenberger, R., & Plecas, D. (2002). Physical evidence of police officer stress. *Policing: An International Journal of Police Strategies & Management*, *25*(2), 399–420. https://doi.org/10.1108/13639510210429437

Armstrong, J., Clare, J., & Plecas, D. (2014). Monitoring the impact of scenario-based use-of-force simulations on police heart rate: Evaluating the Royal Canadian Mounted Police Skills Refresher Program. *The Journal of Criminology, Criminal Justice, Law & Society*, *15*(1), 51.

Baldwin, S., Bennell, C., Andersen, J. P., Semple, T., & Jenkins, B. (2019). Stress-activity mapping: Physiological responses during general duty police encounters. *Frontiers in Psychology*, *10*, 2216. https://doi.org/10.3389/fpsyg.2019.02216

Baldwin, S., Bennell, C., Blaskovits, B., Brown, A., Jenkins, B., Lawrence, C., McGale, H., Semple, T., & Andersen, J. P. (2021). A reasonable officer: Examining the relationships among stress, training, and performance in a highly realistic lethal force scenario. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.759132

Bertilsson, J., Niehorster, D. C., Fredriksson, P., Dahl, M., Granér, S., Fredriksson, O., Mårtensson, J., Magnusson, M., Fransson, P. -A., & Nyström, M. (2020). Towards systematic and objective evaluation of police officer performance in stressful situations. *Police Practice & Research*, *21*(6), 655–669. https://doi.org/10.1080/15614263.2019.1666006

Bertram, J., Moskaliuk, J., & Cress, U. (2015). Virtual training: Making reality work? *Computers in Human Behavior*, *43*, 284–292. https://doi.org/10.1016/j.chb.2014.10.032

Binsch, O., Oudejans, N., van der Kuil, M. N., Landman, A., Smeets, M. M., Leers, M. P., & Smit, A. S. (2022). The effect of virtual reality simulation on police officers' performance and recovery from a real-life surveillance task. *Multimedia Tools and Applications*, 1–22. https://doi.org/10.1007/s11042-022-14110-5

Chang, E., Kim, H. T., & Yoo, B. (2020). Virtual reality sickness: A review of causes and measurements. *International Journal of Human–Computer Interaction*, *36*(17), 1658–1682. https://doi.org/10.1080/10447318.2020.1778351

Clark, R. C., Nguyen, F., & Sweller, J. (2011). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. John Wiley & Sons.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.

Corrigan, S. L., Roberts, S., Warmington, S., Drain, J., & Main, L. C. (2021). Monitoring stress and allostatic load in first responders and tactical operators using heart rate variability: A systematic review. *BMC Public Health*, *21*(1), 1–16. https://doi.org/10.1186/s12889-021-11595-x

Cushion, C. J. (2020). Exploring the delivery of officer safety training: A case study. *Policing: A Journal of Policy and Practice*, *14*(1), 166–180.

Cushion, C. J. (2022). Changing police personal safety training using scenario-based-training: A critical analysis of the 'dilemmas of practice' impacting change [conceptual analysis]. *Frontiers in Education*, *6*, 6. https://doi.org/10.3389/feduc.2021.796765

da Silva Marinho, A., Terton, U., & Jones, C. M. (2022). Cybersickness and postural stability of first time VR users playing VR videogames. *Applied Ergonomics*, *101*, 103698. https://doi.org/10.1016/j.apergo.2022.103698

DiNota, P. M., & Huhta, J. -M. (2019). Complex motor learning and police training: Applied, cognitive, and clinical perspectives. *Frontiers in Psychology*, *10*, 1797. https://doi.org/10.3389/fpsyg.2019.01797

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336. https://doi.org/10.1037/1528-3542.7.2.336

Famega, C. N. (2005). Variation in officer downtime: A review of the research. *Policing: An International Journal of Police Strategies & Management*, *28*(3), 388–414. https://doi.org/10.1108/13639510510614528

Gershon, R. R., Barocas, B., Canton, A. N., Li, X., & Vlahov, D. (2009). Mental, physical, and behavioral outcomes associated with perceived work stress in police officers. *Criminal Justice and Behavior*, *36*(3), 275–289. https://doi.org/10.1177/0093854808330015

Giessing, L. (2021). The potential of virtual reality for police training under stress: A SWOT analysis. In Patrick Arble, E., & B. Arnetz, B. (Eds.), *Interventions, training, and technologies for improved police well-being and performance* (pp. 102–124). IGI Global.

Giessing, L., Frenkel, M. O., Zinner, C., Rummel, J., Nieuwenhuys, A., Kasperk, C., Brune, M., Engel, F. A., & Plessner, H. (2019). Effects of coping-related traits and psychophysiological stress responses on police recruits' shooting behavior in reality-based scenarios [original research]. *Frontiers in Psychology*, *10*, 10. https://doi.org/10.3389/fpsyg.2019.01523

Groer, M., Murphy, R., Bunnell, W., Salomon, K., Van Eepoel, J., Rankin, B., White, K., & Bykowski, C. (2010). Salivary measures of stress and immunity in police officers engaged in simulated critical incident scenarios. *Journal of Occupational and Environmental Medicine*, *52*(6), 595–602. https://doi.org/10.1097/JOM.0b013e3181e129da

Harris, D. J., Hardcastle, K. J., Wilson, M. R., & Vine, S. J. (2021). Assessing the learning and transfer of gaze behaviours in immersive virtual reality. *Virtual Reality*, *25*(4), 961–973. https://doi.org/10.1007/s10055-021-00501-w

Haskins, J., Zhu, B., Gainer, S., Huse, W., Eadara, S., Boyd, B., Laird, C., Farantatos, J., & Jerald, J. (2020). *Exploring VR training for first responders*. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Atlanta, GA, USA.

Houtman, I., & Bakker, F. (1989). The anxiety thermometer: A validation study. *Journal of Personality Assessment*, *53*(3), 575–582. https://doi.org/10.1207/s15327752jpa5303_14

Jenkins, B., Semple, T., Quail, J., & Bennell, C. (2021). Optimizing scenario-based training for law enforcement. *Interventions, Training, and Technologies for Improved Police Well-Being and Performance*, 18–37.

Kleygrewe, L., Oudejans, R. R., Koedijk, M., & Hutter, R. (2022). Police training in practice: Organization and delivery according to European law enforcement agencies. *Frontiers in Psychology*, *12*, 6515. https://doi.org/10.3389/fpsyg.2021.798067

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence: Teleoperators & Virtual Environments*, *10*(3), 282–297. https://doi.org/10.1162/105474601300343612

MacDonald, J. M., Manz, P. W., Alpert, G. P., & Dunham, R. G. (2003). Police use of force: Examining the relationship between calls for service and the balance of police force and suspect resistance. *Journal of Criminal Justice*, *31*(2), 119–127. https://doi.org/10.1016/S0047-2352(02)00219-2

Miles, J. (2005). Tolerance and variance inflation factor. *Encyclopedia of Statistics in Behavioral Science*. https://doi.org/10.1002/0470013192.bsa683

Mugford, R., Corey, S., & Bennell, C. (2013). Improving police training from a cognitive load perspective. *Policing: An International Journal of Police Strategies & management: Policing*, *36*(2), 312–337. https://doi.org/10.1108/13639511311329723

Muñoz, J. E., Quintero, L., Stephens, C. L., & Pope, A. T. (2020). A psychophysiological model of firearms training in police officers: A virtual reality experiment for biocybernetic adaptation. *Frontiers in Psychology*, *11*, 683. https://doi.org/10.3389/fpsyg.2020.00683

Murtinger, M., Jaspaert, E., Schrom-Feiertag, H., & Egger Lampl, S. (2021). Cbrne training in virtual environments: SWOT analysis & practical guidelines. *International Journal of Safety and Security Engineering*, *11*(4), 295–303. https://doi.org/10.18280/ijsse.110402

Nazari, G., Bobos, P., MacDermid, J. C., Sinden, K. E., Richardson, J., & Tang, A. (2018). Psychometric properties of the Zephyr bioharness device: A systematic review. *BMC Sports Science, Medicine and Rehabilitation*, *10*(1), 1–8. https://doi.org/10.1186/s13102-018-0094-4

Nieuwenhuys, A., Caljouw, S. R., Leijsen, M. R., Schmeits, B. A., & Oudejans, R. R. (2009). Quantifying police officers' arrest and self-defence skills: Does performance decrease under pressure? *Ergonomics*, *52*(12), 1460–1468. https://doi.org/10.1080/00140130903287981

Nieuwenhuys, A., & Oudejans, R. R. (2011). Training with anxiety: Short-and long-term effects on police officers' shooting behavior under pressure. *Cognitive Processing*, *12*(3), 277–288. https://doi.org/10.1007/s10339-011-0396-x

OmniSense Analysis Help. 2016. Zephyr Technology. https://www.zephyranywhere.com/media/download/omnisense-analysis-help-20-apr-2016.pdf

Oudejans, R. (2008). Reality-based practice under pressure improves handgun shooting performance of police officers. *Ergonomics*, *51*(3), 261–273. https://doi.org/10.1080/00140130701577435

Oudejans, R. R., & Pijpers, J. R. (2009). Training with anxiety has a positive effect on expert perceptual–motor performance under pressure. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1631–1647. https://doi.org/10.1080/17470210802557702

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional: Design Recent developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Pengnate, S., Riggins, F. J., & Zhang, L. (2020). Understanding users' engagement and responses in 3D virtual reality: The influence of presence on user value. *Interacting with Computers*, *32*(2), 103–117. https://doi.org/10.1093/iwc/iwaa008

Pletz, C. (2021). Which factors promote and inhibit the technology acceptance of immersive virtual reality technology in teaching-learning contexts? Results of an expert survey. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(13), 248–272. https://doi.org/10.3991/ijet.v16i13.20521

Rajakaruna, N., Henry, P. J., Cutler, A., & Fairman, G. (2017). Ensuring the validity of police use of force training. *Police Practice & Research*, *18*(5), 507–521. https://doi.org/10.1080/15614263.2016.1268959

Renden, P. G., Landman, A., Savelsbergh, G. J., & Oudejans, R. R. (2015). Police arrest and self-defence skills: Performance under anxiety of officers with and without additional experience in martial arts. *Ergonomics*, *58*(9), 1496–1506. https://doi.org/10.1080/00140139.2015.1013578

Rosa, P. J., Morais, D., Gamito, P., Oliveira, J., & Saraiva, T. (2016). The immersive virtual reality experience: A typology of users revealed through multiple correspondence analysis combined with cluster analysis technique. *Cyberpsychology, Behavior and Social Networking*, *19*(3), 209–216. https://doi.org/10.1089/cyber.2015.0130

Saunders, J., Davey, S., Bayerl, P. S., & Lohrmann, P. (2019). *Validating virtual reality as an effective training medium in the security domain*. 2019 IEEE conference on virtual reality and 3D user interfaces (VR), Osaka, Japan.

Scarfe, P., & Glennerster, A. (2019). The science behind virtual reality displays. *Annual Review of Vision Science*, *5*(1), 529–547. https://doi.org/10.1146/annurev-vision-091718-014942

Tabachnick, B. G., & Fidell, L. S. (2018). *Using multivariate statistics* (7 ed.). Pearson.

van Dammen, L., Finseth, T. T., McCurdy, B. H., Barnett, N. P., Conrady, R., Leach, A. G., Deick, A. F., van Steenis, A. L., Gardner, R., Smith, B. L., Kay, A., & Shirtcliff, E. A. (2022). Evoking stress reactivity in virtual reality: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *138*, 104709. https://doi.org/10.1016/j.neubiorev.2022.104709

Van Merrienboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Vonk, K. (2008). *Police performance under stress*. Hendon publishing company.

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. *Frontiers in Psychology*, *10*, 158. https://doi.org/10.3389/fpsyg.2019.00158

Wilson, M., Smith, N. C., & Holmes, P. S. (2007). The role of effort in influencing the effect of anxiety on performance: Testing the conflicting predictions of processing efficiency theory and the conscious processing hypothesis. *British Journal of Psychology*, *98*(3), 411–428. https://doi.org/10.1348/000712606X133047

Zijlstra, F. R. H. (1993). Efficiency in work behaviour: A design approach for modern tools

Appendix

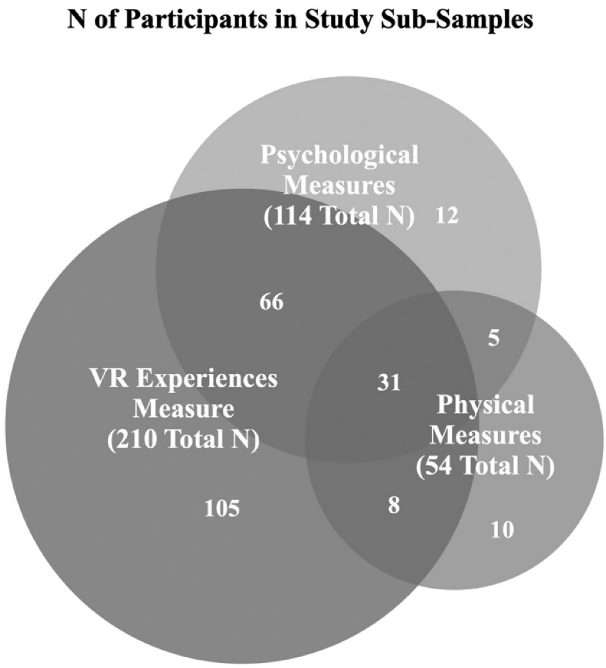

Figure A1. Venn diagram displaying number of participants completing various measures.

Table A1. Descriptive statistics for sub-samples of the primary measures in the study.

	Sex			Age				Experience			
Sub-Sample	Male	Female	Other	Min	Max	Mean	*SD*	Min	Max	Mean	*SD*
Physical Measures	51	3	.	25	55	39.22	7.29	6	28	15.47	5.89
Psychological Measures	110	4	.	25	62	39.74	7.74	3	42	16.24	7.08
VR Experience Measures	201	7	2	25	59	39.11	7.77	3	40	15.00	6.66
Total	227	8	2	25	62	39.39	7.82	3	42	15.16	6.78

Table A2. Descriptive statistics for hierarchical multiple regression model predicting perceived exertion of mental effort in VR.

	Mean	Std. Deviation	N
VAS Mental Effort	50.81	24.62	156
Age	39.38	7.79	238
VR Knowledge	1.95	.83	210
Gaming Frequency	2.12	1.12	210
Prior VR Experience	.62	.49	210
Spatial Presence	3.57	.48	209
Engagement	3.67	.49	209
Negative Effects	2.93	.88	209

Table A3. Pearson correlation coefficients for hierarchical multiple regression model predicting perceived exertion of mental effort in VR.

	1	2	3	4	5	6	7	8
1. Mental Effort	-	-	-	-	-	-	-	-
2. Age	−.116	-	-	-	-	-	-	-
3. VR Knowledge	.066	.005	-	-	-	-	-	-
4. Gaming Frequency	−.218*	−.197*	.350**	-	-	-	-	-
5. Prior VR Experience	.029	.003	−.387**	−.258**	-	-	-	-
6. Spatial Presence	.226*	−.105	.096	−.008	.062	-	-	-
7. Engagement	.364**	−.106	.204*	.087	−.007	.767**	-	-
8. Ecological Validity	.255*	−.157*	.040	.018	.080	.801**	.718**	-
9. Negative Effects	.195*	.041	−.131*	−.180*	.032	−.127*	−.266**	−.140*

* $p < .05$. ** $p < .001$.

Table A4. Descriptive statistics for hierarchical multiple regression model predicting perceived stress in VR.

	Mean	Std. Deviation	N
VAS Stress	36.26	20.28	154
Age	39.40	7.82	236
VR Knowledge	1.94	.83	208
Gaming Frequency	2.10	1.10	208
Prior VR Experience	.63	.48	208
Spatial Presence	3.58	.46	207
Engagement	3.67	.48	207
Ecological Validity	3.64	.50	207
Negative Effects	2.93	.88	207

Table A5. Pearson correlation coefficients for hierarchical multiple regression model predicting perceived stress in VR.

	1	2	3	4	5	6	7	8
1. VAS Stress	-	-	-	-	-	-	-	-
2. Age	.067	-	-	-	-	-	-	-
3. VR Knowledge	.139	.013	-	-	-	-	-	-
4. Gaming Frequency	−.025	−.200*	.342**	-	-	-	-	-
5. Prior VR Experience	−.030	.001	−.380**	−.239**	-	-	-	-
6. Spatial Presence	.275*	−.110	.122*	.056	.030	-	-	-
7. Engagement	.368**	−.105	.214*	.131*	−.026	.759**	-	-
8. Ecological Validity	.333**	−.176*	.090	.112	.034	.785**	.724**	-
9. Negative Effects	.111	.042	−.132*	−.177*	.029	−.143*	−.279**	−.165*

* $p < .05$. ** $p < .001$.