



Kent Academic Repository

Sambrooks, Katie, Sebalo, Ivan, Allen, John, Majid, Shahid, Barrowcliffe, Emma R., Lockerbie, Lona and Gannon, Theresa A. (2026) *Assessing inappropriate fire interest using virtual reality: An investigation in community and clinical populations.* Psychology, Crime & Law . ISSN 1068-316X.

Downloaded from

<https://kar.kent.ac.uk/115946/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1080/1068316X.2026.2729724>

This document version

Publisher pdf

DOI for this version

Licence for this version

CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

Additional information

Versions of research works

Versions of Record

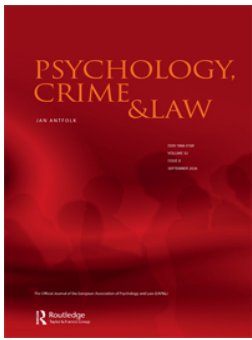
If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).



Assessing inappropriate fire interest using virtual reality: an investigation in community and clinical populations

Katie Sambrooks, Ivan Sebaló, John Allen, Shahid Majid, Emma R. Barrowcliffe, Lona Lockerbie & Theresa A. Gannon

To cite this article: Katie Sambrooks, Ivan Sebaló, John Allen, Shahid Majid, Emma R. Barrowcliffe, Lona Lockerbie & Theresa A. Gannon (16 Sep 2026): Assessing inappropriate fire interest using virtual reality: an investigation in community and clinical populations, *Psychology, Crime & Law*, DOI: [10.1080/1068316X.2026.2729724](https://doi.org/10.1080/1068316X.2026.2729724)

To link to this article: <https://doi.org/10.1080/1068316X.2026.2729724>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 16 Sep 2026.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)



This article has been awarded the Centre for Open Science 'Preregistered' badge.

Assessing inappropriate fire interest using virtual reality: an investigation in community and clinical populations

Katie Sambrooks ^a, Ivan Sebaló^b, John Allen^c, Shahid Majid^d, Emma R. Barrowcliffe^e, Lona Lockerbie^{c,d} and Theresa A. Gannon ^{c,d}

^aSchool of Psychology, Keele University, Staffordshire, UK; ^bSchool of Psychology, University of New York in Prague, Prague, Czech Republic; ^cSchool of Psychology, University of Kent, Canterbury, UK; ^dForensic Care Group, Kent and Medway NHS Care and Social Partnership Trust, Maidstone, UK; ^eSchool of Psychology, Canterbury Christ Church University, Canterbury, UK

ABSTRACT

Inappropriate fire interest is a key risk factor for deliberate firesetting, yet current assessments rely heavily on self-report measures involving imagined scenarios. This two-part study investigated whether virtual reality (VR) offers a more ecologically valid alternative for assessing fire interest. In Study 1, 140 adult males from the community were randomly assigned to one of three conditions: an immersive Virtual Reality Fire Assessment (VRFA), a Televised Fire Assessment (TFA), or an Imagined Fire Assessment (IFA). Measures included heart rate, verbal appraisals, and behavioural engagement. While VR elicited greater presence and more extreme emotional responses, high fire interest did not predict greater interaction in the VRFA. Study 2 replicated the design in a pilot study with a secure psychiatric sample ($N=32$), comparing individuals with and without a history of firesetting. Firesetting individuals provided more positive fire appraisals, and some recalled personal firesetting experiences, suggesting potential clinical insight. However, behavioural measures did not reliably distinguish firesetting individuals from controls. Findings suggest VR is a feasible and ethically acceptable method for assessing fire-related interest, but its diagnostic utility remains limited.

ARTICLE HISTORY

Received 12 April 2026
Accepted 14 August 2026

KEYWORDS

Firesetting; arson; fire interest; virtual reality; VR

Deliberate firesetting and inappropriate fire interest

Deliberate firesetting is a prevalent, problematic behaviour that has devastating consequences spanning fiscal, physical, and psychological impacts. For example, in the year ending March 2025, there were 61,610 deliberate fires attended by Fire and Rescue Services in England (Home Office, 2025a), which resulted in 55 fatalities and 949 injuries (Home Office, 2025b). The National Fire Chiefs Council (NFCC, 2019) estimated the potential annual fiscal loss associated with deliberate firesetting to be up to £11.46 billion in the

CONTACT Katie Sambrooks  K.Sambrooks@Keele.ac.uk  School of Psychology, Keele University, ST5 5BG, Staffordshire, UK

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

UK. Similarly problematic figures are reported internationally (Gannon et al., 2022), emphasising the global scale of this issue and prompting calls to consider deliberate firesetting an international public health concern (Tyler et al., 2019).

Deliberate firesetting is a heterogeneous behaviour associated with a wide range of motivations, including revenge, vandalism, crime concealment, profit, attention seeking, emotional expression, mental disorder, and excitement (Gannon & Pina, 2010). However, despite this diversity, an attraction to, fascination with, or excitement about fire appears consistently across both historical and contemporary classification systems and theoretical accounts of firesetting (Tyler & Gannon, 2020). For example, the most recent theoretical explanation of adult-perpetrated firesetting, the Multi-Trajectory Theory of Adult Firesetting (M-TTAF; Gannon et al., 2012, 2022), recognises inappropriate fire interest as a key psychological vulnerability for firesetting and a distinct firesetting subtype.

There has been a wealth of empirical support for the association between fire interest and deliberate firesetting. For example, Gannon and Barrowcliffe (2012) found that community members who self-reported having engaged in deliberate firesetting (but had not been apprehended for this behaviour) reported increased fascination and arousal to fire when compared to non-firesetting community members. Similarly, Gannon et al. (2013) established that an interest in serious fires (e.g. a hotel or house fire) distinguishes between firesetting prisoners and matched prisoners with no firesetting history. Inappropriate fire interest has also been shown to be a significant risk factor for repeat firesetting. Tyler et al. (2015) found patients with a diagnosed mental disorder and a history of firesetting were 15 times more likely to have set multiple fires if their clinical notes recorded that they held an inappropriate interest in fire.

Given its association with both firesetting behaviour and recidivism, fire interest is routinely considered during firesetting assessments and may help clinicians identify fire-specific treatment needs, formulate offending behaviour, and evaluate risk. Therefore, the accurate measurement of fire interest is vital to inform risk assessment and treatment planning for individuals who have set deliberate fires. Currently, clinicians tasked with assessing inappropriate fire interest typically administer questionnaire measures that require participants to imagine fire-related stimuli and self-report their affect (Gannon et al., 2022). For example, for non-apprehended firesetting individuals, Gannon and Barrowcliffe (2012) developed the Fire Interest subscale of the Fire Setting Scale. It is made up of 10 items (e.g. 'I get excited thinking about fire'), rated on a 7-point Likert scale from 1 ('not at all like me') to 7 ('very strongly like me'). Similarly, the Serious Fire Interest subscale of the Four Factor Fire Scales (FFFS; Ó Ciardha, Tyler, et al., 2015) is comprised of seven items detailing destructive or life-threatening fires (e.g. 'Watching a house burn down' or 'Seeing a hotel on fire in the TV news') to which participants are asked to rate how interested they would be in the situation on a 7-point scale (1 = *upsetting/frightening* to 7 = *exciting, fun, or lovely*).

While these measures are able to distinguish between individuals with a history of firesetting and those without (Barrowcliffe & Gannon, 2016; Gannon et al., 2013) and have clinical utility (Ó Ciardha, Tyler, et al., 2015), they rely heavily on imagined responding and conscious self-report. Self-report measures may be influenced by factors such as socially desirable responding, impression management, and a reluctance to disclose potentially stigmatising interests or experiences (Hildebrand et al., 2018; Paulhus, 1986). In addition, such approaches assume that individuals possess sufficient insight

into their own thoughts, emotions, and behaviour (Schmidt et al., 2015), which may be challenging for some forensic populations (Snowden et al., 2011). More fundamentally, responses to imagined scenarios may not accurately reflect how individuals react when confronted with fire-related stimuli in a more realistic setting. Consequently, there is a need to explore alternative assessment approaches that can capture responses to fire-related stimuli in a more behaviourally observable way.

One alternative to questionnaire-based assessments of fire interest is the fire-specific pictorial Stroop task developed by Gallagher-Duffy et al. (2009). In a sample of male adolescents, those with a history of deliberate firesetting demonstrated greater attentional bias toward fire-related images, shown through slower and less accurate responses when identifying the colour of fire stimuli relative to control participants. Interestingly, this attentional bias did not positively correlate with self-reported fire interest, suggesting that behavioural tasks may capture different aspects of fire salience than traditional self-report measures. However, this research was limited to adolescents, and it remains unclear whether these findings generalise to adults with a history of firesetting. Moreover, while the pictorial Stroop task provides a behavioural measure of attention to fire-related stimuli, participants are still responding to static images presented in a highly controlled laboratory setting. There remains a need to explore assessment approaches that expose individuals to fire-related stimuli in a more realistic and immersive context.

The potential of virtual reality

The emergence of new technologies presents opportunities to develop novel assessment methodologies for deliberate firesetting (Gannon et al., 2022). One potential avenue for new firesetting assessments may be to incorporate Virtual Reality (VR) to display fire-related stimuli. Although there is not a standardised definition of VR within the research literature (Kardong-Edgren et al., 2019), VR typically refers to technology that allows the user to navigate around a computer-generated, three-dimensional environment (Fox et al., 2009; Freeman et al., 2017). The virtual environment is immersive with realistic images and stimuli typically presented via a head-mounted device. These images are continuously rendered and synchronised to the position of the user's head and their body movements are tracked, allowing users to freely move around and interact with items in the virtual environment (Bell et al., 2020). The virtual environments are designed to elicit a sense of presence, wherein the user experiences the virtual setting as real (Diemer et al., 2015; Heeter, 1992). This sense of presence is thought to facilitate more naturalistic responses to virtual stimuli, with research suggesting that the experience of presence is associated with individuals' mental imagery abilities (Iachini et al., 2019). Consequently, VR may offer an advantage over assessment approaches that require participants to imagine scenarios internally, by providing a standardised and immersive representation of those situations.

Over the past decade, the application of VR has expanded and its use within the assessment and treatment of various mental health disorders has increased (Freeman et al., 2017; Geraets et al., 2022). Arguably the most mature field within this research relates to work with clients who have been diagnosed with anxiety disorders (Freeman et al., 2017; Geraets et al., 2021). However, research has also examined VR use in the assessment and treatment of psychosis (Freeman et al., 2008, 2017; Rus-Calafell et al., 2017),

substance-related issues (Segawa et al., 2020), and eating disorders (Clus et al., 2018). It is thought that the application of VR to the assessment and treatment of mental disorders will continue to expand, with possible adoption into regular clinical practice in the near future (Geraets et al., 2021).

The application of VR to forensic contexts

In addition to being a useful tool in the care of several mental disorders, VR appears to have significant potential for the assessment and treatment of individuals with a history of offending (Benbouriche et al., 2014; Fromberger et al., 2018; Kip et al., 2018; Ticknor & Tillinghast, 2011). While research in this area has been limited, several researchers have described the possible benefit of incorporating VR into forensic assessments of an individual's treatment needs or risk (see Sygel & Wallinius, 2021). One of the primary reasons for this is that VR offers the opportunity to expose individuals to stimuli that would otherwise be difficult to access due to security or ethical issues (Kip et al., 2019). VR allows clinicians to safely observe their clients within simulated offence-related situations (Cornet & Van Gelder, 2020). VR has successfully been used to investigate the role of expertise in the decision-making of individuals convicted of burglary by placing them in a virtual environment depicting a residential neighbourhood (Nee et al., 2019).

Research examining other offence-related interests has also highlighted the value of behavioural responses to relevant immersive stimuli. For example, Chartier et al. (2006) demonstrated that patterns of gaze behaviour towards visual sexual stimuli could be used to distinguish between those with a history of sexual offending and those without. This suggests that observing individuals' responses to salient stimuli may provide useful assessment information regarding underlying interests that may not be fully captured through self-report alone. In addition, Renaud and colleagues (2014) investigated the use of VR to expose individuals with a history of sexual offending to unclothed child avatars whilst using penile plethysmography. The erectile responses of 42 control participants and 22 individuals with a history of sexual offences against children were examined. While both the VR stimuli and traditional auditory stimuli generated responses that were significantly different across the participant groups, classification was more accurate using responses from the VR stimuli. Therefore, Renaud et al. concluded that VR represented a superior paradigm for assessing inappropriate sexual arousal, relative to traditional auditory stimuli. These findings underscore VR's potential to enhance forensic assessment in contexts where traditional methods fall short, though further research is needed to determine its effectiveness across a broader range of offending behaviours.

Virtual reality and firesetting

To date, no research has examined the utility of VR for the assessment and treatment of deliberate firesetting. However, clinicians working with individuals who have a history of firesetting hold positive views about the potential use of VR for firesetting. In a survey conducted by Sambrooks et al. (2022), 78% ($n = 57$) of clinicians primarily recruited from the UK agreed that a virtual environment involving a fire had the potential to add value to assessments. Presenting adults who have set deliberate fires with fire-related stimuli via VR could potentially represent a superior methodology for assessing fire

interest, relative to the current protocol of having them imagine fire-related situations. However, due to the lack of research in this area, it is currently unknown whether an immersive stimulus would be necessary to gain this greater insight, or whether non-immersive pictorial stimuli would be sufficient. There is a need for empirical evidence to support the assertion that VR represents an improved paradigm for the assessment of fire-specific treatment needs, particularly inappropriate fire interest. Therefore, the current research compares responses to one of three fire assessments: an immersive fire assessment presented through VR (VRFA), a non-immersive two-dimensional televised fire assessment (TFA), and an imagined fire assessment (IFA). The first study examines this within a community sample, while Study 2 represents a pilot examination of its use with a secure psychiatric sample.

Study 1

Open science framework preplanned hypotheses

Our predefined hypotheses are publicly available via the open science framework repository (<https://osf.io/hg6wj>). We expect that, because a VRFA has an immersive component, this will generate more extreme (i.e. positive or negative) verbal appraisals of the fire and higher physiological heart rate levels relative to the non-immersive TFA and IFA conditions. We also expect that the VRFA will generate more positive appraisals of fire (i.e. indicative of fire interest) and higher physiological heart rate levels (indicating higher levels of simulation) relative to the non-immersive TFA and IFA conditions for individuals who hold higher levels of fire interest. Finally, we expect the VRFA will generate more movement towards the fire and interaction with the fire for individuals who hold higher levels of fire interest within the VRFA condition and relative to the non-immersive TFA condition. In addition to these key predictions, we will also explore (1) the levels of presence and pleasantness associated with the VRFA relative to the non-immersive conditions, (2) the safety of the VRFA procedure in terms of self-reported adverse experiences two weeks post VRFA, and (3) any possible relationship between impression management and variables such as fire interest, feelings of presence, and pleasantness.

Method

We report our method in line with our publicly available Open Science Framework study plan. The study was approved by the University of Kent Ethics Committee (ID 201915519714255616).

Design

Participants were randomly allocated to one of three assessment conditions using graphpad.com, with the aim of achieving approximately equal numbers of participants in each condition: the virtual reality fire assessment (VRFA), the televised two-dimensional fire assessment (TFA), or the imagined fire assessment (IFA). Those in the VRFA condition entered a virtual reality environment involving fire, while those in the TFA and IFA conditions saw a televised version of the same fire scene or were asked to imagine it, respectively.

Participants

One hundred and forty adult males were recruited from the local community using online and offline advertisements requesting participation in a 'study examining assessment measures' in exchange for £10. Recruitment materials described the study as examining different assessment methods. However, when participants were contacted to arrange a testing session, they were provided with additional information about the fire-related nature of the assessment. Participants were also informed that individuals with a history of PTSD or epilepsy should not participate due to the potential for distress or adverse reactions associated with the fire-related stimuli and virtual reality equipment. Screening for these conditions was conducted via participant self-report during the consent process. No participants were excluded. Participants were aged between 18 and 78 years, were predominantly students, and the majority held some or no experience of virtual reality (see Table 1 for further information on demographics). Only males were recruited because deliberate firesetting research has historically focused on male populations, who account for the majority of identified firesetting cases (Gannon & Pina, 2010). The present study was intended as an initial evaluation of the assessment methodology within a relatively homogeneous sample before extending this work to more diverse populations.

Measures

Virtual reality fire assessment (VRFA). The virtual reality prototype environment was created by Mindwave Ventures using Unity software and ran on a RGB 144 Hz Scan 3XS LG15 Vengeance Max Q gaming laptop. Participants experienced the virtual reality

Table 1. Distribution of demographic and pre-experience measures across the VRFA, TFA, and IFA groups.

Variables	VRFA (<i>n</i> = 40) <i>M</i> (<i>SD</i>)	TFA (<i>n</i> = 40) <i>M</i> (<i>SD</i>)	IFA (<i>n</i> = 40) <i>M</i> (<i>SD</i>)	Total (<i>n</i> = 120) <i>M</i> (<i>SD</i>)	<i>F</i> / χ^2	<i>df</i>	<i>p</i> - value
Age	31.92 (16.43)	29.08 (15.75)	29 (14.48)	30 (15.5)	0.46	2	0.63
VR experience							
Moderate to Lots (1 + hour)	7 (36.84)	7 (36.84)	5 (26.32)	19 (100.00)	2.19	4	0.70
Some (<1 h)	19 (36.54)	18 (34.62)	15 (28.85)	52 (100.00)	NA	NA	NA
No experience	14 (28.57)	15 (30.61)	20 (40.82)	49 (100.00)	NA	NA	NA
Ethnicity							
All other	19 (29.69)	22 (34.38)	23 (35.94)	64 (100.00)	0.87	2	0.65
White & Irish	21 (37.50)	18 (32.14)	17 (30.36)	56 (100.00)	NA	NA	NA
Employment							
Full-time	13 (46.43)	4 (14.29)	11 (39.29)	28 (100.00)	7.61	4	0.11
Part-time	7 (41.18)	6 (35.29)	4 (23.53)	17 (100.00)	NA	NA	NA
Retired/Out of Work/ Unable to Work	20 (26.67)	30 (40.00)	25 (33.33)	75 (100.00)	NA	NA	NA
Measures							
Impression Management	4.42 (3.05)	4.42 (3.05)	4.8 (3.58)	4.55 (3.22)	0.18	2	0.84
WASI-II Full Scale IQ	108.03 (10.46)	110.17 (13.31)	104.85 (13.07)	107.68 (12.44)	1.88	2	0.16
Fire Interest	35.48 (13.31)	35.38 (10.87)	34.45 (13.35)	35.1 (12.46)	0.08	2	0.92

Note. In line with our preregistration, categories for categorical variables were merged where expected cell frequencies were <5. For ethnicity, all categories other than British/Irish were merged. For employment, the categories of Retired, Out of work, and Unable to work were merged. For virtual reality experience, the categories of Moderate, Quite a bit, and Lots were merged.

environment via a cabled Oculus Rift headset which provided a high-fidelity view of their virtual environment with integrated audio. The Oculus Rift features per-eye displays with a 1080×1200 resolution, running at 90 Hz refresh rate, 360-degree positional tracking (via three external base stations).

The virtual experience was 6 min in total and was made up of a plain bedsit containing a sofa, armchair, single bed, dining table and chairs, and basic kitchenette (see [Figure 1](#)). Five objects were also displayed in the virtual bedsit (a newspaper, magazine, roll of paper towel, deodorant can, and fire extinguisher). Following familiarisation with the environment (3 min), a virtual fire appeared – without any apparent cause – on the bed. This fire spread across the bed and was accompanied by black smoke and the sound of crackling for a further 3 min (see [Figure 1](#)). Participants were able to explore their virtual environment naturalistically through walking. They also carried a tracked hand controller in their dominant hand which enabled them to pick up and drop/throw objects using a button press on the Oculus Touch controller (see [Figure 2](#)). Four of these items were flammable (a newspaper, a magazine, an aerosol can, and kitchen roll) and the remaining item was a fire extinguisher. As with the other items, this fire extinguisher could be picked up and moved around, but it did not function (i.e. participants could not spray it).

Televised fire assessment (TFA). A televised non-immersive version of exactly the same virtual reality bedsit described above ran on the same gaming laptop. This was experienced by participants through viewing the 15.6' 144 Hz FHD laptop screen and listening to it using a standard set of headphones. Participants were able to explore the televised version of the bedsit through using the controls on the keyboard and mousepad. They could also use this controller to pick up and drop/ throw objects using a button press.

Imagined fire assessment (IFA). The imagined version of the bedsit was presented verbally by the researcher using the following pre-defined script:



Figure 1. Studio flat and fire as depicted in the TFA and VRFA.

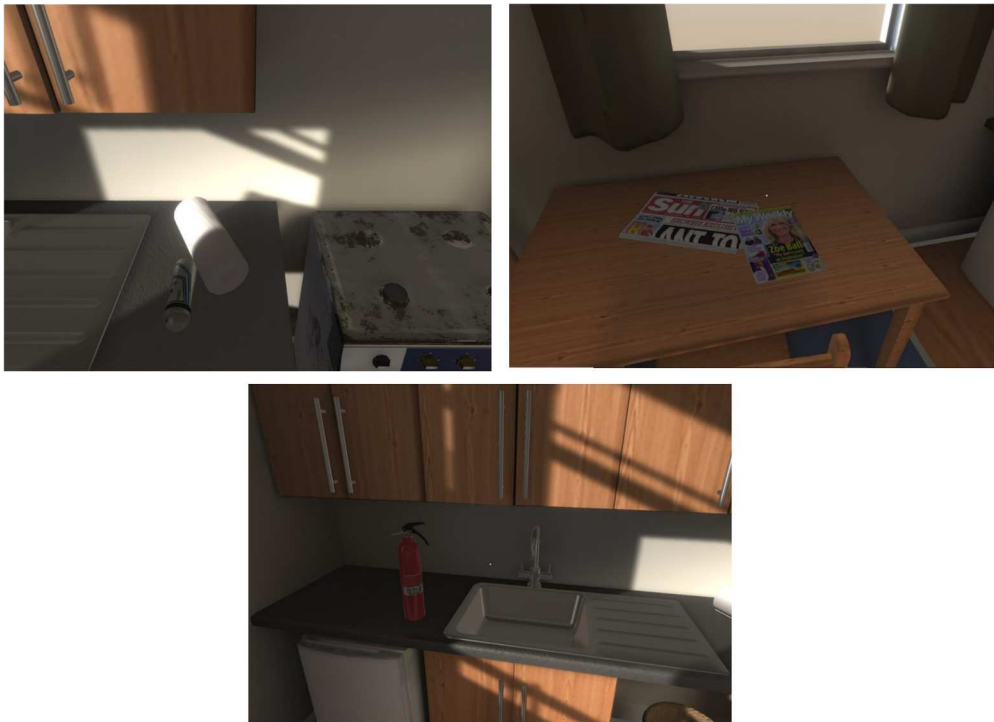


Figure 2. Items within the virtual environment.

I would like you to imagine the inside of a home. It is a type of studio flat or bedsit with a blue carpet and a dark blue rug in the centre of the room. To the right of you there is a small wooden dining table with a Sun newspaper and a 'My Weekly' magazine on it. Above the table is a window with olive green curtains. In front of you there is a single bed with a metal bed frame, a pillow on it and olive-green sheet. To the left of the bed there is a comfortable looking grey, three-seater sofa and a waste paper basket. Further to the left is the door into the bedsit and as you turn to the left again there is a large comfortable red arm-chair. Directly opposite the bed is a small kitchen area with a fridge to the left and a free standing, old and shabby looking cooker to the right. Between the fridge and the cooker is a sink. On the work surface above the fridge is a red fire extinguisher which is standing upright. On the other side of the worktop to the right of the sink there is an aerosol can lying on its side and some kitchen roll also lying on its side. There are kitchen units above the sink and the work tops. To the right of the cooker there is a wooden free standing television unit with one shelf with a folded newspaper and on the top of the unit there is a television and the remote control.

Do you have any questions before we begin?

After 3 min, when the fire starts:

I would like you to continue to imagine the inside of the home I described. This time there has been a development. A fire has started in the centre of the bed. The flames are burning brightly and the fire quickly spreads across the whole bed. Shortly after the bed is covered in flames you can see the sparks flying and black plumes of smoke and notice that the three-seater sofa has now also caught alight. Again, the fire spreads quickly and some flames travel to the floor. Then the comfy looking red arm chair also catches fire.

Pre-experience measures

Several questionnaire measures were completed immediately prior to the fire experience. Where relevant, we report reliabilities according to George and Mallery's (2018) criteria: $\geq .90$ excellent, $.89$ to $\geq .80$ good, $.79$ to $\geq .70$ acceptable, and $.69$ to $.60$ marginal.

Demographics. A demographic questionnaire collected information about participants' age, ethnicity, employment status, and previous experience with VR. The order in which the remaining pre-experience measures were presented to participants was randomised.

Impression management. The Impression Management subscale of the Paulhus Deception Scales (Paulhus, 1998) is a recognised measure of intentional impression management for both forensic and non-forensic individuals. The 20 items (e.g. 'I never swear') are rated on a 5-point Likert scale from 1 ('not true') to 5 ('very true'). Ten of these are uncommon but desirable behaviours (e.g. 'I never swear'), and ten items describe undesirable behaviours (e.g. 'There have been occasions when I have taken advantage of someone'). These latter items are reverse scored such that higher scores on this measure indicate a tendency to provide overly positive self-descriptions. Paulhus (1998) reported excellent internal consistency with both the general population ($\alpha = .84$) and an imprisoned sample ($\alpha = .84$). In this study, internal consistency was acceptable (.76).

Cognitive ability (WASI-II). The short two-subtest (vocabulary and matrix reasoning) format of the Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler, 2011) was used to reliably estimate participants' full-scale intelligence quotient. The Vocabulary subtest is made up of 31 items (3 picture items and 28 verbal items). Individuals are required to verbally define and/or describe the word presented to them. This subtest measures an individual's word knowledge, verbal concept formation, learning ability and degree of language development. The Matrix Reasoning subtest has 30 items. The individual is asked to complete a series of incomplete matrices by selecting one of five response options. This subtest measures an individual's visual intelligence, classification and spatial ability, perceptual organisation, and knowledge of part-whole relationships. Wechsler (2011) reported excellent internal consistency, with reliability coefficients of .92 for the Vocabulary subtest, .90 for the Matrix Reasoning subtest, and .94 for FSIQ-2 composite score.

Fire interest. The Fire Interest subscale of the Fire Setting Scale (Gannon & Barrowcliffe, 2012) is a short and reliable measure of fire interest for individuals in the community. The 10 items (e.g. 'I like watching fire') are rated on a 7-point Likert scale from 1 ('not at all like me') to 7 ('very strongly like me'). High scores indicate greater levels of fire interest. Internal consistency was excellent in the current sample (.91).

Fire experience measures

During each of the fire experiences a number of measures were captured.

Heart rate. In all conditions, participants' heart rate was recorded at baseline and throughout the fire experience using a pulse oximeter. Change in heart rate was then examined.

Free response verbal fire appraisal. Throughout the fire experience (virtual, televised, or imaginary), participants were encouraged to verbalise, aloud, all thoughts and feelings. A blind rater coded these into positive (e.g. 'that looks lovely') or negative (e.g. 'I'm scared') fire appraisals. For each participant, negative appraisals are taken away from positive ones to produce one **overall verbal fire appraisal score**. Values > 0 indicate a greater number of positive fire appraisals relative to negative ones (i.e. a positive fire appraisal) while values < 0 indicate a greater number of negative fire appraisals relative to positive ones (i.e. a negative fire appraisal).

Interaction with the fire. In the VR and TV fire experiences, the virtual environment allowed further measures to be collected which examined participants' interaction with the fire. These were not collected for the Imagined fire experience. Movement within the virtual environments was examined using the participants' average distance travelled from and to the fire and their closest distance to the fire. Participants' interaction with items was examined in terms of the number of flammable items picked up and the number of flammable items burned. The time the fire was watched was recorded using the number of seconds the fire was in the centre of the field of view.

Post-experience measures

Immediately after the fire experience, participants completed two post-experience measures.

Presence. The Sense of Presence Scale (Slater et al., 1998) is a 6-item measure of Virtual Reality immersion. This measure was adapted slightly for the purpose of our study so that questions assessed immersion generally for the bedsit fire rather than VR per se (e.g. During the time of the experience, did you often just think to yourself that you were actually just in a laboratory or did the experience of the house with the fire overwhelm you?). Higher scores on this measure indicate a greater experience of presence. In this study, internal consistency was good (.84).

Pleasantness. A visual analogue scale from 1 ('not at all pleasant') to 10 ('extremely pleasant') was created to examine participants' self-reported pleasantness in relation to their virtual, televised, or imagined fire experience.

Follow-up interview. Participants experiencing the virtual reality fire assessment received a phone call or email two weeks after their experience and were invited to partake in a short semi-structured interview probing for any unintended adverse experiences. Participants were asked 'Looking back at the study, has it had any impact on the way you feel at all? For example, do you feel positive/negative/neutral about having taken part in the study? Why is this?' The overall interview was coded as positive, negative, or neutral.

The two-week timeframe was selected to identify any short-term adverse effects associated with participation whilst minimising participant attrition.

Procedure

Once informed consent was obtained, participants completed the pre-experience questionnaire measures individually beginning with the demographic questionnaire. Subsequent measures were presented in a randomised order to counteract any ordering effects. All measures were read aloud to the participant to ensure adequate comprehension and understanding. All participants immediately completed a familiarisation stage that lasted for three minutes. For those in the VR condition, familiarisation consisted of participants engaging in the virtual studio flat without the fire present so that they could become familiar with the environment and the associated equipment. They were encouraged to use the Oculus Touch controller to pick up objects and move around the room. In the televised non-immersive condition, participants were shown how to navigate around the televised studio flat and interact with objects using the controls on the keyboard and mousepad. In the imagined condition, participants were asked to imagine the studio flat without the fire present. Across all versions of the fire experience, a standard set of instructions was developed to encourage individuals to explore their virtual/televised/imagined environment and verbalise their thoughts and feelings:

You can pick up objects and put down objects using the controller/your imagination. Okay, now you are ready, we would like you to explore your environment, and practice picking up and putting down an object in the room using the controller/your imagination. As you explore your environment, please speak aloud so that we know what you are thinking and feeling.

Once the three minutes of familiarisation elapsed, participants were informed that the fire experience would now begin.

Please continue to explore the environment however you wish. Please remember that we would like you to speak aloud so that we know what you are thinking and feeling.

Each fire assessment lasted for three minutes. Participants were able to stop the fire assessment at any time if they wished. Following the fire assessment, participants completed the measures of presence and pleasantness (in a randomised order). Throughout the assessment, participants were monitored by a researcher and reminded that they could withdraw without penalty at any stage. Following participation, all individuals received a full debrief and were provided with information about relevant support services should they experience any distress following the study. All study procedures were conducted within a university psychology laboratory. Assessments and interviews were administered by a member of the research team with experience conducting psychological research with forensic and community populations. Approximately two weeks later, participants were invited by the same researcher to complete the follow-up interview via phone or email.

Data analysis

Obtained data was analysed using R 4.3.1 and followed the steps specified in our pre-registration document. This included testing for potential covariates using univariate

and Chi-Squared analyses as well as moderated regressions to verify the hypotheses. However, in a departure from our regression, analysis of variance (ANOVA) was performed before moderated regressions to first confirm that there was a difference between groups. Post hoc comparisons with Bonferroni adjustment were also performed to identify the location of between group differences. The primary data for this research are not available due to lack of consent for data sharing from participants.

Results

Pre-experience measures

Table 1 presents the distribution of demographic and clinical factors across the three conditions. There were no notable differences across the conditions on demographic or clinical factors. In particular, there were no significant differences between participants across conditions on levels of fire interest or impression management. Across conditions, the mean level of fire interest ($M = 35.10$) was comparable to that reported by community participants with a self-reported history of deliberate firesetting ($M = 34.45$) and notably higher than that reported by community participants with no history of deliberate firesetting ($M = 29.34$) in Gannon and Barrowcliffe (2012). This finding suggests that, despite being recruited from the general community, participants reported levels of fire interest consistent with those observed among individuals who had previously engaged in deliberate firesetting.

Item 13 on the Impression Management scale ('I sometimes drive faster than the speed limit') had four missing values. Consequently, Little's MCAR test was performed. It was not significant suggesting that these values were missing completely at random ($p = .11$). Participants who did not provide a response to this item indicated that they did not drive and were therefore unable to answer the question. Because of this, predictive mean matching was used to impute values for these four participants. The Impression Management and Fire Interest subscales showed a weak, non-significant relationship, $r = -0.12$, $p = .18$. Thus, impression management was not included in any further analyses as a covariate.

Fire experience measures

Heart rate. An ANOVA was performed to assess difference in changes of heart rate expressed in percentages (see Table 2). However, as there was no significant difference between the groups no follow up tests were performed.

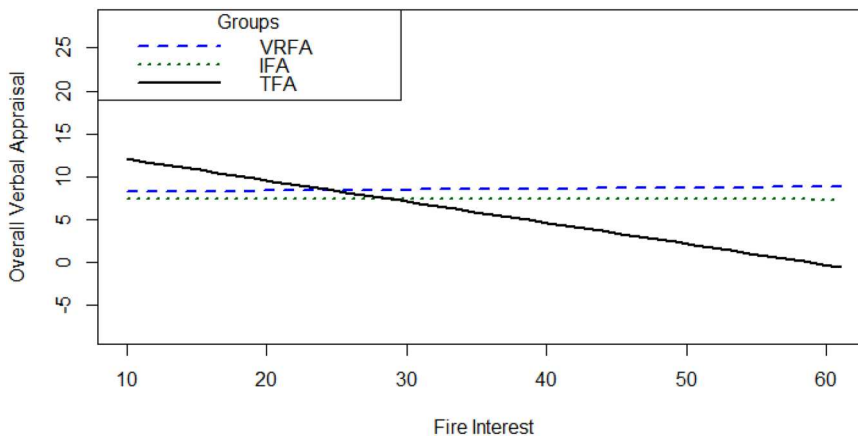
Verbal fire appraisals. ANOVA indicated that there were significant differences between conditions in Overall Verbal Fire Appraisal (see Table 2). Post hoc Bonferroni-adjusted comparisons specified that only the VRFA and TFA conditions differed with participants in the VRFA condition indicating a greater number of negative (relative to positive) fire appraisals ($p = .02$, see Table 2). Furthermore, the moderated regression model was a good fit for the data, adjusted $R^2 = 0.14$, $p < .001$ and the interaction terms between IFA and VRFA groups and Fire Interest was significant (See Table 3). A follow up simple slopes analysis revealed that fire interest had a significant negative association with overall verbal fire appraisal only in the TFA group (slope -0.25 , $p < .001$), but not in the VRFA group (slope 0.01 , $p = .82$), or in the IFA group (slope 0 , $p = .99$) (Figure 3).

Table 2. Means, standard deviations, and one-way ANOVA statistics for key study variables.

	VRFA (<i>n</i> = 40) <i>M</i> (<i>SD</i>)	TFA (<i>n</i> = 40) <i>M</i> (<i>SD</i>)	IFA (<i>n</i> = 40) <i>M</i> (<i>SD</i>)	<i>F</i>	<i>df</i>	<i>p</i> value	partial η^2
Heart rate change	−1.9 (13.61)	−0.46 (9.2)	0.93 (11.2)	0.61	(2,117)	.55	0.01
Overall verbal fire appraisal	8.55 (4.55)	5.82 (5.44)	7.38 (2.77)	3.83	(2,115)	.03	0.06
Movement							
Average distance travelled to fire	24.09 (8)	37.2 (20.65)	–	14.01	(1,78)	.00	0.15
Average distance travelled from fire	2.22 (0.2)	2.15 (0.32)	–	1.42	(1,78)	.24	0.02
Closest distance participants got to fire	1.59 (0.32)	0.48 (0.56)	–	117.28	(1,78)	< .001	0.60
Interaction							
Flammable objects picked up	1.15 (1.21)	1.43 (1.43)	–	0.86	(1,78)	.36	0.01
Flammable objects burned	0.12 (0.52)	0.5 (1.01)	–	4.35	(1,78)	.04	0.05
Watch time	36.35 (16.28)	25.18 (13.29)	–	11.31	(1,78)	< .001	0.13

Table 3. Moderated regression on overall verbal appraisal.

Effect	Estimates	SE	Statistic	<i>p</i>
Intercept (TFA)	14.54	2.27	6.41	< .001
VRFA	−6.40	2.95	−2.17	.03
IFA	−7.13	2.91	−2.45	.02
Fire Interest	−0.25	0.06	−4.03	< .001
VRFA x Fire Interest	0.26	0.08	3.27	.001
IFA x Fire Interest	0.25	0.08	3.12	.002
Observations	118			
<i>R</i> ² / <i>R</i> ² adjusted	0.18 / 0.15			

**Figure 3.** Simple slopes analysis for overall verbal appraisals.

Interaction with the fire. Next, we investigated how participants from TFA and VRFA groups interacted with the fire. As Table 2 shows, there was a significant difference in the average distance travelled to fire, the closest distance participants got to the fire,

number of flammable objects burned, and the time spent watching the fire. Table 4 shows the fit of each model, (adjusted $R^2 = 0.23$, $p < .001$, adjusted $R^2 = 0.63$, $p < .001$, adjusted $R^2 = 0.17$, $p < .001$, adjusted $R^2 = 0.15$, $p < .01$ for each model respectively) and the interaction effects. Given that the interaction between VRFA and Fire interest had the p value above .1 for watch time, it was the only model that was not followed by a simple slopes analysis. Meanwhile, for average distance travelled to fire, a simple slopes analysis showed the strength of the association between fire interest and distance travelled to fire was significant only in the TFA (slope 0.72, $p < .01$), but not in the VRFA (slope 0.03, $p = .87$) (Figure 4). For the closest distance that participants got to the fire, the simple slopes analysis yielded a significant negative effect of fire interest in the TFA group (slope -0.02 , $p < .01$), but not in the VRFA group (slope 0, $p = .53$) (Figure 5). Likewise, fire interest had a significant positive effect on the total number of burnt objects in the TFA group only (slope 0.04, $p < .001$), but not in the VRFA group (slope 0, $p = .93$) (Figure 6).

Post experience measures

Prior to investigating the differences in pleasantness and presence rating between the groups, both were correlated with the Impression Management Scale. The correlation between impression management and pleasantness was weak and non-significant ($r = -0.09$, $p = .33$), but the correlation between impression management and sense of presence was moderate and significant ($r = 0.26$, $p < .01$). As Table 5 shows, while the reported pleasantness of the experience did not differ between the groups, the feeling of presence was significantly different between the groups, when Impression Management was taken into account. Specifically, participants in the VRFA condition reported more presence than those in the TFA condition, $p = .03$. The majority of participants in the VRFA condition (60%, $n = 24$) completed the follow-up interview. The majority of these interviews were coded as positive ($n = 17$), with seven neutral and no negative responses.

Discussion

The primary aim of Study 1 was to evaluate whether an immersive Virtual Reality Fire Assessment (VRFA) could more effectively elicit indicators of inappropriate fire interest than two traditional methods: a non-immersive televised fire assessment (TFA) and an imagined fire assessment (IFA). We hypothesised that the immersive VRFA would result in more extreme verbal appraisals, greater physiological arousal, and increased behavioural engagement with fire, particularly among individuals with higher levels of self-reported fire interest. This study represents a foundational step in assessing the feasibility and utility of VR in fire-related forensic assessments.

Verbal appraisal

Verbal fire appraisals varied significantly across assessment conditions, with participants in the VRFA condition making more negative verbal appraisals (relative to positive appraisals) than those in the TFA condition. This finding suggests that immersion may amplify the perceived threat or seriousness of the scenario, eliciting more aversive verbal reactions. This aligns with previous work indicating that immersive VR environments can heighten emotional salience (Diemer et al., 2015), particularly in simulations involving potential harm or risk.

Table 4. Moderated regression models.

Effect	Average Distance Travelled to Fire			Distance Closest to Fire			Total Number of Objects Burned			Time Spent Watching the Fire		
	Estimates	SE	Statistic	p	Estimates	SE	Statistic	p	Estimates	SE	Statistic	p
Intercept (TFA)	11.69	8.08	1.45	.15	1.14	0.24	4.80	<.001	-0.97	0.41	-2.39	.02
VRFA	11.39	10.53	1.08	.28	0.56	0.31	1.81	.08	1.07	0.53	2.01	.05
Fire Interest	0.72	0.22	3.30	.001	-0.02	0.01	-2.90	.005	0.04	0.01	3.78	<.001
VRFA x Fire Interest	-0.69	0.28	-2.45	.02	0.02	0.01	1.85	.07	-0.04	0.01	-2.87	.005
Observations	80				80				80			
R ² / R ² adjusted	0.26 / 0.23				0.64 / 0.63				0.20 / 0.17			
										0.18 / 0.15		

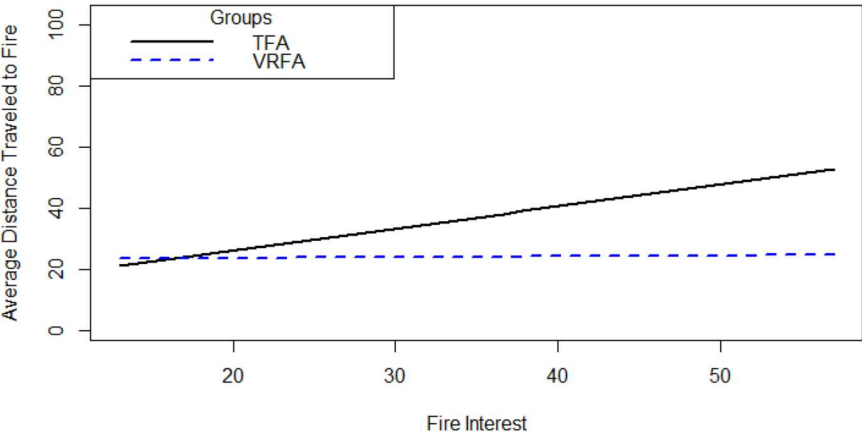


Figure 4. Simple slopes analysis for average distance travelled to the fire.

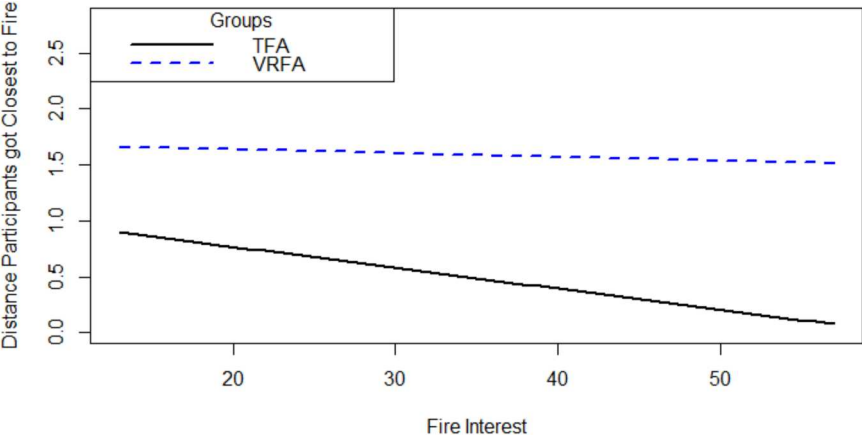


Figure 5. Simple slopes analysis for closest distance to the fire.

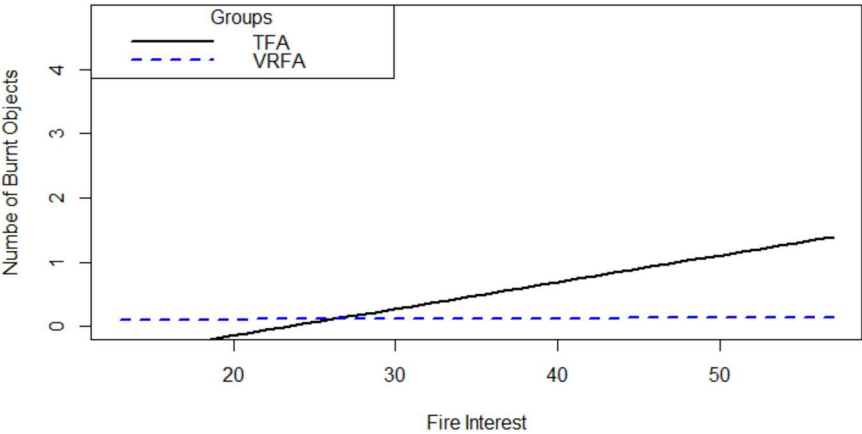


Figure 6. Simple slopes analysis for number of objects burned in the fire.

Table 5. Means, standard deviations, and one-way ANOVA statistics for ratings of pleasantness and presence.

	VRFA (<i>n</i> = 40)	TFA (<i>n</i> = 40)	IFA (<i>n</i> = 40)	<i>F</i> Statistic	df	<i>p</i> value	partial η^2
Pleasantness Difference Between Groups							
Pleasantness	4.22 (2.8)	4.1 (2.54)	4.19 (2.69)	0.02	(2, 117)	.98	0.00
Presence Difference between Groups							
Presence	23.4 (7.6)**	18.88 (7.99)**	20.75 (7.79)	3.63	(2, 116)	.03	0.64
Impression Management	4.42 (3.05)	4.42 (3.05)	4.8 (3.58)	8.79	(1,116)	<.001	0.07

** = significant, $p < .05$. All tests are one-sided.

However, the relationship between fire interest and verbal appraisals did not follow the predicted pattern. While we expected individuals higher in self-reported fire interest to produce more positive verbal responses – particularly in the VRFA condition – this was not supported. In fact, fire interest was only significantly associated with verbal appraisal scores in the TFA condition, where higher fire interest predicted lower overall verbal fire appraisal scores (i.e. more negative appraisals). No such relationship was observed in the VRFA or IFA groups. One plausible explanation for this pattern is the moderate level of realism and emotional intensity offered by the TFA condition. While the TFA involved audiovisual footage of fire, it was not immersive, and did not engage the participant's body or spatial awareness in the same way as the VRFA. Research has shown that non-immersive video formats elicit weaker emotional arousal and a lower sense of presence compared to immersive VR (Diemer et al., 2015; Slater & Sanchez-Vives, 2016). The visually realistic – but psychologically distant – fire footage may have activated ambivalence or caution, especially in individuals aware of the risks or consequences associated with fire-related thoughts or behaviours. This could result in more negative verbal appraisals, despite underlying interest.

In contrast, the VRFA and IFA conditions may have triggered stronger, more uniform affective responses across participants. Immersive VR in particular has been shown to produce heightened physiological arousal, presence, and emotional engagement, even when the content is controlled across formats (Freeman et al., 2017; Geraets et al., 2021). Under such conditions, overt expressions of personal interest may be suppressed or overridden by more immediate affective reactions (e.g. fear, awe, anxiety). This may explain why no significant association emerged between fire interest and verbal appraisals in these conditions, despite theoretical expectations. In summary, the TFA condition may have provided the optimal balance of realism and psychological distance for individual differences in fire interest to be expressed in verbal appraisals. The immersive VRFA and IFA conditions may have dampened this effect by evoking stronger, more homogenous emotional responses that masked underlying differences in fire interest.

Interaction with the fire

A similar pattern of results were apparent in the interaction measures. While there were significant differences across the VRFA and TFA in terms of movement, engagement with items and watch time, fire interest was associated with approach behaviour only in the TFA condition. This suggests that immersive realism may discourage exploration in individuals who otherwise demonstrate fascination with fire-related cues. While previous work in anxiety-related VR studies (e.g. Freeman et al., 2008) has shown that perceived threat leads to avoidance behaviours, our findings suggest that similar dynamics may

be at play in the assessment of morally or socially sensitive behaviours such as firesetting. The lack of association between fire interest and movement in the VRFA condition likely reflects heightened salience and perceived immediacy of risk, which could override internal drives to explore or engage.

Consistent with the movement data, participants in the TFA condition with high fire interest were more likely to burn flammable objects, whereas no such pattern was observed in the VRFA. This absence of behavioural indicators in the immersive condition challenges the presumption that VR necessarily elicits more ecologically valid engagement. This is consistent with recent arguments by Corneille and Gawronski (2024), who caution that implicit measures – such as observed behaviour – can often underperform compared to self-reports, particularly when context introduces ambiguity or performance constraints. Rather than acting on their fire-related interest, participants in the VRFA may have experienced a conflict between internal interest and external realism, leading to behavioural suppression. While previous studies using VR for sexual interest assessment (e.g. Renaud et al., 2014) have noted increased engagement with immersive stimuli, our findings suggest that this effect may not generalise to all forensic domains, particularly when risk-related cues (like fire) are presented in high-fidelity formats.

Presence and pleasantness

Participants reported significantly greater presence in the VRFA condition than the TFA, replicating established findings on the immersive capacity of VR (Slater et al., 1998). This finding reinforces the notion that VR systems – particularly those incorporating head-mounted displays and responsive environmental cues – can elicit a strong sense of ‘being there,’ even in simulated high-risk or morally complex contexts. In contrast, pleasantness ratings did not differ significantly across conditions, indicating that realism alone does not dictate affective experience. While VRFA was perceived as more immersive, it was not necessarily experienced as more emotionally engaging or enjoyable. However, no participants in the VRFA reported a negative experience during the follow-up interview. This is particularly noteworthy given that this was the first use of a fire-based VR scenario, where the potential for distress or discomfort was high due to the inherently threatening nature of the stimulus. The absence of reported negative experiences suggests that the simulation was effective in balancing realism with psychological safety, allowing participants to participate meaningfully without being overwhelmed. This finding supports the feasibility of using VR to explore fire-related scenarios in a controlled and ethically responsible manner.

Limitations

Several limitations should be considered when interpreting these findings. First, although the VR fire assessment was associated with greater presence and emotional engagement than the televised and imagined fire assessments, it did not consistently outperform the televised fire assessment across all outcome measures. This suggests that increased immersion does not necessarily translate into improved assessment performance. It is possible that the realism of the VR environment altered participants’ responses in ways that reduced the expression of fire-related thoughts, feelings, or behaviours. Future research should therefore continue to examine the circumstances under which immersive environments provide advantages over less immersive assessment approaches. In

addition, the VR fire assessment represented a novel assessment environment and its psychometric properties have not yet been established. Responses were assessed following a single exposure to the fire stimuli and it therefore remains unclear whether the observed effects would remain stable across repeated administrations. Future research should examine the reliability and validity of VR-based fire assessments across multiple time points and settings.

A further limitation relates to the administration of questionnaire measures. To ensure adequate comprehension, all measures were read aloud to participants. While this approach standardised administration and reduced the likelihood of misunderstanding, it may also have increased participants' awareness of being observed and influenced responding on socially sensitive measures, such as fire interest and impression management.

Finally, participants were recruited from the general community rather than on the basis of elevated fire interest or a history of deliberate firesetting. Although participants reported levels of fire interest comparable to those reported by community individuals with a self-reported history of deliberate firesetting (Gannon & Barrowcliffe, 2012), the findings may not generalise to populations with more severe fire-related difficulties. Furthermore, only male participants were recruited. While males account for the majority of identified deliberate firesetting cases (Gannon & Pina, 2010), emerging evidence suggests that females who set fires may differ from males in terms of their motivations, psychological characteristics, and offence patterns (Alleyne et al., 2016; Nanayakkara et al., 2020). Consequently, it remains unclear whether females would respond similarly to the assessment paradigms evaluated in the present study. Future research should therefore examine the utility of these assessment approaches among individuals with higher levels of fire interest, known histories of deliberate firesetting, and more diverse demographic characteristics.

Conclusion

The findings of Study 1 provide several insights for the future use of VR in the assessment of firesetting behaviour. First, while immersive environments reliably enhance feelings of presence, this does not automatically translate into more overt behavioural indicators of fire interest, such as visible excitement, fixation, or engagement with fire-related stimuli. In fact, the heightened realism of a virtual fire scenario may increase self-monitoring, anxiety, or social desirability concerns, leading individuals – particularly those without a clinical history of firesetting – to suppress behaviours that might otherwise indicate curiosity or attraction. This raises important questions about when and how to deploy VR in forensic contexts. Less immersive methods, such as televised fire scenarios or static imagery, may sometimes elicit more diagnostically relevant responses by reducing perceived risk and encouraging more spontaneous reactions.

Study 2

As noted in Study 1, it is unclear why participants exhibiting higher levels of fire interest in the VRFA condition did not exhibit the behaviours we expected (e.g. higher levels of Overall Verbal Fire Appraisals or objects burnt in the fire) whilst participants in the TFA condition with higher levels of fire interest seemingly did. This discrepancy raised important questions about the sensitivity and diagnostic value of VR in detecting fire-related interests and

behaviours, particularly in non-clinical populations. To explore this further, Study 2 was designed as a pilot study focused on a clinically relevant sample – patients with a known history of deliberate firesetting – to assess whether VR can more effectively capture meaningful behavioural and psychological responses in individuals at higher risk.

Study 2 extended the methodology of Study 1 to a pilot sample of patients from secure psychiatric units, comparing a group with a history of firesetting to a group of control patients, with no known firesetting history. It aimed to investigate whether VR represents a superior methodology than traditional imagination-based tasks or the use of non-immersive pictorial stimuli, by examining the responses of adult patients with a history of firesetting (relative to those of non-firesetting control patients) to the three fire assessments: VRFA, TFA, or IFA. Participants' responses to these fire experiences were assessed – similarly to study 1 – via a variety of measures, including heart rate, verbalisations of thoughts and feelings, movement towards the fire, and interaction with the fire. The study also examined associations between these responses to the fire assessments and scores on a traditional questionnaire measure of fire interest (the FFFS).

Hypotheses

Due to the increased immersion and sense of presence associated with VR (Diemer et al., 2015; Freeman et al., 2017), it was hypothesised that participants in the VRFA condition would generate more extreme verbal appraisals and greater heart rate change than participants in the TFA and IFA conditions. Previous research suggests that individuals with a history of deliberate firesetting report greater fascination with, attraction to, and excitement about fire than individuals without a history of firesetting (Barrowcliffe & Gannon, 2016; Gannon et al., 2013). Consequently, it was hypothesised that the firesetting group would exhibit greater heart rate change (reflecting heightened emotional responding to fire-related stimuli), more positive verbal appraisals, greater movement towards the fire, and greater interaction with flammable items than the non-firesetting control group. In addition, given their greater interest in fire-related stimuli and lower levels of fire safety awareness relative to non-firesetting offenders (Gannon et al., 2013), it was expected that the firesetting group would be less likely to engage in behaviours aimed at suppressing the fire, i.e. picking up the fire extinguisher. Finally, because the Four Factor Fire Scales (FFFS) were designed to assess fire interest and fire-related attitudes (Ó Ciardha, Tyler, et al., 2015), it was hypothesised that scores on the traditional questionnaire measure of fire interest (the FFFS) would be positively correlated with heart rate change, verbalisation scores, watch time, and the number of flammable items picked up and burned. Negative correlations were predicted for closest and average distance from the fire, and interaction with the fire extinguisher.

Method

Ethics approval for this study was granted by the University's Research Ethics Committee (Ref: 201915701929305878) and South Central – Oxford C National Health Service Research Ethics Committee (Ref:19/SC/0537).

Design

As in Study 1, participants were allocated to one of three assessment conditions: the virtual reality fire assessment (VRFA), the televised two-dimensional fire assessment (TFA), or the imagined fire assessment (IFA). However, this study examined the responses of adult patients with a history of firesetting, relative to those of non-firesetting control patients. Thus, the study had a 3 (fire assessment condition: VRFA; TFA; IFA) x 2 (firesetting status: firesetting; control) design.

Participants

As in Study 1, patients were excluded from participating if they had received a diagnosis of Epilepsy or PTSD. Prior to recruitment, two individuals were identified as ineligible to participate: one due to a diagnosis of epilepsy and one because they used a wheelchair and would have been unable to navigate the virtual environment and simultaneously operate the hand controls required for the assessment. A total of 32 participants were recruited from across two forensic mental health settings (one low secure, one medium secure). Participants with a history of deliberate firesetting or fire-related risk behaviours (e.g. repeated threats to set fires) as noted in their hospital records were recruited to the 'firesetting' group ($n = 17$). Participants with no known history of firesetting were recruited to the 'control' group ($n = 15$). As in Study 1, patients were excluded from participating if they had received a diagnosis of Epilepsy or PTSD. Participants' ages ranged from 21 to 72 years ($M = 39.38$, $SD = 12.59$) and the majority identified themselves as White (78.10%, $n = 25$). The vast majority of participants were male, with one female participant in the Fire-setting-Imagined group.

Measures

Fire assessments. The three fire assessments utilised in Study 1 were employed in this study too. However, in this study each of the fire assessments lasted for two minutes rather than three to streamline the assessment for clinical use. The only other change in these fire assessments was that a heat source was placed in the approximate position of the fire during the VRFA. The heat source was activated simultaneously with the appearance of the fire within the virtual environment and remained active for the duration of the assessment. All other equipment, instructions, and measures captured remained the same.

Pre-experience measures. As in Study 1, participants completed a demographic questionnaire, the short two-subtest format of WASI-II (as a brief measure of general cognitive functioning) and the PDS-IM (to measure socially desirable responding). A questionnaire measure of fire interest was administered, however the measure differed from Study 1. Study 2 utilised the Four Factor Fire Scales (FFFS; Ó Ciardha, Tyler, et al., 2015) as this had been recommended for use with individuals who had been apprehended for their firesetting. Further details on this measure are below.

The FFFS combine items from three fire-related measures: the Fire Interest Rating Scale (Murphy & Clare, 1996), the Fire Attitude Scale (Muckley, 1997), and the Identification with Fire Questionnaire (Gannon et al., 2011). The Fire Interest Rating Scale examines an individual's fascination with or attraction to fire and consists of 14 items describing fire-

related situations (e.g. 'Watching a house burn down'). Participants are asked to rate how interested they would be in each of the situations on a scale of 1 (*upsetting/frightening*) to 7 (*exciting, fun, or lovely*). The Fire Attitude Scale consists of 19 items and examines an individual's attitudes towards fire. Participants respond to items such as 'Setting just a small fire can make you feel a lot better' on a scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). The Identification with Fire Questionnaire consists of 17 items and assesses the extent to which an individual relates to or identifies with fire (e.g. 'Fire is almost part of my personality'). Participants also respond to this measure on a 5-point scale (1 = *Strongly Disagree* to 5 = *Strongly Agree*).

The FFFS combine the above measures to form four subscales that have been empirically determined via factor analysis (see Ó Ciardha, Tyler, et al., 2015). These subscales examine (1) identification with fire (e.g. 'Fire is almost part of my personality'; 11 items), (2) serious fire interest (e.g. 'Striking a match to set fire to a building'; 7 items), (3) perceived fire safety awareness (e.g. 'I know a lot about how to prevent fires'; 6 items), and (4) firesetting as normal (e.g. 'Most people have set a few small fires just for fun'; 7 items). The total score on the Four Factor Fire Scales reflects an individual's overall fire interest, attitudes, and affiliation to fire, and perceived fire safety awareness (Ó Ciardha, Tyler, et al., 2015). Gannon et al. (2013) have reported questionable to good internal reliability for these subscales when administered with imprisoned males with a history of firesetting (identification with fire $\alpha = .88$, serious fire interest $\alpha = .86$, perceived fire safety awareness $\alpha = .68$, normalisation of firesetting $\alpha = .73$) and excellent internal reliability for the total score ($\alpha = .90$).

Post-experience measures. Immediately after the fire experience, participants completed the same two post-experience measures as in Study 1 (namely the Sense of Presence Scale and the Visual Analogue Scale). In this study all participants (regardless of condition) were invited to complete the follow-up interview approximately two weeks later.

Procedure

The procedure for Study 2 was largely the same as described in Study 1, with the key difference being the duration of the fire experience, which lasted for two minutes instead of three. In Study 2, participants followed the same steps as in Study 1: providing informed consent, completing the demographic questionnaire and subsequent measures in a randomised order, and participating in a three-minute familiarisation stage. Instructions for exploring and verbalising thoughts in the VRFA, TFA, and IFA conditions remained consistent. Following the fire experience, they completed the measures of presence and pleasantness, and were invited to complete the follow-up interview approximately two weeks later. One participant completed the study with the aid of a translator.

Data analysis

All analyses were conducted in R 4.1.0. To assess whether there were any significant differences on the fire experience measures, factorial ANOVAs were conducted, with the exception of picking up the fire extinguisher for which a Fisher's Exact Test was employed. Where appropriate, post hoc comparisons with Bonferroni adjustment were also performed to identify the location of between group differences. Due to the significant

differences across the fire assessment conditions in FFFS Poor Fire Safety scores, two-way ANCOVAs were also conducted for all measures to control for self-reported fire safety awareness. The relationship between measures captured during the fire experience and the questionnaire measure of inappropriate fire interest (the FFFS Serious Fire Interest subscale) was examined through correlations. The primary data for this research are not available due to lack of consent for data sharing from participants.

Due to the small sample size of this pilot study, marginally statistically significant associations ($p < .10$) are reported, in addition to statistically significant associations ($p < .05$). This practice has been used in many other studies (e.g. Daljeet et al., 2022; Maragh-Bass et al., 2017). For all ANOVAs, partial eta squared (η_p^2) is reported as an effect size measure, which, as advised by Richardson (2011), is interpreted according to Cohen's (1988) guidelines: .01 represents a small effect, .06 represents a medium effect, and .14 represents a large effect.

Results

Pre-experience measures

Descriptive statistics for participant demographics and their scores on pre-experience measures are displayed in Table 6.

Demographics. Participant groups did not significantly differ in terms of age or ethnicity. For age, there was no significant main effect of firesetting status, $F(1, 26) = 0.02$, $p = .887$, $\eta_p^2 = .001$, or fire assessment condition, $F(2, 26) = 2.32$, $p = .118$, $\eta_p^2 = .152$, and no significant interaction, $F(2, 26) = 1.46$, $p = .250$, $\eta_p^2 = .101$. Similarly, there was no significant association between ethnicity and status ($p = .378$, Fisher's Exact Test), or ethnicity and condition ($p = .660$, Fisher's Exact Test).

Experience with VR. When examining prior experience with VR, a Fisher's Exact Test indicated there was no significant association between previous VR experience and firesetting status ($p = .491$). There was also no significant association between VR experience and condition ($p = .196$, Fisher's Exact Test).

Impression management. There were eight responses missing to item 12 ('I have never dropped litter on the street'), three responses missing to item 10 ('I always declare everything at customs'), and one response missing to item 18 ('I have never damaged a library book or store merchandise without reporting it'). Consequently, Little's MCAR test was performed. It was not significant suggesting that these values were missing completely at random ($p = .252$). Because of this, predictive mean matching was used to impute values.

There was no significant main effect of status, $F(1, 26) = 1.79$, $p = .192$, $\eta_p^2 = .065$, or condition, $F(2, 26) = 0.36$, $p = .704$, $\eta_p^2 = .027$, on impression management scores. There was no significant status by condition interaction, $F(2, 26) = 0.10$, $p = .905$, $\eta_p^2 = .008$. Thus, impression management was not included in any further analyses as a covariate.

Cognitive ability (WASI-II). When the two subtests of the WASI-II were combined to form the full-scale intelligence quotient (FSIQ-2), there was no significant main effect of status,

Table 6. Descriptive statistics for Study 2 pre-experience measures.

Table 1. Descriptive statistics for Study 1: fire experience indicators												
	Firesetting Participants						Control Participants					
	VRFA (n = 6)	TFA (n = 6)	IFA (n = 5)	n or M (SD)	VRFA (n = 5)	TFA (n = 5)	IFA (n = 5)	n or M (SD)	VRFA (n = 5)	TFA (n = 5)	IFA (n = 5)	
	n or M (SD)	n or M (SD)	n or M (SD)		n or M (SD)	n or M (SD)	n or M (SD)		n or M (SD)			
Age	40.33	(9.20)	38.33	(18.90)	40.40	(7.44)	43.60	(12.74)	27.40	(9.13)	46.20	(10.52)
Ethnicity												
White	5		6		4		3		4		3	
Black	1		0		0		1		1		1	
Mixed race	0		0		1		1		0		1	
Gender												
Male	6		6		4		5		5		5	
Female	0		0		1		0		0		0	
R Experience												
No experience	4		2		3		5		3		2	
Any experience	2		4		2		0		2		3	
Impression Management	9.00	(10.18)	8.67	(6.15)	7.60	(5.37)	12.80	(10.06)	13.60	(2.61)	9.60	(7.40)
MAISI-II Full Scale IQ	85.00	(12.81)	81.17	(16.68)	87.80	(8.05)	87.80	(16.42)	84.80	(19.28)	85.20	(9.93)
Fire-related factors												
Serious fire interest	10.67	(5.01)	12.50	(4.42)	12.80	(3.11)	7.80	(3.03)	6.80	(2.49)	11.80	(3.49)
Identification with fire	20.00	(11.15)	18.33	(5.24)	26.60	(10.88)	18.80	(8.04)	12.80	(2.49)	22.40	(6.88)
Poor fire safety	10.50	(2.66)	14.50	(4.32)	12.40	(0.55)	8.00	(4.47)	11.60	(3.78)	13.20	(1.30)
Firesetting as normal	18.83	(4.79)	21.33	(3.93)	21.20	(1.79)	16.00	(6.75)	19.00	(2.92)	17.80	(3.90)
Total score	58.50	(13.63)	64.17	(5.31)	70.80	(9.36)	51.00	(16.57)	50.20	(7.40)	62.80	(4.97)

$F(1, 26) = 0.01$, $p = .937$, $\eta_p^2 = .001$, or condition, $F(2, 26) = 0.31$, $p = .735$, $\eta_p^2 = .023$. There was no significant interaction between status and condition, $F(2, 26) = 0.29$, $p = .748$, $\eta_p^2 = .022$.

Four Factor Fire Scales (FFFS). Scores on each scale of the Four Factor Fire Scales, as well as the total score, were examined to see if there were any differences across the participant groups. Participants with a history of firesetting had significantly higher Serious Fire Interest scores ($M = 11.94$, $SD = 4.16$) than control participants with no firesetting history ($M = 8.80$, $SD = 3.59$), $F(1, 26) = 5.64$, $p = .025$, $\eta_p^2 = .178$. There was a medium but non-significant main effect of condition, $F(2, 26) = 1.99$, $p = .157$, $\eta_p^2 = .133$, and no significant interaction between status and condition, $F(2, 26) = 1.03$, $p = .372$, $\eta_p^2 = .073$.

For the Identification with Fire scale there was no significant main effect of status, $F(1, 26) = 1.61$, $p = .216$, $\eta_p^2 = .058$, or condition, $F(2, 26) = 3.19$, $p = .058$, $\eta_p^2 = .197$, and no significant status*condition interaction, $F(2, 26) = 0.20$, $p = .817$, $\eta_p^2 = .015$.

On the Poor Fire Safety scale, there was no significant main effect of status, $F(1, 26) = 1.77$, $p = .195$, $\eta_p^2 = .064$. There was a significant main effect of condition, $F(2, 26) = 4.62$, $p = .019$, $\eta_p^2 = .262$. Post hoc Bonferroni-adjusted comparisons showed a significant difference only between the VRFA condition ($M = 9.36$, $SD = 3.64$) and the TFA condition ($M = 13.18$, $SD = 4.17$) on Poor Fire Safety scores ($p = .033$). The difference between VRFA and the IFA condition ($M = 12.80$, $SD = 1.03$) was marginally significant, $p = .072$. The IFA condition did not significantly differ from the TFA condition ($p = 1.00$). There was no significant status*condition interaction, $F(2, 26) = 1.01$, $p = .379$, $\eta_p^2 = .072$.

On the Firesetting as Normal scale, there was no significant main effect of status, $F(1, 26) = 3.50$, $p = .073$, $\eta_p^2 = .118$, no significant main effect of condition, $F(2, 26) = 1.21$, $p = .316$, $\eta_p^2 = .085$, and no significant interaction between status and condition, $F(2, 26) = 0.04$, $p = .961$, $\eta_p^2 = .003$.

Finally, for the FFFS total score there was a significant main effect of status, $F(1, 26) = 7.02$, $p = .014$, $\eta_p^2 = .213$. Participants with a history of firesetting had significantly higher FFFS total scores ($M = 64.12$, $SD = 10.71$) than control participants with no firesetting history ($M = 54.67$, $SD = 11.69$, $p = .023$). There was also a significant main effect of condition, $F(2, 26) = 3.84$, $p = .035$, $\eta_p^2 = .228$. Post hoc Bonferroni-adjusted comparisons showed a marginally significant difference only between the IFA condition ($M = 66.80$, $SD = 8.23$) and the VRFA condition ($M = 55.09$, $SD = 8.23$; $p = .073$). FFFS total scores for the TFA condition ($M = 57.82$, $SD = 9.44$) did not significantly differ from the other conditions. There was no significant status by condition interaction, $F(2, 26) = 0.32$, $p = .729$, $\eta_p^2 = .024$.

Fire experience measures

Means and standard deviations for each of the measures captured during the fire experiences can be found in Table 7. Factorial ANOVAs were conducted to examine differences across firesetting status and conditions for each of the measures. Due to the significant main effect of condition on FFFS Poor Fire Safety scores, ANCOVAs were also conducted for each measure, controlling for these scores as a covariate.

Heart rate. There were small, non-significant main effects of firesetting status, $F(1, 26) = 0.03$, $p = .869$, $\eta_p^2 = .001$, and condition, $F(2, 26) = 0.09$, $p = .913$, $\eta_p^2 = .007$, on change in

Table 7. Means and standard deviations for measures captured during fire assessments.

	Firesetting Participants						Control Participants					
	VRFA (n = 6)		TFA (n = 6)		IFA (n = 5)		VRFA (n = 5)		TFA (n = 5)		IFA (n = 5)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Heart rate change	−0.02	(12.30)	4.97	(8.46)	0.00	(2.62)	4.70	(6.51)	−0.38	(8.85)	2.04	(2.65)
Overall verbal fire appraisals	−3.83	(0.98)	−2.17	(1.47)	−1.80	(1.79)	−6.20	(3.27)	−4.00	(2.45)	−1.60	(1.34)
Closest distance	1.61	(0.55)	1.32	(1.03)	−	−	1.48	(0.48)	0.57	(0.56)	−	−
Average distance	2.25	(0.27)	2.30	(0.18)	−	−	2.14	(0.45)	2.08	(0.29)	−	−
Watch time	49.12	(36.69)	19.48	(12.99)	−	−	38.22	(11.53)	39.04	(21.21)	−	−
Number of flammable items picked up	1.00	(1.26)	0.83	(1.60)	−	−	0.60	(1.34)	1.20	(1.30)	−	−
Number of flammable items burned	0.00	(0.00)	0.67	(1.21)	−	−	0.00	(0.00)	0.20	(0.45)	−	−
	% yes	(n)	% yes	(n)			% yes	(n)	% yes	(n)		
Fire extinguisher picked up	66.67	(4)	83.33	(5)	−	−	60.00	(3)	100.00	(5)	−	−

heart rate. There was no significant interaction between status and condition on this measure, $F(2, 26) = 1.16$, $p = .329$, $\eta_p^2 = .082$. A two-way ANCOVA was also performed to examine the effects of status and condition on heart rate, after controlling for FFFS Poor Fire Safety scores. There were non-significant main effects of firesetting status, $F(1, 25) = 0.002$, $p = .965$, $\eta_p^2 < .001$, and condition, $F(2, 25) = 0.11$, $p = .900$, $\eta_p^2 = .008$, on change in heart rate. There was also no significant interaction between status and condition on this measure, whilst controlling for Fire Safety, $F(2, 25) = 1.21$, $p = .316$, $\eta_p^2 = .088$.

Verbal fire appraisals. When examining overall verbal fire appraisals, there was a marginally significant main effect of status, $F(1, 26) = 3.58$, $p = .070$, $\eta_p^2 = .121$. Control participants voiced more negative (relative to positive) fire appraisals ($M = -3.96$, $SD = 3.01$) than participants with a history of firesetting ($M = -2.65$, $SD = 1.62$). There was a significant main effect of condition, $F(1, 26) = 47.39$, $p = .003$, $\eta_p^2 = .362$. Post hoc Bonferroni-adjusted comparisons showed that participants in the VRFA condition ($M = -4.91$, $SD = 2.51$) voiced significantly more negative (relative to positive) fire appraisals than participants in the IFA condition ($M = -1.70$, $SD = 1.49$; $p = .005$). Total verbalisations in the TFA condition ($M = -3.00$, $SD = 2.10$) did not significantly differ from the other conditions. There was not a significant status*condition interaction, $F(2, 26) = 1.20$, $p = .317$, $\eta_p^2 = .084$.

A two-way ANCOVA was performed to examine the effects of status and condition on verbal fire appraisals, after controlling for FFFS Poor Fire Safety scores. There was no significant main effect of status, $F(1, 25) = 2.35$, $p = .137$, $\eta_p^2 = .086$, but after controlling for FFFS Poor Fire Safety scores, the main effect of condition remained significant, $F(2, 25) = 4.05$, $p = .030$, $\eta_p^2 = .245$. There was not a significant status*condition interaction, $F(2, 25) = 0.76$, $p = .479$, $\eta_p^2 = .057$.

Interaction with the fire. Since the Imagined condition did not involve recording movement, interaction with virtual items, or watch time, all further analyses of fire experience measures feature only the TFA and VRFA conditions.

Movement

For average distance from the fire, there was no significant main effect of status, $F(1, 18) = 1.52$, $p = .233$, $\eta_p^2 = .078$, or condition, $F(1, 18) = 0.01$, $p = .941$, $\eta_p^2 < .001$. There was no significant interaction between status and condition, $F(1, 18) = 0.15$, $p = .700$, $\eta_p^2 = .008$. After controlling for FFFS Poor Fire Safety scores in a two-way ANCOVA, there remained no significant main effects of status, $F(1, 17) = 1.97$, $p = .178$, $\eta_p^2 = .104$, or condition, $F(1, 17) = 0.11$, $p = .750$, $\eta_p^2 = .006$. There was no significant interaction, $F(1, 17) = 0.17$, $p = .688$, $\eta_p^2 = .010$.

In terms of the closest distance participants got to the fire, there was a medium but non-significant main effect of status, $F(1, 18) = 2.08$, $p = .166$, $\eta_p^2 = .104$. The control group ($M = 1.03$, $SD = 0.69$) moved closer to the fire than the firesetting group ($M = 1.46$, $SD = 0.80$). There was also a large, marginally significant main effect of condition, $F(1, 18) = 3.90$, $p = .064$, $\eta_p^2 = .178$. Participants in the TFA condition ($M = 0.98$, $SD = 0.90$) got closer to the fire than participants in the VRFA condition ($M = 1.55$, $SD = 0.50$). There was also no significant status*condition interaction, $F(1, 18) = 1.04$, $p = .321$, $\eta_p^2 = .055$. After controlling for FFFS Poor Fire Safety scores in a two-way ANCOVA, there was a large, marginally significant main effect of status, $F(1, 17) = 3.68$, $p = .072$, $\eta_p^2 = .178$.

There was a small, non-significant main effect of condition, $F(1, 17) = 1.02$, $p = .328$, $\eta_p^2 = .056$. There was no significant interaction, $F(1, 17) = 1.20$, $p = .288$, $\eta_p^2 = .066$.

Interaction with items

Examining the number of flammable items picked up, there was no significant main effect of status, $F(1, 18) < 0.01$, $p = .978$, $\eta_p^2 < .001$, or condition, $F(1, 18) = 0.13$, $p = .720$, $\eta_p^2 = .007$. There was also no significant status*condition interaction, $F(1, 18) = 0.41$, $p = .528$, $\eta_p^2 = .022$. A two-way ANCOVA controlling for FFFS Poor Fire Safety scores demonstrated no significant main effects of status, $F(1, 17) = 0.07$, $p = .801$, $\eta_p^2 = .004$, or condition, $F(1, 17) = 0.003$, $p = .955$, $\eta_p^2 < .001$. There was not a significant status*condition interaction, $F(1, 17) = 0.43$, $p = .519$, $\eta_p^2 = .025$.

In addition, the number of flammable items burned was examined. There was a small, non-significant main effect of status, $F(1, 18) = 0.66$, $p = .428$, $\eta_p^2 = .035$, and a medium, non-significant main effect of condition, $F(1, 18) = 2.27$, $p = .150$, $\eta_p^2 = .112$. There was no significant interaction between status and condition, $F(1, 18) = 0.66$, $p = .428$, $\eta_p^2 = .035$. A two-way ANCOVA controlling for FFFS Poor Fire Safety scores demonstrated no significant main effects of status, $F(1, 17) = 0.12$, $p = .721$, $\eta_p^2 = .007$, or condition, $F(1, 17) = 0.622$, $p = .441$, $\eta_p^2 < .035$. There was not a significant status*condition interaction, $F(1, 17) = 0.62$, $p = .444$, $\eta_p^2 = .035$.

The final interaction-based measure examined whether participants picked up the fire extinguisher. A Fisher's Exact Test indicated there was no significant association between picking up the fire extinguisher and firesetting status ($p = 1.00$). There was also no significant association between picking up the fire extinguisher and condition ($p = .311$).

Watch time

There was no significant main effect of status, $F(1, 18) = 0.19$, $p = .672$, $\eta_p^2 = .010$, nor condition, $F(1, 18) = 2.06$, $p = .169$, $\eta_p^2 = .103$ on time spent watching the fire. There was no significant interaction between status and condition, $F(1, 18) = 2.30$, $p = .147$, $\eta_p^2 = .113$. After controlling for FFFS Poor Fire Safety scores in a two-way ANCOVA, there remained no significant main effects of status, $F(1, 17) < 0.01$, $p = .957$, $\eta_p^2 < .001$, or condition, $F(1, 17) = 0.62$, $p = .442$, $\eta_p^2 = .035$. There was no significant interaction, $F(1, 17) = 2.26$, $p = .151$, $\eta_p^2 = .117$.

Relationship between measures and FFFS fire interest scores

To determine how the measures which were captured during the fire experiences related to a traditional questionnaire measure of inappropriate fire interest, correlations between all measures and FFFS Serious Fire Interest scores were calculated and are displayed in Table 8. There was a marginally significant positive correlation between FFFS fire interest scores and the overall verbal appraisals. There was a significant positive correlation between the number of flammable items picked up and FFFS fire interest scores, and a marginally significant positive correlation with the number of flammable items burned. All other correlations were non-significant.

Post experience measures

Presence. For total presence scores, there was no significant main effect of status, $F(1, 26) = 0.22$, $p = .640$, $\eta_p^2 = .009$. There was a large, significant main effect of condition, $F(2, 26) =$

Table 8. Means, standard deviations, and correlations with FFFS serious fire interest scores.

	<i>n</i>	<i>M</i> (<i>SD</i>) or % yes	<i>r</i>	95% CI	<i>p</i> value
Heart rate change	32	1.92 (7.65)	.10	[−.20, 1.00]	.288
Overall verbal fire appraisal	32	−3.25 (2.42)	.29	[−.00, 1.00]	.051*
Closest distance to fire	22	1.26 (0.77)	−.19	[−1.00, .18]	.196
Average distance from fire	22	2.20 (0.29)	.05	[−1.00, .40]	.578
Time fire watched	22	36.27 (24.54)	−.14	[−.48, 1.00]	.732
Number of flammable items picked up	22	0.91 (1.31)	.37	[.01, 1.00]	.044**
Number of flammable items burned	22	0.23 (0.69)	.35	[−.01, 1.00]	.056*
Fire extinguisher picked up	22	0.77 (0.42)	.16	[−1.00, .49]	.759

* = marginally significant, $p < .10$; ** = significant, $p < .05$. All tests are one-sided.

3.76, $p = .037$, $\eta_p^2 = .224$. Presence scores in the IFA ($M = 18.00$, $SD = 8.42$) did not significantly differ from TFA or VRFA. There was a marginally significant status*condition interaction, $F(2, 26) = 2.75$, $p = .082$, $\eta_p^2 = .175$.

Following the marginally significant interaction effect, simple main effects analyses were performed to explore the effects of condition within each level of status. For firesetting participants, there was a significant effect of condition on presence, $F(2, 26) = 4.63$, $p = .019$, $\eta_p^2 = 0.262$. However, the main effect of condition on presence was not significant for the control group, $F(2, 26) = 1.97$, $p = .159$, $\eta_p^2 = 0.132$. Post hoc tests using the Bonferroni correction indicated there was a marginally significant difference between the firesetting VRFA condition ($M = 26.67$, $SD = 4.27$) and the control TFA condition ($M = 14.40$, $SD = 7.16$), $t(26) = 3.06$, $p = .077$. There was also a marginally significant difference between the firesetting VRFA condition ($M = 26.67$, $SD = 4.27$) and the firesetting IFA condition ($M = 14.60$, $SD = 7.64$), $t(26) = 3.01$, $p = .087$.

Pleasantness. Descriptive statistics for all post-experience measures can be seen in Table 9. There was no significant main effect of status, $F(1, 26) = 0.67$, $p = .422$, $\eta_p^2 = .025$, on pleasantness scores. There was a large, but non-significant main effect of condition, $F(2, 26) = 2.04$, $p = .151$, $\eta_p^2 = .136$. The IFA generated the highest pleasantness scores ($M = 4.09$, $SD = 2.68$), followed by the VR fire experience ($M = 3.85$, $SD = 2.38$). The TFA generated the lowest pleasantness scores ($M = 2.21$, $SD = 2.09$). There was no significant status*condition interaction, $F(2, 26) = 0.54$, $p = .587$, $\eta_p^2 = .040$.

Follow-up interviews. A total of 23 participants completed a follow-up interview. Six participants were unable to be contacted due to local or national COVID-19 restrictions, and three declined to participate. As in Study 1, no participants reported negative responses to any of the fire experiences. Across all participant groups, the majority of the interviews were coded as neutral ($n = 16$), with seven positive responses. When examining for potential differences across participant groups, there was no significant association between rating at follow-up and status ($p = .370$, Fisher's Exact Test), or condition ($p = 1.00$, Fisher's Exact Test).

Sensitivity power analyses

To evaluate the adequacy of Study 2 in terms of statistical power, we compared observed effect sizes (partial η_p^2) with the minimum detectable effect sizes (MDEs) at 80% power ($\alpha = .05$). For the overall 3×2 design ($N = 32$), the study was powered to detect large

Table 9. Descriptive statistics for post-experience measures.

	Firesetting Participants			Control Participants		
	VRFA (n = 6)	TFA (n = 6)	IFA (n = 5)	VRFA (n = 5)	TFA (n = 5)	IFA (n = 5)
Pleasantness	3.60 (2.25)	2.80 (2.01)	4.78 (3.29)	4.16 (2.75)	1.50 (2.17)	3.40 (2.04)
Presence	26.67 (4.27)	19.67 (6.65)	14.60 (7.64)	21.80 (5.02)	14.40 (7.16)	21.40 (8.50)
Follow-Up Interview						
Positive	1	2	2	1	1	0
Neutral	3	2	2	2	3	4
Negative	0	0	0	0	0	0

effects of approximately $\eta^2 \geq .25$ ($df_1 = 2$) and $\eta^2 \geq .21$ ($df_1 = 1$). For the 2×2 subset analyses ($N \approx 22$), the study was powered to detect effects of approximately $\eta^2 \geq .29$. The main effect of Condition on verbal fire appraisals ($\eta^2 = .36$) exceeded the detectable threshold and was therefore adequately powered. By contrast, the main effect of Status ($\eta^2 = .12$), the Condition $\times$ Status interaction ($\eta^2 = .04$), and all subset analyses (e.g. closest distance, watch time, and items interacted with; η^2 values = .03–.18) fell below the relevant MDE thresholds, indicating that these effects were underpowered in the present sample. Thus, Study 2 was capable of detecting large effects but was insufficiently powered to reliably identify small-to-medium effects.

Discussion

Study 2 represents an initial pilot investigation of the feasibility of using VR to assess fire interest in adults with a history of deliberate firesetting.

Verbal appraisals

As in study 1, VR-based stimuli were expected to elicit more emotions and offer greater insight into the underlying cognitions. In keeping with this, the VRFA condition generated the most extreme verbal appraisals, with a greater number of negative verbal appraisals relative to positive appraisals. For example, participants said 'I'd feel terrible. I'd wanna get out' and 'I'm feeling a bit apprehensive obviously.' After controlling for self-reported fire safety awareness, the difference between the VRFA and the IFA remained significant. However, there was a lack of significant difference between the VRFA and TFA in terms of the overall verbal appraisal, suggesting that presenting clients with pictorial stimuli may be sufficient to tap into the emotions and cognitions associated with fire, and an immersive experience is not necessary.

Although there was not a significant difference between the firesetting and the control group in terms of the overall verbal appraisal, participants with a history of firesetting did voice more positive appraisals. It is thought these positive verbalisations are an indicator of inappropriate interest in fire, whereby individuals are fascinated by fire and associate it with positive feelings and thoughts. In support of this, positive verbalisations were significantly positively correlated with fire interest as measured by the traditional questionnaire measure, the FFFS (Ó Ciardha, Barnoux, et al., 2015). This was expected since higher scores on this measure indicate participants providing positive affective responses (i.e. 'exciting, fun, or lovely') to items describing serious fire-related situations (e.g. 'Watching a house burn down'), similar to that depicted in the virtual environment.

However, caution is warranted when interpreting verbal responses within a forensic population. Some participants may have been motivated to provide socially desirable responses when discussing fire-related scenarios, particularly given the potentially sensitive nature of firesetting behaviour and the potential consequences associated with disclosure. Socially desirable responding and impression management have long been recognised as challenges in forensic assessment contexts, where individuals may have incentives to present themselves in a favourable manner (Hildebrand et al., 2018). However, there were no significant differences between participant groups on the measure of impression management, suggesting that any such tendency was not unique to the firesetting group. Nevertheless, it remains possible that verbal appraisals

were influenced by broader self-presentation concerns or participants' awareness of being observed. Future research should continue to examine the extent to which verbal responses during fire assessments are affected by these factors.

Despite this potential limitation, the verbal appraisal data also highlighted a potentially valuable clinical application of VR. Some firesetting participants in the VRFA condition reflected on their own experiences of having set deliberate fires. For example, one participant said, 'It makes me realise what I did probably wasn't the best of things to have done.' Such reflections suggest that immersive fire-related scenarios may encourage self-reflection regarding previous firesetting behaviour. This highlighted a potential further avenue for the use of VR with adults who have set fires; to improve their fire safety awareness which is a key treatment need for this adults who have set deliberate fires (Gannon et al., 2013). This avenue holds promise since VR has successfully been used to teach children fire safety skills (Çakiroğlu & Gökoğlu, 2019; Smith & Ericson, 2009).

Movement

In contrast, measures of participants' movement towards the fire did not significantly correlate with FFFS fire interest scores. This was unexpected; it was hypothesised that those with higher FFFS scores would move closer to the fire since they would be more fascinated by it. These findings suggest that movement towards virtual fires may not be a valid indicator of fire interest for future VR-based assessments. This aligns with recent commentary by Corneille and Gawronski (2024), who argue that despite their intuitive appeal, implicit measures often lack the reliability and validity of well-constructed self-report instruments, especially in complex, emotionally charged tasks like those simulated in VR. However, due to the pilot nature of the current research, it utilised a sample that would only have reliably identified large effects on these movement measures. The current study also found that the TFA generated greater movement towards the fire than the VRFA. As was discussed in Study 1, this may reflect the lower presence in the TFA condition; perhaps participants felt more comfortable moving towards a fire that felt less real and therefore represented less of a danger.

Interaction with items

Measures of whether participants interacted with flammable items appeared to be a better indicator of inappropriate fire interest as they were positively correlated with FFFS scores. However, there was low engagement with the flammable items across participant groups and conditions. Less than half of all participants picked up a flammable item (40.91%; $n = 9$). Fewer still burned these flammable items in the fire (9.09%; $n = 2$). Interestingly, both of the participants that burned items were in the TFA condition, while in the VRFA condition no participants burned an item. It is worth noting that Poor Fire Safety scores from the FFFS were significantly higher in the TFA condition and so this finding may actually reflect lower fire safety awareness among participants in the TFA condition. Indeed, after controlling for FFFS Poor Fire Safety scores, there was no significant main effect of condition on the number of flammable items burned. Further research is needed to investigate the extent to which interaction with flammable items within virtual environments could provide a useful gauge of various fire-related treatment needs, and whether an immersive or non-immersive environment best facilitates such interaction.

Surprisingly, there was a small, non-significant positive correlation between the fire extinguisher being picked up and FFFS interest scores. Adults with higher fire interest scores on the traditional questionnaire measure were more likely to pick up the virtual fire extinguisher. This finding was unexpected, as it had been hypothesised that individuals with greater fire interest would be less likely to engage in behaviours aimed at extinguishing the fire. One possible explanation is that some participants may have been interested in the fire extinguisher itself. Individuals with elevated fire interest may demonstrate broader interest in fire-related objects and paraphernalia, rather than solely fire itself (Gannon et al., 2022). Consequently, interaction with the fire extinguisher may not necessarily reflect a desire to extinguish the fire and could instead represent curiosity about, or engagement with, a fire-related object. Future research should therefore examine whether interactions with different virtual objects reflect distinct aspects of fire-related cognition, interest, and fire safety awareness.

However, it should also be noted that many participants realised during the familiarisation phase that there was no way they could operate the virtual fire extinguisher. Therefore, the number of participants that picked up the fire extinguisher during the fire experience may not be an accurate reflection of how many participants wanted to act to put out the fire. Future applications should therefore incorporate a virtual fire extinguisher that is able to be operated in order to better assess fire safety awareness behaviours. Research examining the use of VR for the training of fire and rescue service personnel has advocated for the use of realistic fire extinguishers rather than the usual VR controllers in such assessments (Saghafian et al., 2020; Seo et al., 2019). The use of more realistic or haptic fire extinguishers may further improve the ecological validity of such assessments and represents an important avenue for future research. More realistic interactions may also provide a richer source of behavioural data, including response latency, accuracy of extinguisher use, and persistence in attempting to suppress the fire. Consequently, future research should examine whether the inclusion of functional fire-safety equipment improves the assessment of fire-related treatment needs within virtual environments.

Watch time

Time spent watching the fire did not significantly differ across the fire experiences or across participant groups. It was also not significantly correlated with FFFS fire interest scores. This was surprising since it was hypothesised that adults with an inappropriate interest in fire would be transfixed by a virtual fire. However, this finding does align with the findings of Study 1, and previous research examining juveniles with a history of firesetting that has also failed to demonstrate an association between attending to pictorial fire-related stimuli and a questionnaire measure of inappropriate fire interest (Gallagher-Duffy et al., 2009). The equipment used in the current study did not allow for eye tracking, so instead watch time was calculated using the time the fire was in the centre of the field of view. Future research could incorporate more advanced eye tracking technology (as has been employed in research examining sexual interest among individuals who have sexually offended; e.g. Chartier et al., 2006; Renaud et al., 2009) to further examine its utility in the assessment of inappropriate fire interest.

It is also important to interpret these findings in light of the pilot nature of the study. Although several movement and behavioural measures did not significantly differ across

participant groups, the relatively small sample size limited statistical power and may have reduced our ability to detect small-to-medium effects. Consequently, some potentially meaningful differences in behaviours such as movement towards the fire, watch time, and interaction with virtual objects may not have reached statistical significance. Larger-scale studies are therefore required to establish the reliability and sensitivity of these behavioural indicators.

Post-experience

Given that clinicians involved in the assessment and treatment of individuals who have set fires have previously voiced concern about retraumatising and eliciting negative emotions (Sambrooks et al., 2022), this study was particularly interested in the psychological outcomes of participating in these fire experiences. Crucially, no participants reported any negative effects from partaking in any of the fire experiences during a follow-up interview approximately two weeks later. This is a promising initial indicator that VR can be safely used to expose patients to fire-related stimuli.

The VRFA also appeared to be superior to the other fire experiences in terms of its realism. As hypothesised, the VRFA elicited significantly greater feelings of presence than the non-immersive TFA. This seemed to be more apparent for patients with a history of firesetting. Several firesetting participants reported during their verbalisations that they found the virtual fire on the bed to be very similar to fires they had previously set, particularly if they were cell fires (e.g. 'It does remind me of when I set a fire in my cell in prison'). This aligns with previous work that has demonstrated a positive correlation between participants' feelings of presence within a virtual environment and their ability to generate vivid mental images (Iachini et al., 2019). It is likely that participants in the firesetting group were able to generate more vivid fire-related imagery due to their own experiences with fire. Therefore, future research should endeavour to collect information on the context in which participants' past firesetting occurred and examine the impact of this on feelings of presence.

Conclusion

Although a pilot study with a small sample, Study 2 provides initial evidence that VR can be safely and feasibly used in the assessment of adults who have a history of firesetting. No ill effects from partaking in any of the fire experiences were reported by participants. This finding should alleviate some of the concerns around re-traumatisation that have previously been raised by clinicians (Sambrooks et al., 2022). Furthermore, the results indicate that VR can offer clinicians greater insight into the cognitions and emotions associated with firesetting, since the VR-based fire experience generated the largest number of verbalisations from participants. Further research is needed to establish whether an immersive experience is necessary to elicit these insights, or if two-dimensional pictorial stimuli would be sufficient. Future studies should also investigate the construct validity of movement and watch time measures in VR-based assessments of inappropriate fire interest. However, consistent with Corneille and Gawronski's (2024) assertion that implicit measures often lack robustness, researchers should remain cautious about overinterpreting behaviour-based indicators in the absence of corroborating self-report data.

General discussion

This research represents the first empirical investigation of the feasibility and validity of Virtual Reality (VR) as a novel assessment tool for measuring inappropriate fire interest in both community and forensic populations. Across two innovative studies, we compared an immersive VR fire assessment (VRFA) with two traditional methods: a televised fire assessment (TFA) and an imagined fire assessment (IFA). Participants' cognitive, emotional, and behavioural responses were assessed across a range of modalities, including verbal appraisals, heart rate, movement, item interaction, and post-experience measures of presence and pleasantness. Together, these studies offer critical insights into the utility and limitations of VR as a forensic assessment tool for deliberate firesetting.

One of the most consistent findings across both studies was that the VRFA elicited significantly greater feelings of presence. This supports a growing consensus that immersive VR technologies produce heightened emotional and sensory realism (Bell et al., 2020; Diemer et al., 2015). Participants described the VR experience as realistic and evocative, with some individuals in the firesetting group noting that the virtual fire reminded them of real fires they had set. This feedback underscores VR's potential to simulate offence-relevant contexts that traditional assessments cannot easily replicate. Notably, several participants from the firesetting group spontaneously engaged in reflective discussions about their offence-related behaviours during or following the VRFA. This unprompted introspection suggests that, beyond its utility as an assessment tool, VR may also hold promise as a therapeutic or rehabilitative medium – facilitating self-awareness and reflection potentially important for treatment engagement and behavioural change.

However, heightened realism did not universally translate to increased behavioural or verbal expressions of fire interest. Individuals with elevated fire interest (Study 1) or with a history of firesetting (Study 2) showed more overt engagement with fire stimuli (e.g. positive appraisals, burning flammable items) in the TFA condition rather than the VRFA. This suggests that the immersive realism of VR may activate inhibition, anxiety, or moral self-regulation, limiting the expression of inappropriate fire interest. These findings raise important considerations for the clinical and forensic use of VR in assessing interests or motivations. While VR can provide ecologically valid simulations, its immersive nature may necessitate alternative behavioural indicators or more subtle response metrics (e.g. gaze fixation, micro-movements, implicit association tasks). In contrast, non-immersive methods such as the TFA condition appeared to provide a safer cognitive distance, allowing participants to reveal more of their internal responses. The TFA environment elicited more behavioural engagement from high fire-interest individuals. This suggests that non-immersive fire stimuli may actually be more diagnostic of inappropriate interest than immersive stimuli, particularly in the context of assessment.

An additional contribution of this research lies in its demonstration that VR fire assessments can be administered safely, even with forensic psychiatric populations. Across both studies, no adverse effects were reported during the two-week follow-up interviews, providing reassuring evidence for the clinical safety of VR-based assessments in this context. This directly addresses concerns previously raised by clinicians (Sambrooks et al., 2022)

regarding the potential for VR fire simulations to retraumatise individuals with fire-related experiences or mental health vulnerabilities. Nevertheless, the ethical implications of using VR within forensic assessment contexts warrant careful consideration. Although no adverse effects were reported in either study, VR assessments expose individuals to offence-relevant stimuli and may therefore evoke strong emotional, cognitive, or behavioural responses. Consequently, careful consideration should be given to issues of informed consent, psychological wellbeing, appropriate monitoring during assessments, and the potential implications of assessment findings for clinical decision-making and risk management. As VR technologies become increasingly sophisticated and accessible, the development of clear ethical guidance for their use within forensic assessment and intervention will be essential.

As with any scientific study, limitations must be noted. Sample sizes were modest for our specialist sample (Study 2) limiting statistical power. Nevertheless, we did exhaust our local hospital NHS trust to ensure that the VR assessment was both safe and feasible for use in the assessment of adult patients with a history of firesetting. The present research was conducted exclusively within the United Kingdom. Consequently, caution is warranted when generalising these findings to other countries and cultural contexts. International variation in the prevalence, legal classification, and management of deliberate firesetting has previously been highlighted (Gannon et al., 2022; Tyler et al., 2019). Attitudes towards fire, firesetting, risk, and authority may vary across cultures and legal systems (Gannon et al., 2022), potentially influencing how individuals respond to fire-related stimuli and engage with virtual environments. In addition, differences in forensic mental health services, criminal justice practices, and approaches to firesetting assessment may affect the implementation and utility of VR-based assessments. Future research should therefore examine the applicability of these findings across a range of cultural, clinical, and forensic settings.

In addition, there are concerns about the utility of some of the fire experience measures utilised. For example, the VR fire extinguisher was non-operational, which may have reduced ecological validity for assessing fire safety behaviours. Additionally, movement data were limited to gross spatial indicators and did not include more precise attentional metrics such as eye-tracking or response latency. Heart rate, used as an indicator of physiological arousal, did not significantly differ across conditions. While we anticipated that individuals experiencing the immersive VR fire scenario might exhibit elevated heart rate responses, the absence of such effects is perhaps unsurprising given the known variability and limited specificity of heart rate in detecting subtle emotional responses (Kreibig, 2010). Moreover, heart rate lacks the temporal precision of measures like skin conductance, which are more responsive to phasic changes in autonomic arousal (Boucsein, 2012; Critchley, 2002). In the context of forensic VR assessments, where stimuli may evoke a mix of affective, cognitive, and moral responses, relying on heart rate and movement data alone may fail to capture other nuanced physiological markers of engagement or internal conflict.

The limitations of heart rate measurement may be particularly relevant within immersive VR environments. Indeed, a recent systematic review of VR-based psychiatric assessments found that physiological measures often demonstrated less discriminatory value than self-report measures, with patients and controls frequently exhibiting similar

physiological responses despite differences in subjective experience (Geraets et al., 2022). The authors suggested that physiological measures obtained during immersive VR experiences may be particularly susceptible to noise, as participants are simultaneously moving, navigating complex environments, and responding to multiple stimuli, in contrast to the controlled conditions under which such measures are typically validated. Consequently, the absence of significant heart rate differences in the present studies should not necessarily be interpreted as evidence that participants responded similarly to the fire-related stimuli. Rather, it may reflect the limitations of relying on a single physiological indicator to capture complex emotional and cognitive responses within immersive environments.

Future research should therefore adopt a multi-modal assessment approach that combines physiological, behavioural, and self-report measures. Rather than relying solely on heart rate, incorporating additional indicators such as electrodermal activity, eye-tracking, and pupillometry may provide complementary information regarding emotional arousal, attentional engagement, and responses to fire-related stimuli. This may enhance measurement precision and better elucidate the interplay between emotional salience and observable behaviour in virtual environments. This recommendation is consistent with previous VR research demonstrating the value of such measures for identifying attentional preferences and patterns of engagement that may not be evident through self-report or behavioural observation alone (Chartier et al., 2006; Iachini et al., 2019; Nee et al., 2019). Integrating these measures with behavioural indicators, such as movement and interaction with virtual objects, may provide a more comprehensive understanding of how individuals attend and respond to fire-related stimuli. Such an approach may also help distinguish between emotional arousal, attentional engagement, and consciously reported fire interest, thereby providing a richer assessment of fire-related treatment needs and risk-relevant processes.

In conclusion, the use of VR in the assessment of inappropriate fire interest offers both promise and complexity. While it enhances realism and presence, it may simultaneously suppress overt behavioural expression, especially in ethically charged contexts. Clinicians and researchers must therefore be cautious in interpreting the absence of overt indicators as the absence of inappropriate interest. Rather, a multi-method approach that combines fire-related stimuli with traditional self-report and behavioural tasks may offer the most comprehensive and valid assessment strategy.

Open Scholarship



This article has earned the [Center for Open Science](#) badge for Preregistered. The materials are openly accessible at .

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the EIRA (Enabling Innovation: Research to Application) Proof of Concept Fund POC001.

ORCID

Katie Sambrooks  <http://orcid.org/0000-0002-7981-1871>

Theresa A. Gannon  <http://orcid.org/0000-0002-5810-4158>

References

- Alleyne, E., Gannon, T. A., Mozova, K., Page, T. E., & Ó Ciardha, C. (2016). Female fire-setters: Gender-associated psychological and psychopathological features. *Psychiatry*, 79(4), 364–378. <https://doi.org/10.1080/00332747.2016.1185892>
- Barrowcliffe, E., & Gannon, T. A. (2016). Comparing the psychological characteristics of un-apprehended firesetters and non-firesetters living in the UK. *Psychology, Crime & Law*, 22(4), 382–404. <https://doi.org/10.1080/1068316X.2015.1111365>
- Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in Clinical Neuroscience*, 22(2), 169–177. <https://doi.org/10.31887/DCNS.2020.22.2/lvalmaggia>
- Benbouriche, M., Nolet, K., Trottier, D., & Renaud, P. (2014). Virtual reality applications in forensic psychiatry. *VRIC '14: Proceedings of the 2014 Virtual Reality International Conference*, 1–4. <https://doi.org/10.1145/2617841.2620692>
- Boucsein, W. (2012). *Electrodermal activity*. Springer US. <https://doi.org/10.1007/978-1-4614-1126-0>
- Çakiroğlu, Ü., & Gökoğlu, S. (2019). Development of fire safety behavioral skills via virtual reality. *Computers & Education*, 133, 56–68. <https://doi.org/10.1016/j.compedu.2019.01.014>
- Chartier, S., Renaud, P., Bouchard, S., Proulx, J., Rouleau, J. L., Fedoroff, P., & Bradford, J. (2006). Sexual preference classification from gaze behavior data using a multilayer perceptron. *Annual Review of CyberTherapy and Telemedicine*, 4, 149–157.
- Clus, D., Larsen, M. E., Lemey, C., & Berrouguet, S. (2018). The use of virtual reality in patients with eating disorders: Systematic review. *Journal of Medical Internet Research*, 20(4), e7898. <https://doi.org/10.2196/jmir.7898>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Corneille, O., & Gawronski, B. (2024). Self-reports are better measurement instruments than implicit measures. *Nature Reviews Psychology*, 3(12), 835–846. <https://doi.org/10.1038/s44159-024-00376-z>
- Cornet, L. J. M., & Van Gelder, J.-L. (2020). Virtual reality: A use case for criminal justice practice. *Psychology, Crime & Law*, 26(7), 1–17. <https://doi.org/10.1080/1068316X.2019.1708357>
- Critchley, H. D. (2002). Review: Electrodermal responses: What happens in the brain. *The Neuroscientist*, 8(2), 132–142. <https://doi.org/10.1177/107385840200800209>
- Daljeet, K. N., Feeney, J. R., & Goffin, R. D. (2022). Reactions to negative feedback: The role of resiliency and implications for counterproductivity. *Journal of Behavioral and Applied Management*, 22(2), 185–214.
- Diemer, J., Alpers, G. W., Peperkorn, H. M., Shibani, Y., & Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: A review of research in virtual reality. *Frontiers in Psychology*, 6, 26. <https://doi.org/10.3389/fpsyg.2015.00026>
- Fox, J., Arena, D., & Bailenson, J. N. (2009). Virtual reality: A survival guide for the social scientist. *Journal of Media Psychology*, 21(3), 95–113. <https://doi.org/10.1027/1864-1105.21.3.95>
- Freeman, D., Pugh, K., Antley, A., Slater, M., Bebbington, P., Gittins, M., Dunn, G., Kuipers, E., Fowler, D., & Garety, P. (2008). Virtual reality study of paranoid thinking in the general population. *British Journal of Psychiatry*, 192(4), 258–263. <https://doi.org/10.1192/bjp.bp.107.044677>

- Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, 47(14), 2393–2400. <https://doi.org/10.1017/S003329171700040X>
- Fromberger, P., Jordan, K., & Müller, J. L. (2018). Virtual reality applications for diagnosis, risk assessment and therapy of child abusers. *Behavioral Sciences & the Law*, 36(2), 235–244. <https://doi.org/10.1002/bsl.2332>
- Gallagher-Duffy, J., MacKay, S., Duffy, J., Sullivan-Thomas, M., & Peterson-Badali, M. (2009). The pictorial fire stroop: A measure of processing bias for fire-related stimuli. *Journal of Abnormal Child Psychology*, 37(8), 1165–1176. <https://doi.org/10.1007/s10802-009-9339-8>
- Gannon, T. A., & Barrowcliffe, E. (2012). Firesetting in the general population: The development and validation of the fire setting and fire proclivity scales. *Legal and Criminological Psychology*, 17(1), 105–122. <https://doi.org/10.1348/135532510X523203>
- Gannon, T. A., Ó Ciardha, C., Barnoux, M., Tyler, N., Mozova, K., & Alleyne, E. K. (2013). Male imprisoned firesetters have different characteristics than other imprisoned offenders and require specialist treatment. *Psychiatry: Interpersonal and Biological Processes*, 76(4), 349–364. <https://doi.org/10.1521/psyc.2013.76.4.349>
- Gannon, T. A., Ó Ciardha, C., & Barnoux, M.-F. (2011). *The identification with fire questionnaire*. Unpublished manuscript.
- Gannon, T. A., Ó Ciardha, C., Doley, R. M., & Alleyne, E. (2012). The Multi-Trajectory Theory of Adult Firesetting (M-TTAF). *Aggression and Violent Behavior*, 17(2), 107–121. <https://doi.org/10.1016/j.avb.2011.08.001>
- Gannon, T. A., & Pina, A. (2010). Firesetting: Psychopathology, theory and treatment. *Aggression and Violent Behavior*, 15(3), 224–238. <https://doi.org/10.1016/j.avb.2010.01.001>
- Gannon, T. A., Tyler, N., Ó Ciardha, C., & Alleyne, E. (2022). *Adult deliberate firesetting: Theory, assessment and treatment*. Wiley-Blackwell.
- George, D., & Mallery, P. (2018). Reliability analysis. In *IBM SPSS statistics 25 step by step* (pp. 249–260). Routledge.
- Geraets, C. N. W., van der Stouwe, E. C. D., Pot-Kolder, R., & Veling, W. (2021). Advances in immersive virtual reality interventions for mental disorders: A new reality? *Current Opinion in Psychology*, 41, 40–45. <https://doi.org/10.1016/j.copsyc.2021.02.004>
- Geraets, C. N. W., Wallinius, M., & Sygel, K. (2022). Use of virtual reality in psychiatric diagnostic assessments: A systematic review. *Frontiers in Psychiatry*, 13, 1–12. <https://doi.org/10.3389/fpsyt.2022.828410>
- Heeter, C. (1992). Being there: The subjective experience of presence. *Presence Teleoper. Virtual Environ*, 1, 262–271. <https://doi.org/10.1162/pres.1992.1.2.262>
- Hildebrand, M., Wibbelink, C. J. M., & Verschuere, B. (2018). Do impression management and self-deception distort self-report measures with content of dynamic risk factors in offender samples? A meta-analytic review. *International Journal of Law and Psychiatry*, 58, 157–170. <https://doi.org/10.1016/j.ijlp.2018.02.013>
- Home Office. (2025a). *FIRE0401: Deliberate fires attended by the Fire and Rescue Services*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1154900/fire-statistics-data-tables-fire0401-110523.xlsx.
- Home Office. (2025b). *FIRE0402: Fatalities and non-fatal casualties in deliberate fires attended*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1154903/fire-statistics-data-tables-fire0402-110523.xlsx.
- Iachini, T., Maffei, L., Masullo, M., Senese, V. P., Rapuano, M., Pascale, A., Sorrentino, F., & Ruggiero, G. (2019). The experience of virtual reality: Are individual differences in mental imagery associated with sense of presence? *Cognitive Processing*, 20(3), 291–298. <https://doi.org/10.1007/s10339-018-0897-y>
- Kardong-Edgren, S. S., Farra, S. L., Alinier, G., & Young, H. M. (2019). A call to unify definitions of virtual reality. *Clinical Simulation in Nursing*, 31, 28–34. <https://doi.org/10.1016/j.ecns.2019.02.006>
- Kip, H., Bouman, Y. H. A., Kelders, S. M., & van Gemert-Pijnen, L. J. E. W. C. (2018). Ehealth in treatment of offenders in forensic mental health: A review of the current state. *Frontiers in Psychiatry*, 9, 1–19. <https://doi.org/10.3389/fpsyt.2018.00042>

- Kip, H., Kelders, S. M., Weerink, K., Kuiper, A., Brüningshoff, I., Bouman, Y. H. A., Dijkslag, D., & van Gemert-Pijnen, L. J. E. W. C. (2019). Identifying the added value of virtual reality for treatment in forensic mental health: A scenario-based, qualitative approach. *Frontiers in Psychology*, 10, 1–15. <https://doi.org/10.3389/fpsyg.2019.00406>
- Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *The Biopsychology of Emotion: Current Theoretical and Empirical Perspectives*, 84(3), 394–421. <https://doi.org/10.1016/j.biopsycho.2010.03.010>
- Maragh-Bass, A. C., Torain, M., Adler, R., Ranjit, A., Schneider, E., Shields, R. Y., Kodadek, L. M., Snyder, C. F., German, D., Peterson, S., Schuur, J., Lau, B. D., & Haider, A. H. (2017). Is it okay to ask: Transgender patient perspectives on sexual orientation and gender identity collection in health-care. *Academic Emergency Medicine*, 24(6), 655–667. <https://doi.org/10.1111/acem.13182>
- Muckley, A. (1997). *Firesetting: Addressing offending behaviour, A resource and training manual*. Redcar and Cleveland Psychological Service.
- Murphy, G. H., & Clare, I. C. (1996). Analysis of motivation in people with mild learning disabilities (mental handicap) who set fires. *Psychology, Crime and Law*, 2(3), 153–164. <https://doi.org/10.1080/10683169608409774>
- Nanayakkara, V., Ogloff, J. R. P., Davis, M. R., & McEwan, T. E. (2020). Gender-based types of firesetting: Clinical, behavioural and motivational differences among female and male firesetters. *The Journal of Forensic Psychiatry & Psychology*, 31(2), 273–291. <https://doi.org/10.1080/14789949.2020.1720266>
- National Fire Chiefs Council. (2019). *Arson Reduction Strategy 2019-2022*. https://www.nationalfirechiefs.org.uk/write/MediaUploads/NFCC_Arson_Reduction_Strategy_Final_070619.pdf.
- Nee, C., Gelder, J.-L. v., Otte, M., Vernham, Z., & Meenaghan, A. (2019). Learning on the job: Studying expertise in residential burglars using virtual environments. *Criminology*, 57(3), 481–511. <https://doi.org/10.1111/1745-9125.12210>
- Ó Ciardha, C., Barnoux, M. F. L., Alleyne, E. K. A., Tyler, N., Mozova, K., & Gannon, T. A. (2015). Multiple factors in the assessment of firesetters' fire interest and attitudes. *Legal and Criminological Psychology*, 20(1), 37–47. <https://doi.org/10.1111/lcrp.12065>
- Ó Ciardha, C., Tyler, N., & Gannon, T. A. (2015). A practical guide to assessing adult firesetters' fire-specific treatment needs using the Four Factor Fire Scales. *Psychiatry*, 78(4), 293–304. <https://doi.org/10.1080/00332747.2015.1061310>
- Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires* (pp. 143–165). Springer.
- Paulhus, D. L. (1998). *Paulhus Deception Scales (PDS): User's manual*. Multi-Health Systems.
- Renaud, P., Chartier, S., Rouleau, J.-L., Proulx, J., Trottier, D., Bradford, J. P., Fedoroff, P., Bouchard, S., & Décarie, J. (2009). Gaze behavior nonlinear dynamics assessed in virtual immersion as a diagnostic index of sexual deviancy: Preliminary results. *JVRB - Journal of Virtual Reality and Broadcasting*, 6(3), 1–8. <https://doi.org/10.20385/1860-2037/6.2009.3>
- Renaud, P., Trottier, D., Rouleau, J.-L., Goyette, M., Saumur, C., Boukhalfi, T., & Bouchard, S. (2014). Using immersive virtual reality and anatomically correct computer-generated characters in the forensic assessment of deviant sexual preferences. *Virtual Reality*, 18(1), 37–47. <https://doi.org/10.1007/s10055-013-0235-8>
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135–147. <https://doi.org/10.1016/j.edurev.2010.12.001>
- Rus-Calafell, M., Garety, P., Sason, E., Craig, T., & Valmaggia, L. (2017). Virtual reality in the assessment and treatment of psychosis: A systematic review of its utility, acceptability and effectiveness. *Psychological Medicine*, 48, 1–30. <https://doi.org/10.1017/S0033291717001945>
- Saghafian, M., Laumann, K., Akhtar, R. S., & Skogstad, M. R. (2020). The evaluation of virtual reality fire extinguisher training. *Frontiers in Psychology*, 11, 1–16. <https://doi.org/10.3389/fpsyg.2020.593466>
- Sambrooks, K., Lockerbie, L., Majid, S., & Gannon, T. (2022). Clinicians' perceptions of virtual reality for firesetting. *The Journal of Forensic Practice*, 24(4), 404–419. <https://doi.org/10.1108/JFP-05-2022-0027>

- Schmidt, A. F., Banse, R., & Imhoff, R. (2015). Indirect measures in forensic contexts. In T. M. Ortner, & F. J. R. van de Vijver (Eds.), *Behavior-based assessment: Going beyond self-report in the personality, affective, motivation. And social domains* (pp. 173–194). Hogrefe.
- Segawa, T., Baudry, T., Bourla, A., Blanc, J.-V., Peretti, C.-S., Mouchabac, S., & Ferreri, F. (2020). Virtual Reality (VR) in assessment and treatment of addictive disorders: A systematic review. *Frontiers in Neuroscience*, 13, 1–14. <https://doi.org/10.3389/fnins.2019.01409>
- Seo, S.-W., Kwon, S., Hassan, W., Talhan, A., & Jeon, S. (2019). Interactive virtual-reality fire extinguisher with haptic feedback. *25th ACM Symposium on Virtual Reality Software and Technology*, 1–2. <https://doi.org/10.1145/3359996.3364725>
- Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, 3, 1–47. <https://doi.org/10.3389/frobt.2016.00074>
- Slater, M., Steed, A., McCarthy, J., & Maringelli, F. (1998). The influence of body movement on subjective presence in virtual environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(3), 469–477. <https://doi.org/10.1518/001872098779591368>
- Smith, S., & Ericson, E. (2009). Using immersive game-based virtual reality to teach fire-safety skills to children. *Virtual Reality*, 13(2), 87–99. <https://doi.org/10.1007/s10055-009-0113-6>
- Snowden, R. J., Craig, R. L., & Gray, N. S. (2011). Indirect behavioral measures of cognition among sexual offenders. *Journal of Sex Research*, 48(2–3), 192–217.
- Sygel, K., & Wallinius, M. (2021). Immersive virtual reality simulation in forensic psychiatry and adjacent clinical fields: A review of current assessment and treatment methods for practitioners. *Frontiers in Psychiatry*, 12, 1–16. <https://www.frontiersin.org/articles/10.3389/fpsy.2021.673089>
- Ticknor, B., & Tillinghast, S. (2011). Virtual reality and the criminal justice system: New possibilities for research, training, and rehabilitation. *Journal For Virtual Worlds Research*, 4(2), 2. <https://doi.org/10.4101/jvwr.v4i2.2071>
- Tyler, N., & Gannon, T. A. (2020). The classification of deliberate firesetting. *Aggression and Violent Behavior*, 101458. <https://doi.org/10.1016/j.avb.2020.101458>
- Tyler, N., Gannon, T. A., Dickens, G. L., & Lockerbie, L. (2015). Characteristics that predict firesetting in male and female mentally disordered offenders. *Psychology, Crime & Law*, 21(8), 776–797. <https://doi.org/10.1080/1068316X.2015.1054382>
- Tyler, N., Gannon, T. A., Ó Ciardha, C., Ogloff, J. R., & Stadolnik, R. (2019). Deliberate firesetting: An international public health issue. *The Lancet Public Health*, 4(8), e371–e372. [https://doi.org/10.1016/S2468-2667\(19\)30136-7](https://doi.org/10.1016/S2468-2667(19)30136-7)
- Wechsler, D. (2011). *Wechsler abbreviated scale of intelligence—Second edition*. PsychCorp. <https://doi.org/10.1037/t15171-000>