



## Clinical Psychology

# Evidence That Tetris Reduces Immediate but Not Subsequent Daily Intrusions of a Trauma Film: A Multilab Replication Study

Ineke Wessel<sup>1</sup>, Julie Krans<sup>2,3,4</sup>, Casper J. Albers<sup>5</sup>, Nidhi Chauhan<sup>1</sup>, Diana S. Ferreira de Sá<sup>6</sup>, Alexander Hauck<sup>6</sup>, Lars Jaswetz<sup>7</sup>, Tanja Michael<sup>6</sup>, Reginald D. V. Nixon<sup>8</sup>, David G. Pearson<sup>9</sup>, Isabelle Slattery<sup>8</sup>, Tom Smeets<sup>10</sup>, Melanie K. T. Takarangi<sup>8</sup>, Thomas R.A. Willems<sup>10</sup>, Kevin van Schie<sup>10</sup>

<sup>1</sup> Department of Clinical Psychology and Experimental Psychopathology, University of Groningen, Groningen, Netherlands, <sup>2</sup> Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands, <sup>3</sup> Pro Persona Research, Pro Persona, Nijmegen, the Netherlands, <sup>4</sup> Behavior, Health, and Psychopathology, KU Leuven, Leuven, Belgium, <sup>5</sup> Department of Psychometrics & Statistics, University of Groningen, Groningen, Netherlands, <sup>6</sup> Clinical Psychology and Psychotherapy, Saarland University, Saarbrücken, Germany, <sup>7</sup> Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands, <sup>8</sup> College of Education, Psychology and Social Work, Flinders Institute for Mental Health and Wellbeing, Flinders University, Adelaide, South Australia, <sup>9</sup> School of Psychology and Sports Science, Anglia Ruskin University, Cambridge, United Kingdom, <sup>10</sup> Department of Medical and Clinical Psychology, Tilburg University, Tilburg, the Netherlands

Keywords: trauma film paradigm, Tetris, intrusive memory, mental imagery, replication

<https://doi.org/10.1525/collabra.130791>

## Collabra: Psychology

Vol. 11, Issue 1, 2025

Reactivating a target memory and subsequently playing the computer game Tetris is thought to reduce intrusive memories and is being explored clinically. However, the current literature on the effect of Tetris on intrusions has limitations. To examine whether the previous finding from experimental research that Tetris reduces trauma film related intrusions is replicable in a large sample, we conducted a preregistered, multi-site study in healthy participants. Experiment 1 ( $N = 141$ ) showed similar intrusion rates of an updated trauma film that was then used in Experiment 2. In line with previous findings, Experiment 2 ( $N = 433$ ) showed that compared to a perceptual vigilance control task, Tetris reduced retrospective ratings of trauma film related intrusions during the task. In contrast to such an immediate effect, the hypothesis that Tetris reduces intrusions over the course of one week was not corroborated. Future research may examine the role of the format of delivering the experiment (i.e., in-person versus online sessions) and type of control condition (i.e., no-task versus active control). This may inform studies on whether, and in what settings, Tetris is a suitable intervention for clinical use.

### Introduction

Many people experience a potentially traumatic event in their lifetime. Depending on the event type, up to 19% of people develop post-traumatic stress disorder (PTSD; R. C. Kessler et al., 2017). One core PTSD symptom is persistent re-experiencing of the traumatic event (i.e., intrusions or flashbacks; American Psychiatric Association, 2013). This may provide a good target for early interventions to prevent PTSD (Holmes et al., 2010).

Holmes and colleagues (2009) were the first to suggest that playing the computer game Tetris after trauma may prevent intrusion development. They tested this possibility using the trauma film paradigm, in which healthy participants watch a film containing aversive scenes and sub-

sequently report film-related intrusions. In Tetris, differently shaped and colored blocks fall down from the top of the screen and must be rotated to neatly interlock at the bottom of the screen and clear filled lines. Holmes et al (2009) found that compared to a control condition, Tetris reduced the number of film-related intrusions during the game, in the week after (per a diary), and on a retrospective questionnaire. The authors proposed that Tetris would demand visuospatial resources in working memory, and that because such resources are limited, the competing demand would interfere with the (re)consolidation of visual memories of the trauma film. This idea has developed into an imagery-competing hypothesis (e.g., Agren et al., 2021;

<sup>a</sup> Correspondence concerning this article should be addressed to Ineke Wessel, Department of Clinical Psychology and Experimental Psychopathology, University of Groningen, Grote Kruisstraat 2-1, 9712 TS Groningen, The Netherlands, [j.p.wessel@rug.nl](mailto:j.p.wessel@rug.nl)  
<https://www.rug.nl/staff/j.p.wessel/>

Ramineni et al., 2023) emphasizing the role of pre-Tetris memory reactivation and mental rotation during gameplay.

A literature search<sup>1</sup> on experimental evidence comparing the effect of Tetris with a control condition rendered 14 published follow-up experiments in 12 articles. Data from 10 experiments (Badawi et al., 2020; Hagenaars et al., 2017; Holmes et al., 2010; James et al., 2015, Experiments 1 & 2; H. Kessler et al., 2020; Lau-Zhu et al., 2019, Experiments 1 & 2; 2021; Page & Coxon, 2017) suggest playing Tetris—and other visuospatially demanding games—effectively reduces long-term intrusions (in a daily diary) whereas four do not (Asselbergs et al., 2018, Experiment 2; Brennen et al., 2021; Brühl et al., 2019; Hemi et al., 2023). These discrepancies may be due to methodological differences (e.g., Brennen et al., 2021) or the absence of a procedure to reactivate the memory shortly before playing Tetris (Brühl et al., 2019). The latter is thought to be crucial for reducing intrusion development (James et al., 2015)<sup>2</sup>. Thus, at first sight, most findings seem to corroborate the effectiveness of Tetris on the occurrence of intrusions.

There are some limitations to this body of literature. First, only four (29%) of the 14 experiments were published independently of the original research team (Asselbergs et al., 2018; Badawi et al., 2020; Brühl et al., 2019; Page & Coxon, 2017). Only two of these independent studies (Badawi et al., 2020; Page & Coxon, 2017) showed that Tetris reduced the number of intrusions compared to a control condition. Second, only one of the 14 experiments was preregistered (Badawi et al., 2020). Importantly, preregistration, where the method and the analysis plans are determined in advance, limits the probability of false positive results (Nelson et al., 2018). Third, the Tetris literature is based on relatively small samples (between 10 to 40 per group in Page & Coxon, 2017, and Brennen et al., 2021, respectively). The larger the sample, the more accurate the effect size estimate, and the more sensitive the research.

Given its simplicity, Tetris seems to be a promising candidate for a cost-effective and easy-to-implement addition to PTSD prevention programs. Indeed, four studies have investigated the effects of Tetris in survivors of potentially traumatic events such as road traffic accidents (Iyadurai et al., 2018), working in an intensive care unit (Iyadurai et al., 2023) and emergency caesarian section (Deforges et al., 2023; Horsch et al., 2017). Together, the results of these studies can be interpreted to favor the idea that Tetris may reduce the negative sequelae of real-life emotional events (although see for a discussion of Deforges et al.'s, 2023 findings Halvorsen et al., 2024 and Deforges et al., 2024). Finally, two studies in PTSD patients show contradictory findings (i.e., a significant reduction in intrusions, H. Kessler et al., 2018, vs. nonsignificant differences, Kehayan et al., 2024).

Given the limitations of the literature, it seems sensible to confirm the effect of Tetris in an analogue study using a large sample. Experiment 2 in the present paper reports an independent, preregistered, large sample multi-site study investigating the effect of Tetris on trauma film-related intrusion development. We based our study on the very first experiment on the effects of Tetris on intrusions (Holmes et al., 2009), but implemented several changes. In that sense our study does not qualify as a direct replication (i.e., a strict repetition of the procedure) and it could be argued that it would be more accurate to characterize it as a conceptual replication. However, Nosek and Errington (2020) argued that a distinction between 'direct' and 'conceptual' replication is suboptimal regarding the role of replication studies in the advancement of knowledge. Instead, they defined replication as "a study for which any outcome would be considered diagnostic evidence about a claim from prior research. (...) To be a replication, 2 things must be true: outcomes consistent with a prior claim would increase confidence in the claim, and outcomes inconsistent with a prior claim would decrease confidence in the claim" (p. 2). If reliable, the imagery-competing hypothesis that is central to the interpretation of the effect of Tetris should be robust enough to survive procedural variation between studies. In support of this we note that none of the 14 published follow-up experiments identified previously exactly replicate the procedure used by Holmes et al (2009), yet positive outcomes are regularly cited as evidence for the effectiveness of Tetris. Studies by Holmes, Brewin and Hennessy (2004) and Bourne et al. (2010) have reported that visuospatial tapping significantly reduces weekly intrusions for a trauma film without using a Tetris condition at all. We believe that if the imagery-competing hypothesis holds, the procedural changes in our study should not hinder testing the claim that "Playing 'Tetris' after viewing traumatic material reduces unwanted, involuntary memory flashbacks to that traumatic film" (Holmes et al., 2009, p. 1). We therefore consider our experiment testing this claim to be a replication study. The results are reported in Experiment 2. First, we describe a study that was the basis for one important change—developing a new trauma film—in Experiment 1.

## Experiment 1

A problem in replication research is that the exact circumstances of an original study are impossible to recreate, if only because participants differ (Nosek & Errington, 2020). The trauma film used in early Tetris studies (Holmes et al., 2009, 2010; James et al., 2015) was developed over 15 years ago and may be perceived differently nowadays due to societal / technical developments increasing the availability and perhaps tolerance of graphic material (e.g., the rise of social media, streaming services). In addition, some scenes

1 We searched PSYCHINFO (October 2024) using the terms: 1) tetris AND intrus\* AND trauma; 2) tetris or visuospatial intervention or visuospatial task\* or computer game\* AND PTSD or intrus\*

2 Therefore, when we refer to "Tetris" in the context of studies relying on the trauma film paradigm in this article, we imply a procedure including reactivation of the target memory prior to gameplay.

in the original film might appear outdated (e.g., clothing style) and / or are of poor quality (e.g., low resolution; pre-production material for a TV commercial). Therefore, we developed a film with more modern clips of better technical quality. In light of the aim of the Experiment 2, that is, to test Holmes et al.'s (2009) claim that a Tetris intervention decreases the number of intrusions, we reasoned that it would suffice if this new film would elicit at least as many intrusions as the original film.

## Method

For the sake of brevity, we report detailed information on the method in a supplement at <https://osf.io/ydhmg>. This supplement contains screening instrument details, video software, scene selection, procedural instructions (film viewing, Intrusion Reminder task, intrusion diary), exploratory measures, and debriefing questions.

## Deviations from Preregistration

The deviations from preregistration are listed in [Table 1](#), using the template by Willroth and Atherton (2023).

## Design and Power Analysis

We used a two-group non-inferiority design (Schumi & Wittes, 2011) with Film (original versus new) as the between-subjects factor and number of intrusions as the dependent variable. We determined the required sample size employing the TOSTer spreadsheet (version 0.4.6) for equivalence tests (Lakens et al., 2018). TOSTer is equipped for statistically testing whether an ES falls significantly within lower and upper bounds relative to 0. Because we tested specifically for non-inferiority we used only a lower bound. Based on previously reported findings (Holmes et al., 2009; James et al., 2015), we set the lower margin of the difference to 0.25, meaning that the new film should yield at least 75% of the number of intrusions of the original film. The other parameters were  $\alpha = 0.1$  (because we used only one bound),  $\beta = 0.2$  and the SD from previous pilot work ( $SD = 3.79$ ; <https://osf.io/9vtbz>). This analysis resulted in a required sample size of 76 participants per condition (total  $N = 152$ ).

## Participants

Participants were 141 undergraduate students from the University of Groningen (UG, Groningen, The Netherlands;  $n = 81$ ) and Saarland University (SU, Saarbrücken, Germany;  $n = 60$ ). They were recruited through local participant recruitment systems and received course credits. Participants with high levels of depression or PTSD, or a diagnosed psychological disorder were screened out (Saarland University; see [supplement](#) and participant flow diagrams at <https://osf.io/n42he/>). Randomization into the Original and New Film conditions—via Qualtrics—was stratified for gender (i.e., female vs non-female participants). Thus, gender was equally distributed across conditions. See <https://osf.io/n42he/> for demographics.

## Materials

Participants were tested in their preferred language (Dutch, German, or English). Unless stated otherwise, we used and translated the materials and instructions used in Holmes et al. (2009) and James et al. (2015; see <https://osf.io/tgxy4/>). The questionnaires and trauma films were presented via Qualtrics (January 2021 [computer software]). All other tasks were constructed with OpenSesame (Version 3.3.6; Mathôt et al., 2012) and delivered through JATOS (version 3.5.5; Lange et al., 2015). For the detailed experimenter protocol, see <https://osf.io/m8fhc>.

**Trauma films.** The two trauma films were both approximately 12 minutes and included 11 separate scenes. The Original Film (Holmes et al., 2009) included scenes depicting motor vehicle accidents, drowning, medical procedures, an animal attack and the aftermath of the Rwandan genocide. The New Film was based on pilot work conducted for the present purpose of replicating the effect of Tetris on intrusions (<https://osf.io/45rsz>) and involved selecting scenes from two films depicting 1) interpersonal violence and 2) accidents/disasters (see [supplement](#) for details on scene content and selection). The scenes reflected either real-life footage or staged events and were freely available on YouTube at the time. Participants were instructed to wear headphones, watch the film closely, and not look away, as if they were a bystander (see [experimenter protocol](#) for details). Participants were warned against participating if they felt easily overwhelmed by graphic film clips.

## Questionnaires and Tasks

**Negative Emotional State.** The Depression Anxiety Stress Scales (DASS-21; Lovibond & Lovibond, 1995) is a 21-item self-report measure with subscales of Depression, Anxiety and Stress containing seven items referring to the past week. Items were scored on 4-point scales (0 = *Did not apply to me at all*; 3 = *Applied to me very much, or most of the time*; score range 0-21). The internal consistency of the subscales in the present sample was fair to acceptable (Depression  $\alpha = .78$ ; anxiety  $\alpha = .64$ ; stress  $\alpha = .78$ ).

**State Mood.** Six slider scales appeared under the header “Right at this very moment I am feeling”. Participants rated the extent to which they felt sad, hopeless, fearful, horrified, anxious, and depressed (range: 0 = *not at all* to 100 = *extremely*; see James et al., 2015). These six scales were averaged to reflect mood at the time of testing. The internal consistency of the combined scores was good (Pre-film  $\alpha = .84$ ; Post-film  $\alpha = .90$ ).

**Film Ratings.** Questions about the film were adapted from James et al. (2015) and rated using slider scales: “How stressful did you find the film?” (0 = *not at all* to 100 = *extremely*), and “How much attention did you pay to the film you just watched?” (0 = *none at all* to 100 = *total attention*). We added the question “To what extent did you close your eyes or look away during the film you just watched?” (0 = *not at all* to 100 = *the whole film*).

**Filler Task.** A filler task was programmed in OpenSesame/Jatos, presenting 15 classical music excerpts. After each excerpt, participants rated the pleasantness of the

**Table 1. Preregistration Deviations Table for Study 1**

Deviations					
#	Details		Original Wording	Deviation Description	Reader Impact
1	Type	Sample	Originally ( <a href="https://osf.io/w7384">https://osf.io/w7384</a> ), we aimed for N = 152, to be tested at 3 sites between February and December 2021. Upon evaluating progress in September 2021, we concluded that reaching the sample size might be difficult, and included a fourth site for data collection. This was described in an addendum dated October 19, 2021 ( <a href="https://osf.io/eaqmf">https://osf.io/eaqmf</a> ; OSF did not have the update function at that time): “A fourth site has been added, data collection will start soon. 4. Flinders University, Adelaide, Australia: First year undergraduate Psychology students have the opportunity to receive course credit for their participation. We aim to collect a minimum of 10 students by the end of Semester 2, 2021.”	Recruiting participants from the fourth site before the end date proved not to be possible – participants just did not sign up, presumably because it was the end of academic year and prospective participants did not need any more study credits.	We believe that not testing participants from the fourth site after all has a minimal impact on the interpretation of the study, especially because it had not been part of the original plan to begin with.
	Reason	Plan not possible			
	Timing	During data collection			
2	Type	Sample	From the addendum ( <a href="https://osf.io/eaqmf">https://osf.io/eaqmf</a> ): “The four sites will collect data (and, if possible, disregard the N per site that was aimed for previously) until November 4, 2021. If by that date the total N across sites is lower than the minimum of 152 valid cases, data collection will continue until this minimum is reached.”	Data collection at the other sites continued until November 19, 2021. We purposely oversampled to increase chances of obtaining the minimum of N = 152 valid cases. At the time of terminating data collection, there were n = 141 valid participants from the other sites, and another 25 participants from the British site. The eligibility of the latter was unclear at that time. Given that the Dutch and German sites combined only had n = 6 non-eligible participants, we thought it safe to assume that we would end up with N = 152 valid cases. As our time schedule for the subsequent Tetris study did not allow further delays, we terminated data collection. Thus, we deviated from the protocol in that we decided to stop data collection without certainty that all cases were indeed valid.	In itself, the decision to stop on the assumption that we had enough valid cases should not impact the interpretation of the study. However, see point 3 below.
	Reason	Plan not possible			
	Timing	During data collection			
3	Type	Sample	The four sites will collect data (and, if possible, disregard the N per site that was aimed for previously) until	Upon checking the British data (November 2022 – March 2023) some issues arose that cast doubts on its reliability:	Omitting the British sample rendered the final sample N = 141, which is lower than the minimum of N = 152. This means that our non-
	Reason	Miscommunication			

	Timing	After results known	November 4, 2021. If by that date the total N across sites is lower than the minimum of 152 valid cases, data collection will continue until this minimum is reached.	<ol style="list-style-type: none"> <li>1. There were several cases in the separate datafiles for sessions 2 and 3 that did not match (i.e., missing or duplicate participant numbers, film condition unknown) rendering the data incomplete. These issues could not be solved because the original experimenter was unavailable;</li> <li>2. all participants had been instructed to start their diaries on the morning of the day after the film viewing session rather than immediately after this session.</li> </ol> <p>Especially because of the latter we decided to disregard the British sample in the final report of the data in the present paper altogether.</p>	inferiority test may have been less sensitive than anticipated in the preregistration.
4	Type	Sample	<p>Cases will be excluded and replaced when (...)</p> <ul style="list-style-type: none"> <li>• The participant is unavailable for the entire day that the post-diary online session should be scheduled (i.e., 7 days after film-viewing)</li> </ul>	<p>A few participants in the Dutch (<math>n = 5</math>) and German (<math>n = 1</math>) sample had a follow up session that was later than exactly 7 days after film-viewing. As full data were available for these participants, we left them in the datafile that is publicly available. They were excluded from the analyses.</p>	As these cases are not included in the analyses, there is no impact on the interpretation of the study.
	Reason	Typo/Error: Procedural error			
	Timing	During data collection			
5	Type	Sample	<p>Planned Sample / Pre-selection rules: Exclusion criteria (screening) * Scores indicating moderate to severe depression (<math>\geq 11</math>) on the Quick Inventory of Depressive Symptomatology (QIDS; Rush et al., 2003; see <a href="http://www.ids-qids.org/">http://www.ids-qids.org/</a>)</p>	<p>We discovered an error in the coding of the items in Qualtrics such that scores for 2 items were affected. (QIDS 15: 0,1,4,2 and QIDS 16: 0,1,1,2 rather than 0,1,2,3). We checked the extent to which this had affected participant selection. For the Dutch sample <math>N = 107</math> were tested for eligibility. Of those, the outcome should have differed for 3 participants: <math>n = 2</math> were tested but should have been excluded (both participants' QIDS scores changed from 10 to 11 after correction). One participant was erroneously excluded (QIDS changed from 12 to 10 after correction). For the German sample <math>N = 68</math> were assessed for eligibility. The coding error affected scores of 7 participants. Their corrected total scores were still <math>&lt; 11</math>, and thus, the outcome did not differ.</p>	Relatively few cases were affected. The impact on interpretation of the study should be minimal.
	Reason	Typo/Error in screening			
	Timing	After data access			
6	Type	Variables	The description of a rating scale for assessing distraction during the Sitting Quietly Period retrospectively is missing.	<p>The Qualtrics Questionnaire contained an item that was administered immediately following the Sitting Quietly period: "To what extent did you deliberately try to distract yourself from film-related thoughts/images?" (0 = not at all to 100 = the whole time). We forgot to mention it in the preregistration; the item had been in the materials from the start.</p>	No impact – it is a variable that can be used for exploring what happened in the period during which participants recorded intrusions – we report the descriptives in table 1.
	Reason	Typo/Error			
	Timing	Before data collection			
7	Type	Study Design	Power Analysis: "Using the same	The SD in the previous study was based on $N = 206$ .	None. The sample size would not have been

		(Power analysis)	parameters (bound = 0.25; $\alpha$ = 0.1, $\beta$ = 0.2) and <b>an SD based on our previous results (study 1; SD = 3.79</b> or 65% of the mean in Holmes et al., 2009) indicated (...)”.	Later we discovered that one case in that data file met exclusion criteria. Therefore, the SD reported for the previous experiment ( <a href="https://osf.io/ju5nf">https://osf.io/ju5nf</a> ) differs slightly (SD = 3.71)	different, the discrepancy is reported here for sake of transparency.
	Reason	New knowledge			
	Timing	After results known			
Unregistered Steps					
#	Details	Original Wording	Original Wording	Unregistered Step Description	Reader Impact
1	Type	Analysis	Unregistered preliminary analyses.	In December 2021 we included the diary data (i.e., primary outcome measure) from all sites, including the UK, in a preliminary non-inferiority analysis in order to determine which film to use in the Tetris study. We knew that $n$ = 6 participants in the Dutch and German sample had a follow up session that was later than exactly 7 days after film-viewing. We anticipated that that could be the case for several participants in this UK sample as well, but at the time we had no information about the exact testing day. In addition, we observed several (extreme) outliers which still had to be checked against the lab logs. Therefore, we looked at the data in three separate non-inferiority tests (1. Including all $N$ = 173; 2. Excluding outliers ( $n$ = 153) and 3. Excluding outliers and all participants who had been tested later than the 8 <sup>th</sup> day). All tests suggested that the new film was non-inferior to the original film. We uploaded a report of these analyses as well as the datafile they were based on to the OSF ( <a href="https://osf.io/n42he/">https://osf.io/n42he/</a> ).	Based on these preliminary analyses we decided to use the new film in the subsequent Tetris study. Later analyses according to the preregistered plan, but excluding the British sample, showed that the non-inferiority analysis on the primary outcome measure (diary intrusions) containing extreme outliers was not statistically significant. Yet, there also was no indication that the original film was superior, rendering this analysis truly inconclusive. Given the statistically significant exploratory follow-up analyses of the diary data excluding outliers, the outcomes regarding initial intrusions, IMS and the emotionality measures (Mood, distress), we believe that we would have decided to use the new film even if the non-inferiority test of the Diary data had been nonsignificant in the preliminary analyses. Thus, we believe that reader impact is minimal.
	Timing	After data access			

music on a 9-point scale (1 = *extremely unpleasant*; 9 = *extremely pleasant*).

**Intrusion Reminder / Sitting Quietly Task.** Participants saw one still image from all 11 film scenes in the same order the scenes appeared in the film, for 3 seconds per image. The images contained as few graphic details as possible (e.g., representing a scene just prior to a moment of impact). Participants were instructed to watch the images closely and then sit quietly for 10 minutes with their eyes closed and record any film-related intrusive images and thoughts by pressing separate keys on the computer keyboard.

**Distraction during Initial Intrusion Recording.** Immediately following the Sitting Quietly period, participants answered the question “To what extent did you deliberately try to distract yourself from film-related thoughts/images?” (0 = *not at all* to 100 = *the whole time*).

**Intrusion Diary.** The digital Intrusion Diary (<https://osf.io/7sv6e>) was a Microsoft Word file consisting of instructions, tables for each day (divided into morning, afternoon, evening) where participants recorded intrusion type (image; thought; or a combination) and content. Participants were instructed how to use the Intrusion Diary, and to complete the digital diary each morning, afternoon and evening. Participants received a daily email reminder every afternoon at 5 PM that contained a link to an online Intrusion Summary (see [supplement](#)).

**Diary Compliance Ratings.** To measure diary compliance, participants completed three questions using Qualtrics slider scales: “To what extent is the following true: I have been unable (or forgotten) to record my unpleasant thoughts and images in the diary” (0 = *not at all true of me* to 100 = *extremely true*), “Please indicate how accurate you think the diary you completed is” (0 = *not at all accurate* to 100 = *extremely accurate*), and “To what extent did the daily email affect the frequency of the intrusions of the film?” (0 = *not at all* to 100 = *extremely*). We added one open question: “Do you have any suggestions or comments regarding the diary?”.

**Impact of Movie Scale.** The Impact of Movie Scale (IMS; see James et al., 2015) assesses film-related distress symptoms during the week after watching the film. It contains 22 items rated on a 5-point likert scale (0 = *not at all*; 4 = *extremely*; range 0 - 88). The internal consistency of the IMS in the present sample was good,  $\alpha = .83$ .

## Procedure

Due to the social distancing measures related to COVID-19, participants were tested in two online video conferencing meetings. After digitally signing an informed consent form, participants completed screening questionnaires (see [supplement](#)) and the DASS-21. While participants completed the DASS, the experimenter checked the screening questionnaire for eligibility. Non-eligible participants were thanked and debriefed. Eligible participants continued with the Mood scales, and then watched one of the two trauma films, followed by the Mood scales again and the Film ratings. Next, the Filler task and the Intrusion Reminder / Sitting Quietly task were presented. Afterwards,

the experimenter explained how participants record their intrusions in the diary. Participants returned for an online follow-up session one week later, on the same weekday. The follow up included the Intrusion Summary of the present day, the Diary Compliance ratings, the IMS, and exploratory / debriefing questionnaires (see [supplement](#)). Because Holmes et al. (2009) suggested playing Tetris mitigates intrusions, participants played a generic online Tetris game for 3 minutes as a neutralizer task (Tetris holding, <http://tetris.com/play-tetris/>) and were then debriefed.

## Scoring and Analyses

**Diary Scoring.** The diaries were scored by counting the number of Images (I) and Image-Thought (IT) combinations entered each day into a total Image-Based Diary Intrusions score. The raters checked the content of the intrusions noted on the last page of the diary; entries not matching the content of the film were disregarded.

**Preregistered Analyses.** Descriptives were obtained using SPSS version 29. We checked the three intrusion variables of interest (i.e., Image-based Diary Intrusions, IMS scores and Initial Intrusions) for outliers. Outliers were defined as scores that were 1.5x Interquartile Range below the first or above the third quartile. None of these outliers were due to deviations from the protocol according to the logs recorded during the experimental sessions, and thus these outliers were kept in the analyses.

We used Welch *t*-tests to examine whether the mean Image-Based Diary, Initial Intrusion counts, and IMS-scores in the New Film condition reflected at least 75% of the means in the Original Film condition. These non-inferiority analyses were conducted using the TOSTer spreadsheet (Version 0.4.6, see Lakens et al., 2018). First, for each variable, 25% of the mean value observed in the Original Film condition was calculated and subtracted from zero to obtain lower bounds for the difference between the original and new film conditions. Next, this lower bound as well as the means and standard deviations for each condition were entered into the TOSTer spreadsheet to test the difference between the groups against the lower bound value with a one-sided Welch *t*-test. Because we were interested in non-inferiority, we set the upper bound to zero and disregarded the outcome of the second one-sided *t*-test of the mean difference between groups against this upper bound in the spreadsheet.

**Exploratory Analyses.** We did not preregister group comparisons on variables other than the intrusion measures. Yet, to investigate whether the new film elicited at least as much distress as the original film, we conducted non-inferiority analyses of the mean difference scores in Mood State prior and after watching the film as well as the retrospective measure of Distress during Film viewing. We followed a similar strategy as for the intrusion variables and set the lower bound at 0 - 25% of the mean value observed in the Original Film condition. Given the distributions of the intrusion variables (especially Image-Based Diary Intrusions and Initial Intrusions) were skewed, we repeated the non-inferiority analyses without outliers to determine the robustness of the findings.

**Table 2. Descriptive Statistics for DASS Subscale Scores, Mood and Disgust Ratings before and after Film Watching, and Checks for Film Watching, Distraction during Initial Intrusion Recording and Diary Compliance in the Original Film and New Film Groups in Study 1**

	Original film (n = 68)				New Film (n = 73)			
	M	SD	Median	IQR	M	SD	Median	IQR
DASS								
Depression	1.82	2.41	1	2	2.18	2.34	1	3
Anxiety	1.31	2.02	1	2	1.62	1.83	1	2
Stress	2.43	2.25	2	4	3.96	3.22	4	2
Mood								
Pre-film	7.17	10.6	3.42	8.88	8.68	9.52	4.67	11.4
Post-film	20.6	15.6	17	19.6	24.5	19.6	19.2	26.4
Disgust								
Pre-film	1.15	2.82	0	1	1.15	2.66	0	1
Post-film	40.6	30.4	39	60	51.6	32.0	41	62
During the film								
Distress	46.2	22.9	42	34	52.6	24.2	55	42
Attention	93.7	9.77	97	8	94.1	7.01	96	9
Eyes closed	5.15	8.29	1.50	8	9.36	18.40	3	10
Distraction during Intrusions <sup>1</sup>	43.1	27.3	40.0	42	39.5	25.8	35	25.8
Diary checks								
Unable/Forgotten	8.06	15.6	1	10	8.55	12.8	2	13
Accuracy	83.7	16.9	86	15	84.7	12.3	86	15
Daily Emails	20.8	26.4	10	35	23.7	26.8	12	39

<sup>1</sup> n = 1 missing in the Original Film condition

Note. DASS = Depression Anxiety Stress Scales

## Results

### Compliance and Manipulation Checks

Table 2 summarizes the descriptives of the DASS, Mood State Ratings, Film Ratings and Diary Compliance ratings for each condition separately.

The mean differences in mood before and after film watching were  $m = 15.8$  ( $SD = 18.6$ ) and  $m = 13.4$  ( $SD = 12.8$ ) for the new and original film, respectively. Exploratory analyses showed that the observed mean difference between conditions ( $m = 2.42$ ; 90% CI [-2.01; 6.85]) was significantly higher than the lower bound of -3.35,  $t(127.93) = 2.16$ ,  $p = .02$ .

Likewise, the observed mean difference in distress during film watching between conditions ( $m = 6.36$ ; 90% CI [-0.21; 12.9]) was significantly higher than the lower bound of -11.6,  $t(138.96) = 4.51$ ,  $p < .001$ . Thus, the new film elicited at least as much mood change and distress as the original film. Inspection of Table 2 does not give rise to suspicions that the new film was inferior to the original film on any of the other variables.

### Non-inferiority Analyses

Table 3 shows the descriptives of the three intrusion measures of interest.

The primary outcome was the number of Image-based Intrusions in the diary. The non-inferiority test showed that the observed difference between the conditions ( $m_{\text{dif}} = 0.03$ , 90% CI [-1.91; 1.97]) did not statistically significantly exceed the lower bound of -1.13, Welch  $t(131.5) = 0.99$ ,  $p = .162$ . In contrast, the inferiority tests for the secondary outcome measures were statistically significant. That is, the observed difference between the groups in IMS scores ( $m_{\text{dif}} = 2.26$ , 90% CI [-0.02; 4.54]) exceeded the lower bound of -2.44, Welch  $t(139) = 3.40$ ,  $p < .001$ . Likewise, the observed difference between the groups in Initial Intrusions ( $m_{\text{dif}} = 2.37$ , 90% CI [-1.49; 6.09]) was higher than the lower bound of -3.33, Welch  $t(102.9) = 2.46$ ,  $p = .008$ . Thus, for IMS and Initial Intrusions, the new film was not statistically inferior to the original film, meaning that it rendered at least as many intrusions. However, the inferiority test for the Image-based Diary Intrusions was inconclusive: we cannot reject the hypothesis that the new film rendered fewer intrusions than the original film.

### Exploratory Analyses

Because we observed outliers in the data, we removed those cases (Image-based Diary Intrusions:  $n = 10$ ; IMS:  $n = 2$ ; Initial Intrusions  $n = 7$ ) and conducted exploratory inferiority analyses. As for the Image-based Diary Intrusions, the observed mean difference ( $m_{\text{dif}} = 0.42$ , 90% CI [-0.39; 1.23])

**Table 3. Descriptive Statistics for Image-Based Intrusions in the Diary, IMS and Initial Intrusions in the Original Film and New Film Groups in Study 1**

	Original Film ( <i>n</i> = 68)				New Film ( <i>n</i> = 73)			
	<i>M</i>	<i>SD</i>	Median	IQR	<i>M</i>	<i>SD</i>	Median	IQR
Image-based Intrusions	4.50	7.51	2	4	4.53	6.33	3	5
IMS	9.74	7.88	8.5	10.8	12.0	8.50	10	12
Initial Intrusions <sup>1</sup>	13.3	9.12	11	12	15.6	15.9	9	15

<sup>1</sup> Due to missing values: Original Film *n* = 62; New Film *n* = 65

Note. Image-based Intrusions = Images and Image/Thought combinations in the Diary; IMS = Impact of Movie Scale; Initial Intrusions = Number of Image-Based Intrusions during the Sitting Quietly Period

now significantly exceeded a lower bound of -0.71, Welch  $t(128.8) = 2.31$ ,  $p = 0.011$ . The inferiority analyses of the IMS scores showed that the observed difference between conditions ( $m_{\text{dif}} = 2.37$ , 90% CI [0.29; 4.45]), remained statistically significantly higher than a lower bound of -2.44, Welch  $t(136.6) = 3.83$ ,  $p < .001$ . In contrast, the observed difference in Initial Intrusions ( $m_{\text{dif}} = -2.21$ , 90% CI [-4.92; 0.50]), did not significantly differ from a lower bound of -3.32, Welch  $t(117.9) = 0.68$ ,  $p = .250$ .

## Discussion

Overall, the non-inferiority analyses suggested the new film did not elicit markedly fewer intrusions than the original film, although in part, the results seemed to depend on removing outliers. The analysis of the main outcome variable—Image-based Diary Intrusions—was inconclusive, whereas removing outliers suggested the new film elicited at least as many intrusions as the original film. The other intrusion measures, and the extent to which the films induced mood state and distress, provided no evidence that the new film was inferior to the original film. We therefore assume that the new film is a suitable replacement.

## Experiment 2

In Experiment 2 we compared the effect of Tetris and an active control condition on film-related intrusions. Apart from a different trauma-film, our study procedure included two other notable differences compared to the original Tetris study (Holmes et al., 2009). First, due to lockdowns during the COVID-19 pandemic, experimenters tested participants in online video conferencing sessions rather than in the laboratory. Secondly and relatedly, because we anticipated that asking participants to sit quietly (cf. Holmes et al., 2009) for 10 minutes in their home environment would expose them to distractions that would be hard to resist, we employed a perceptual vigilance control condition. The task required watching a black computer screen and responding to a red dot that occurred at random during 30 second intervals. We chose this cognitively low-demanding task in the visual domain with the idea that it would possess a “game-like” quality to some extent, but lack the defining features of Tetris (e.g., multiple shapes and colors, ongoing visual stimulation, and mental rotation).

Although our procedure differed from that of Holmes et al. (2009), we expected that given the imagery-competing hypothesis, the outcome of our experiment would be informative for the claim that Tetris after a trauma film reduces intrusions of that film.

## Method

### Deviations from Preregistration

Deviations from preregistration (using the template in Willroth & Atherton, 2023) are listed in Table 4.

### Design and Power Analysis

Six sites used a two-group between-participants design. We aimed for a sample size that could detect the effect size of Cohen's  $d = 0.91$  reported in Holmes et al. (2009) with 95% power and a one-tailed  $\alpha$  adjusted for 3 comparisons (3 dependent variables) of 0.0167 in each of the six labs. An a priori power analysis (G\*power v3.1.3; Faul et al., 2009) for a linear model yielded a minimum of  $n = 72$  per site, and thus, a planned total sample size of 432. A sensitivity analysis with the same parameters (power = 95% and  $\alpha = .0167$ ) indicated this number could detect Cohen's  $d = 0.36$ . We also planned negative binomial GLM analyses. Exact power calculations are complicated, but these models are at least as powerful as the linear model (Cundill & Alexander, 2015).

### Participants

Participants ( $N = 433$ ) were undergraduate students recruited at Anglia Ruskin University (ARU, Cambridge, UK), Flinders University (FU, Adelaide, Australia), University of Groningen (UG, Groningen, The Netherlands), Radboud University (RU, Nijmegen, The Netherlands); Saarland University (SU, Saarbrücken, Germany), and Tilburg University (TU, Tilburg, The Netherlands) through local participant recruitment systems, and were tested in English (except for participants at Saarland University who were tested in German). See <https://osf.io/3mgmv/> for demographics per site.

### Materials

Experiment 2's materials were identical to Experiment 1, with the following exceptions: (1) the exploratory disgust

Table 4. Preregistration Deviations Table for Study 2

Deviations					
#	Details		Original Wording	Deviation Description	Reader Impact
1	Type	Sample	<b>Justify planned sample size</b> We aim for a study that is sensitive enough to detect this effect size in each of the 6 labs (...) a minimum of N = 72 per site. <b>Describe data collection termination rule</b> If the researchers at a particular site foresee that they are not likely to reach their required sample size at the end of their projected data collection period (e.g., end of semester), other sites will try to test more participants. We anticipated that the data collection would end on August 1, 2022. However, the projected N was not reached at that date. We deemed it feasible to carry on collecting data (with the exception of RU). We will extend data collection until the end of March 2023 (rather than the dates of December 31, 2022 and February 28, 2023 in earlier versions of the preregistration).	There were two major deviations from the original sampling plan. 1) It appeared that multiple sites had to extend data collection beyond the originally planned date of August 1, 2022 because they did not reach their projected N. RU had tested n = 43 eligible participants, but it was not feasible to continue data collection. UG had already finished, and re-opened data collection in order to compensate for this. Thus, the RU sample is actually a mixed RU/UG sample. 2) Between December 31, 2022 and March 31, 2023, several sites carried on collecting data, and, if possible, tested more participants than their projected N. The reason was that it was uncertain whether the ARU would reach their final N. By March 31, 2023, ARU had, indeed, not reached their projected sample size but students were still collecting data for thesis projects that would extend into the summer semester. We waited until data collection for those student projects was complete. Due to an oversight, the preregistration was not updated.	The deviations aimed at reaching the preregistered total N of 432, so in that respect, impact is minimal. The subsamples, however, differ in size, and this may render the analyses with site as covariate less reliable.
	Reason	Plan not possible			
	Timing	During data collection			
2	Type	Sample	Cases will be excluded and replaced when (...) - The participant is unavailable for the entire day that the post-diary session should be scheduled (i.e., 7 days after film-viewing), unless they hand in the diary on this day	A total of n = 13 participants had a follow up session that was later than day 8 (= exactly 7 days after film-viewing.) As full data were available for these participants, we left them in the publicly available datafile. In addition to the preregistered analyses (excluding these cases), we repeated the analyses including these cases. The results are available on <a href="https://osf.io/3mgny/">https://osf.io/3mgny/</a> (called "lenient analyses")	As these cases are not included in the analyses reported here, there is no impact on the interpretation of the study.
	Reason	Typo/ Error Procedural error			
	Timing	During data collection			
3	Type	Sample	<b>Planned Sample / Pre-selection rules:</b> Exclusion criteria (screening) * Scores indicating moderate to severe depression ( $\geq 11$ ) on the Quick Inventory of Depressive Symptomatology (QIDS; Rush et al., 2003; see <a href="http://www.ids-qids.org/">http://www.ids-qids.org/</a> )	We discovered an error in the coding of the items in Qualtrics such that scores for 2 items were affected. (QIDS 15: 0,1,4,2 and QIDS 16: 0,1,1,2 rather than 0,1,2,3). For the sites where data collection had started, we checked the extent to which this had affected participant selection. In total, n = 2 were tested but should have been excluded (i.e., a total score of 11 instead of 10).	Few cases were affected. The impact on interpretation of the study should be minimal.
	Reason	Typo/ Error in screening			
	Timing	During data collection			
4	Type	Analysis	<b>Data-based outlier criteria</b> The experimenters will note in a lab journal whether they encountered any circumstances of note ('oddities') while	We had already run the analyses on the whole sample and the sample excluding all outliers, when we realized that the preregistration required consulting the lab journals in	We believe that reader impact is minimal, because we report all analyses.
	Reason	New			

		knowledge	testing participants. Based on these oddities in the lab journal we will decide whether or not to exclude outliers. When no oddities are recorded, the data are not excluded.	the latter case to determine what exactly should count as an outlier. We deviated from this step and did not run new analyses because 1) deciding who to exclude based on notes is ambiguous (unless the circumstances were extreme and testing did discontinue in those cases anyway); and 2) excluding all outliers is a more rigorous approach. We report both analyses in the paper.	
	Timing	After data access			
Unregistered Steps					
#	Details		Original Wording	Unregistered Step Description	Reader Impact
1	Type	Sample	It was nowhere explicitly mentioned in the preregistration or protocol that participants should be tested individually. We did, however, expect that to be the case for all participants.	During processing of the ARU data, we discovered that some student experimenters had tested their participants in online groups. This happened in 34 cases. These cases were identified and excluded from the analyses.	This contributed substantially to a lower final number of eligible ARU participants than preregistered (i.e., $n = 47$ rather than 72). For the total sample across sites this should have minimal reader impact because we reached the minimum sample size of $N = 432$ . The analyses with site as a covariate may be less reliable.
	Timing	After data access			

item was removed from the state mood scales; (2) the item assessing distress in the Film Ratings was reworded; (3) the 10 minute Sitting Quietly period following the Intrusion Reminder task was replaced by one of two tasks (i.e., Tetris or Perceptual Vigilance Control); (4) the distraction question about the Sitting Quietly period in Experiment 1 was replaced by retrospective ratings (including initial intrusions); (5) the different sites used different exploratory measures (see <https://osf.io/ut957>) and (6) daily emails were sent at 8 AM each day and contained a reminder to report intrusions in the diary rather than a link to an intrusion-monitoring questionnaire.

**Film Ratings.** The distress item now read “How distressing did you find the film you just watched?” (0 = *not at all* to 100 = *extremely*) to better match earlier studies (James et al., 2015).

**Tetris.** We used a non-public research version of the computer game Tetris, which was programmed by The Tetris Company (see <https://tetris.com/about-us>) with the same specifications as the version they programmed for Badawi et al. (2020)). During the game, participants see differently colored geometric blocks (i.e., tetrominoes) fall from the top of the screen to the bottom. One block is always in play and the upcoming three blocks are visible at the upper right-hand side of the screen. Following the instructions in James et al. (2015, see <https://osf.io/nur8q/>), participants were explicitly instructed to simultaneously focus on these upcoming blocks and plan for their descent. The goal of Tetris is to complete as many horizontal lines as possible by neatly interlocking the blocks without gaps. Completed horizontal lines disappeared from the screen and points were awarded. All participants started at level 1 and moved up one level for every 10 lines cleared (with level 5 being the maximum). Block dropping speed increased as more rows were completed. Participants used the arrow keys to move the blocks left or right, rotate them 90 degrees, or move them down. We used the instructions in James et al. (2015, see <https://osf.io/nur8q/>), asking the participants to focus on the blocks that would be falling immediately after the one that was being played (displayed at the top right of the screen, outside the playing field). Participants were told to try and work out in their mind’s eye where best to place and rotate these blocks together with the block that was ‘in play’ to make the most complete lines and get the best score. Our Tetris version also stimulated visual mental rotation because as a default, the ghost tetromino (i.e., a faintly colored copy of the piece in play showing where it would land at the bottom of the playing field) option was turned off.

**Perceptual Vigilance Task.** As an active control task, we used a perceptual vigilance task (adapted from Wilkinson & Houghton, 1982) programmed in PsyToolKit (Stoet, 2010; <https://www.psychtoolkit.org/>). This task required participants to look at a black screen and press the spacebar every time a red circle was presented. The circle was presented randomly at a low frequency (20 times in total). Each 30 sec trial started and ended with a 2 second wait period. The red circle appeared at a random time within the re-

maining 26 seconds and disappeared upon the participant’s response or after a maximum of 2 sec.

**Retrospective Task Ratings.** The task ratings started with a text field where participants could enter their high score on either Tetris or the Vigilance task. Two VAS, with a slightly different wording for the Tetris and Vigilance conditions asking about 1) Initial Intrusions (“How often did mental images of the film spontaneously pop into your mind while playing the game / during the task you just did?”; 0 = *not at all* to 100 = *the whole time*) and 2) task difficulty (“How difficult or easy did you find the game you just played / task you just did?”; 0 = *not difficult at all/easy* to 100 = *extremely difficult/hard*).

## Procedure

Experiment 2’s procedure was identical to Experiment 1 except for the elements related to the intervention. To keep the time between the Intrusion Reminder task and the intervention as brief as possible (and similar for the Tetris and the Perceptual Vigilance Control groups), after screening, all eligible participants received instructions for playing Tetris and played the game for 3 minutes (cf. Badawi et al., 2020). After the Intrusion Reminder task, participants were randomly assigned to Tetris or the Perceptual Vigilance Task. For both tasks participants were instructed to memorize (or write down) their high score as it appeared on the screen after they finished the task to be able to report it later on. This instruction was intended to enhance task engagement. After either intervention, participants completed the retrospective task ratings. See <https://osf.io/m3f7n> for the experimenter protocol.

## Analyses

**Preregistered Analyses.** Descriptives for demographics and DASS were obtained using SPSS version 29. All other analyses were carried out in R (version 4.3.2, R Core Team, 2023), using the package robustbase (version 0.99, Maechler et al., 2023) for medcouple calculations and the packages apa (version 0.3.4, Gromer, 2023), broom (version 1.0.5, Robinson et al., 2023) and crosstable (version 0.6.2, Chaltiel & Hajage, 2023) for styling the output. We report analyses on the pooled data for the six sites here. The analyses were repeated for each individual site; this output can be found on <https://osf.io/3mgv/>.

Similar to Experiment 1, the three intrusion variables of interest (i.e., Image-based Diary Intrusions, IMS scores and Initial Intrusions) were checked for outliers. Because we expected extreme values based on theory and pilot work, we analyzed the data with and without outliers. The boundaries for determining outliers were adjusted for skewness by means of the medcouple (MC) score (Brys et al., 2004) and scores above  $Q3 + 1.5 * \exp(3MC) * IQR$  were flagged (Hubert & Vandervieren, 2008).

We tested the difference between the Tetris and Perceptual Vigilance Control conditions on three outcomes of interest and adjusted alpha to .0167. Because Image-based Diary Intrusions is a count variable, we tested the difference between the Tetris and Perceptual Vigilance Con-

**Table 5. Descriptive Statistics for DASS Subscale Scores, and Mood Ratings before and after Film Watching, and Checks for Film Watching, Task Difficulty and Diary Compliance in the Tetris and Perceptual Vigilance Control Groups in Study 2**

	Tetris ( <i>n</i> = 209)				Perceptual Vigilance Control ( <i>n</i> = 224)			
	<i>M</i>	<i>SD</i>	Median	IQR	<i>M</i>	<i>SD</i>	Median	IQR
DASS								
Depression	1.75	2.03	1	3	1.94	2.17	1	2
Anxiety	1.64	1.94	1	3	1.53	2.02	1	2
Stress	2.86	2.34	2	4	3.14	2.74	2	4
Mood								
Pre-film	6.62	7.76	4	9	6.41	8.00	4	7.80
Post-film	22.14	18.62	16.50	27	23.19	19.79	16.92	25.88
During the film								
Distress	54.16	26.99	60	43	55.13	27.18	60	45
Attention	93.95	8.22	96	10	95.04	6.78	98	8
Eyes closed	7.89	14.28	3	9	7.22	13.29	2	10
Task Difficulty	34.06	21.93	31	30	16.70	20.35	10	21.25
Diary checks								
Unable/Forgotten	11.79	21.20	1	10	8.21	16.56	1	10
Accuracy	85.72	15.43	90	16	86.81	15.05	91	19
Daily Emails	23.31	27.71	11	40	19.95	25.65	8	31.25

Note. DASS = Depression Anxiety Stress Scales

trol conditions by means of a negative binomial regression analysis. For IMS scores we conducted independent student *t*-tests. We conducted Welch *t*-tests for Initial Intrusions, because the assumption of homoscedasticity was violated.

To control for systematic differences in locations, the negative binomial regression analysis was repeated including test site as a covariate, with ARU serving as the reference group. Analyses for IMS scores and Initial Intrusions were repeated as 2 (Condition)  $\times$  6 (Location) Analyses of Variance (ANOVA), treating test site as a covariate. We report all analyses with and without the outliers.

**Exploratory Analyses.** We did not preregister group comparisons on variables other than the intrusion measures. We conducted two exploratory analyses. The first was a within participants *t*-test of the difference between Mood State prior and after watching the film to see whether it significantly induced negative mood. Second, we looked at differences in Difficulty ratings of the Tetris/Vigilance tasks with a between groups *t*-test. Both analyses were repeated controlling for test site, using a one-way ANOVA including Location as factor and a 2 (Condition)  $\times$  6 (Location) ANOVA, respectively.

## Results

### Compliance and Manipulation Checks

Table 5 summarizes the descriptives of the DASS, Mood State Ratings, Film Ratings and Diary Compliance ratings for both conditions separately.

The film successfully induced negative mood, as was evident from a paired student *t*-test showing a pre-to-post film increase in negative state mood,  $t(431) = 18.99$ ,  $p < .001$ , Cohen's  $d = 0.91$ , 95% CI [0.80; 1.03]. The one-way ANOVA on pre-to-post differences controlling for Location showed differences between sites,  $F(5, 426) = 5.85$ ,  $p < .001$ ,  $\eta_p^2 = .06$ , but the increase in negative state mood still was highly significant,  $F(1, 426) = 381.05$ ,  $p < .001$ ,  $\eta_p^2 = .47$ .

Furthermore, we observed that participants had more difficulty performing Tetris compared to Perceptual Vigilance Control,  $t(431) = 8.55$ ,  $p < .001$ , Cohen's  $d = 0.82$ , 95% CI [0.62; 1.02]. A 2 (Condition)  $\times$  6 (Location) ANOVA showed that the effect of Condition remained statistically significant,  $F(1, 426) = 74.21$ ,  $p < .001$ ,  $\eta_p^2 = .15$ , while controlling for Location,  $F(5, 426) = 2.34$ ,  $p = .041$ ,  $\eta_p^2 = .03$ .

### Intrusion Measures

Table 6 summarizes the descriptive statistics of the intrusion measures.

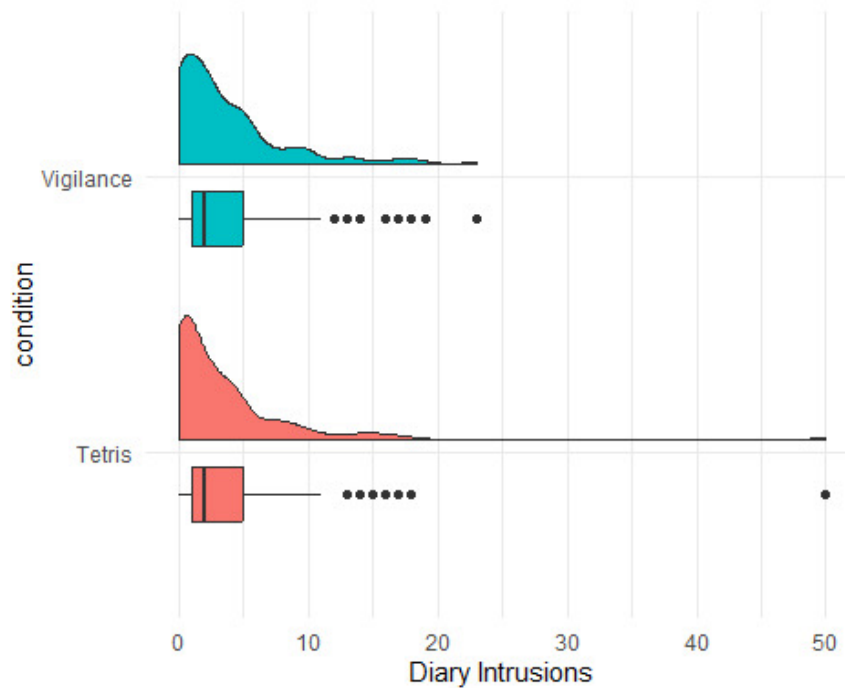
Figures 1-3 show the distributions of the number of Image-Based Diary Intrusions, IMS Scores and ratings of Initial Intrusions in the Tetris and the Perceptual Vigilance Control conditions separately.

**Image-Based Diary Intrusions.** The negative binomial regression analysis revealed the number of Image-Based Diary Intrusions was not statistically significantly associated with playing Tetris (coefficient estimate of 0.04,  $SE = 0.11$ ,  $z = 0.38$ ,  $p = .702$ ) when compared to the Perceptual Vigilance Control condition. Using the medcouple technique, outliers were flagged if scores were higher than

**Table 6. Descriptive Statistics for Image-Based Intrusions in the Diary, IMS and Initial Intrusions in the Tetris and Perceptual Vigilance Control Groups in Study 2**

	Tetris ( <i>n</i> = 209)				Perceptual Vigilance Control ( <i>n</i> = 224)			
	<i>M</i>	<i>SD</i>	Median	IQR	<i>M</i>	<i>SD</i>	Median	IQR
Image-Based Intrusions	3.44	4.99	2	4	3.59	4.12	2	4
IMS	11.90	9.14	10	12	11.87	9.00	10	11
Initial Intrusions	5.88	12.14	0	5	19.91	23.76	10	32

Note. Image-based Intrusions = Images and Image/Thought combinations in the Diary; IMS = Impact of Movie Scale; Initial Intrusions = Number of Image-Based Intrusions during the Intervention (i.e., Tetris vs. Perceptual Vigilance Control)

**Figure 1. The Distributions, Medians and Inter-Quartile Ranges of Image-Based Diary Intrusions in the Tetris and Perceptual Vigilance Control Conditions**

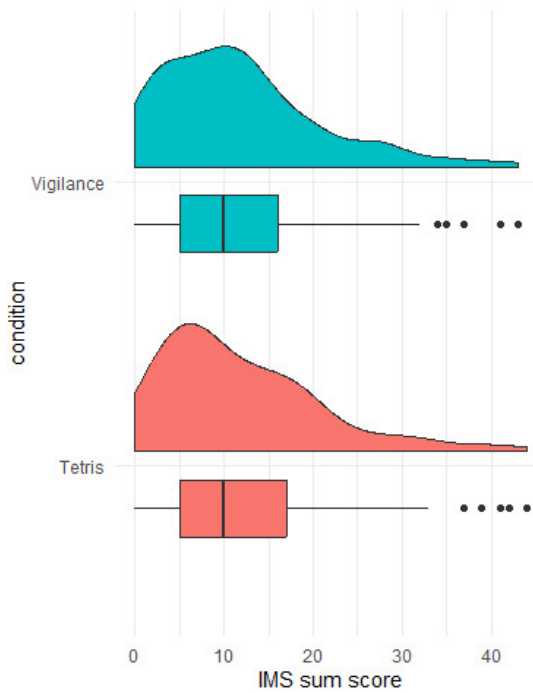
9.374, which happened in 36 cases. Excluding medcouple-outliers did not change the results of the negative binomial regression analysis (coefficient estimate of 0.07,  $SE = 0.11$ ,  $z = 0.67$ ,  $p = .505$ ).

Including test-site as covariate showed Location did not influence the results of the analyses on the full sample (coefficient estimate of 0.02,  $SE = 0.11$ ,  $z = 0.16$ ,  $p = .874$ ). Likewise, including Location as a covariate had no discernible impact on the overall effect in the sample excluding outliers (coefficient estimate of 0.06,  $SE = 0.10$ ,  $z = 0.56$ ,  $p = .577$ ).

**IMS Scores.** Levene's test for assumption of homogeneity of variances was met for IMS scores,  $F(1, 431) = .10$ ,  $p = .753$ . The  $t$ -test showed the difference between Tetris and Perceptual Vigilance Control was statistically nonsignificant,  $t(431) = 0.04$ ,  $p = .969$ , Cohen's  $d < 0.01$ , 95% CI [-0.18; 0.19]. Excluding 35 medcouple-outliers (i.e., scores above 26.48) did not change the results,  $t(396) = 0.16$ ,  $p = .873$ , Cohen's  $d = 0.02$ , 95% CI [-0.18; 0.21].

Controlling for test site, we performed a 2 (Condition)  $\times$  6 (Location) ANOVA, which showed no statistically significant effect of Tetris on IMS intrusions, neither in the full sample,  $F(1, 426) < 0.01$ ,  $p = .968$ , nor when outliers were excluded,  $F(1, 391) = 0.03$ ,  $p = .872$ . The results of the ANOVAs should be interpreted with mild caution because the QQ plots indicate slight violations from normality. However, given our sample size, the central limit theorem ensures the normality assumption (approximately) holds.

**Initial Intrusions.** Given the right-skewed distribution for initial intrusions, we used the medcouple technique to detect outliers. Seventeen cases were flagged because they scored above 67.41. Because Levene's test indicated a violation of the assumption of homogeneity of variances,  $F(1, 429) = 61.96$ ,  $p < .001$ , we report Welch  $t$ -tests. These  $t$ -tests revealed a statistically significant difference in initial intrusions between the Tetris and Perceptual Vigilance Control conditions, regardless of whether outliers were included,  $t(335.46) = -7.8$ ,  $p < .001$ , or excluded,  $t(350.41) = -6.71$ ,  $p$



**Figure 2. The Distributions, Medians and Inter-Quartile Ranges of IMS Intrusions in the Tetris and Perceptual Vigilance Control Conditions in Study 2**

< .001 with the Tetris group showing the lowest number of initial intrusions<sup>3</sup>.

The 2 (Condition) x 6 (Location) ANOVA showed no evidence that adding test location to the analyses changed how Tetris affected initial intrusions,  $F(1, 424) = 58.82, p < .001$ , even when outliers were excluded,  $F(1, 407) = 45.11, p < .001$ . Similar to the IMS Intrusions analysis, the outcomes of the ANOVA on Initial Intrusions should be approached with some caution because the QQ plots indicate slight deviations from normality. Nevertheless, according to the central limit theorem the assumption of normality holds (approximately) given our sample size.

### General Discussion

This independent, multilab replication examined whether Tetris after viewing a distressing film reduces intrusions of that film (Holmes et al., 2009). After developing an alternative trauma film (Experiment 1) we tested the effect of Tetris against an active perceptual vigilance control condition in Experiment 2. We assessed the number of intrusions through a retrospective rating, a one-week intrusion diary, and a questionnaire at one-week follow-up.

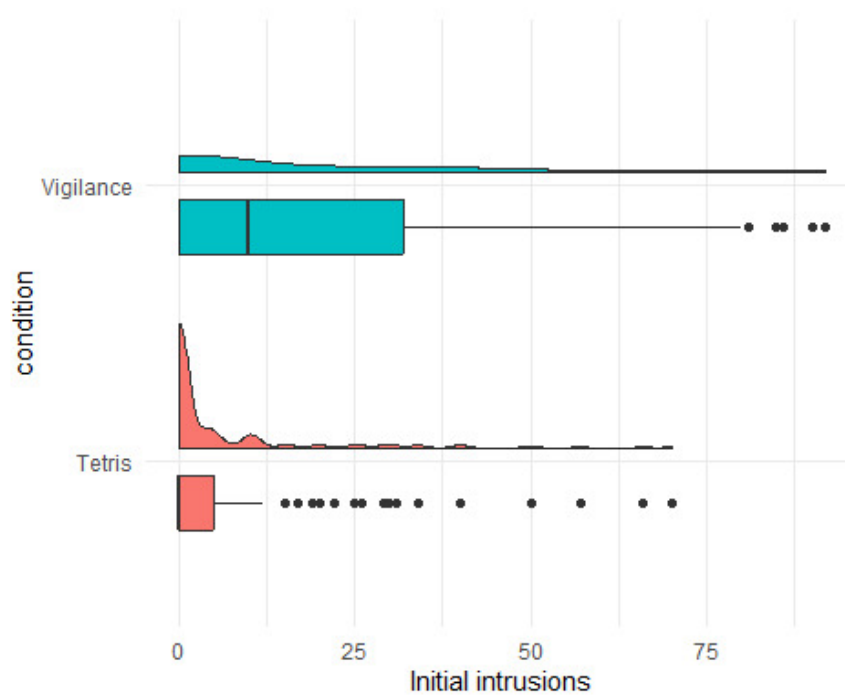
Our results replicate previous findings (e.g., Badawi et al., 2020; Holmes et al., 2009, 2010) that the Tetris condition was associated with fewer *initial* intrusions than the perceptual vigilance control condition, despite procedural

and methodological differences compared to previous studies. In contrast, we did not replicate the beneficial effect of Tetris on *subsequent* diary intrusions and retrospective film-related distress (IMS) seen in these prior experiments. There are several types of explanations for these mixed results.

The first type focuses on methodological differences. To begin with, we used a new trauma film to induce intrusions. This new film provoked fewer intrusions than the film used in earlier studies (Holmes et al., 2009, 2010; James et al., 2015), but was not inferior to that original film (cf. Experiment 1). Moreover, the new film induced an average number of intrusions, comparable to other studies (e.g., Arnau-dova & Hagenars, 2017; van Schie et al., 2019) and had a significant emotional impact in both of the present experiments. Furthermore, in previous studies a lower number of intrusive memories did not prevent the primary intervention from effectively modulating intrusions (e.g., Hagenars et al., 2008; Krans et al., 2009), including Tetris (Hagenars et al., 2017). Nevertheless, we cannot exclude the possibility that a relatively low frequency and resulting skewed distribution may have affected the sensitivity to detect differences in diary intrusions, despite choosing an analytical method that specifically deals with skewed distributions in count data. Importantly, the retrospective intrusion measure (IMS) did not show a highly skewed distribution, increasing confidence in the overall pattern of findings. It therefore seems unlikely that our null-finding is due to a change in materials or the number of intrusions provoked.

A second difference with Holmes et al (2009) was our active rather than passive control condition. The main reason for using an active condition was to ensure that participants remained engaged with the experiment, given the online setting. Active control conditions have been used in this type of research (e.g., logging activity; Iyadurai et al., 2018). In their own follow-up study, Holmes et al (2010) incorporated an active control condition (playing a computer quiz game) to counter the criticism that Tetris was merely distracting participants from focusing attention on the traumatic material. We deliberately used a perceptual vigilance task as a comparison to introduce a computer game-like element, without the complexity of Tetris. That is, we assumed that responding to a brief stimulus appearing randomly over relatively long intervals (30 sec on average) would only minimally tax visual working memory. However, one might wonder whether *any* activity in the visual domain might mitigate intrusions. Although we cannot be sure, there are reasons to believe that this might not be the case. To begin with, we replicated the finding that Tetris significantly reduces immediate intrusions for a trauma film in comparison to a control. The perceptual vigilance control group in our experiment reported more intrusions during the task, suggesting that it was not as

<sup>3</sup> Effect sizes (expressing the number of standard deviations difference between groups) are not reported because Welch *t*-tests do not assume a single pooled standard deviation. Therefore, describing the effect size as a single number is conceptually misleading.



**Figure 3. The Distributions, Medians and Inter-Quartile Ranges of Initial Intrusions in the Tetris and Perceptual Vigilance Control Conditions in Study 2**

resource-demanding as Tetris. Furthermore, Badawi et al. (2020) found that Tetris reduced long-term intrusions relative to a visuospatial working memory task (D-Corsi), which did not significantly differ from a resting control condition. The D-Corsi task would be more complex than our perceptual vigilance task. Relatedly, outside the Tetris literature, Meyer et al (2020; experiment 2) reported that imagining complex visual stimuli after reactivation in a trauma film paradigm did not mitigate intrusions. Thus, employing even more complex visual tasks than our perceptual vigilance task has rendered inconclusive effects on intrusions. Finally, the present Experiment 1's New Film condition may be considered equivalent to a passive control condition (i.e., no intervention) and the intrusions reported in this condition are numerically comparable to those of the perceptual vigilance task condition in Experiment 2. Thus, it is unlikely using a different control condition would have changed the conclusions.

A third difference was our online video conferencing (vs. in-person) study administration, due to the COVID-19 pandemic. We conducted several pilot studies prior to the Tetris study in Experiment 2. One of these was conducted in the lab (see <https://osf.io/9vtbz/>) and rendered comparable Image-based Diary intrusions ( $M = 3.2$ ,  $SD = 3.7$ , Median = 2) as in the online Tetris study reported here. In addition, Iyadurai et al. (2023) successfully used Tetris as an online intervention. Thus, the online format in itself was probably not problematic. Furthermore, the experiment was conducted according to a detailed protocol, with contact established with the experimenter at various points to ensure high compliance throughout all phases.

Nevertheless, we cannot definitively exclude the possibility that procedural differences contributed to incon-

clusive evidence of longer-term benefits of Tetris. Procedural deviations between studies are however inevitable (e.g., Nosek & Errington, 2020). The ability to reliably replicate an effect across varying contexts and settings is a hallmark of robust and meaningful findings that might have clinical utility. An effect that fails to replicate raises concerns about its generalizability and practical significance. Therefore, it is important to examine not only the specific experimental conditions but also the broader reliability and reproducibility of Tetris as an intervention for intrusive memories. Future replication studies may explore whether these changes represent important boundary conditions that limit the generalization of the effect of Tetris in analogue studies, and as such, refine theory (cf. Nosek & Errington, 2020). For example, concerns about our active perceptual vigilance control condition may inspire studies that systematically isolate the various components of Tetris (engagement/distraction; visual domain; mental rotation) to see which component affects intrusions most strongly. Likewise, the impact of testing format (in-person laboratory versus online sessions) may be examined.

Another type of explanation for our findings focuses on what it means when findings are inconclusive. Statistical nonsignificance might reflect either one of two things: a *failure to find* a true effect (i.e., our findings are false negatives), or the *absence* of a true effect (Lakens, 2022). Thus, our results may also mean that Tetris only has a temporary effect, not a long-lasting one. This interpretation casts doubt on one of the hypothesized mechanisms underlying the effect of Tetris on intrusions, i.e., disruption of the (re)consolidation of visuospatial imagery components of a target memory. This interpretation adds to uncertainties regarding another hypothesized mechanism, namely that

the disruption of the reconsolidation of traumatic memories depends on the modality of the interference (i.e., a visuospatial task to target visual mental imagery; e.g., Lau-Zhu et al., 2017). That is, research contrasting Tetris with an apparently equally demanding non-visual intervention renders mixed results; in one study Tetris was superior to a word game (Holmes et al., 2010), in another both were comparable (Hagenaars et al., 2017). Together, mixed findings regarding modality specificity and the present inconclusive evidence regarding the effect of Tetris challenge the imagery-competing hypothesis.

Regarding clinical implications, our results suggest that caution is warranted before assuming that the Tetris task in its current form can be considered an empirically proven intervention. Analogue and clinical studies differ in sample characteristics and procedural and methodological details, making direct comparisons challenging. Although it is difficult to directly compare our study to clinical ones, a prominent perspective in intrusion research supports the idea that psychopathology is dimensional rather than qualitatively different from ‘healthy’ (see Boals et al., 2020). According to this view, more controlled laboratory studies in healthy populations are needed to build robust clinical applications. Future research should investigate the mechanisms that might underlie the potential effect of Tetris on intrusions and examine whether these can be effectively translated to clinical settings. This will help determine if the mechanisms underlying the effect of Tetris on intrusions are applicable in clinical interventions.

In conclusion, our study replicated previous findings by demonstrating that playing Tetris after viewing a distressing film can reduce initial intrusions compared to a perceptual vigilance control condition. However, we did not observe the same beneficial effects on subsequent diary intrusions and retrospective film-related distress. So we only partially replicated previous Tetris findings. The mixed results may be attributed to methodological differences, such as the use of a new trauma film, an active control condition, and the online administration of the study. These factors underscore the importance of examining the broader reliability and reproducibility of Tetris as an intervention for intrusive memories. Although our findings provide valuable insights, further research is needed to fully understand whether Tetris affects intrusions in the long term and, if so, to explore the mechanisms underlying this effect. Future studies should investigate the specific components of Tetris that contribute to its impact and determine whether these findings can be reliably replicated across different contexts and populations. Continued exploration and validation of these findings will enhance our understanding of whether and how Tetris can be effectively utilized to support individuals experiencing intrusive memories.

## Contributions

Contributed to conception and design: IW, JK, CJA, DSfDs, TM, RDVN, DGP, TS, MKTT, KvS

Contributed to acquisition of data: IW, NC, DSfDs, AH, LJ, TM, DGP, IS, TRAW, KvS

Contributed to analysis and interpretation of data: IW, CJA

Drafted and/or revised the article: IW, JK, CJA, NC, DSfDs, AH, LJ, TM, RDVN, DGP, IS, TS, MKTT, TRAW, KvS

Approved the submitted version for publication: IW, JK, CJA, NC, DSfDs, AH, LJ, TM, RDVN, DGP, IS, TS, MKTT, TRAW, KvS

## Acknowledgments

We wish to thank Lucie Berger, Morticia Borocho, Ilse de Bree, Paschalia Chasalevri, Sasha Cox, Milou Douwes, Elisa van der Duim, Anna Heggenberger, Monika Lehnert, Lisann Maahs, Matty Karsten, Erika Miklasevics, Viktoria Perdikiogianni, Annabelle Quathamer, Lara Ruder, Nathalia Rustom, Katharina Staggenborg, Saskia Steinert, Nadine S. J. Stirling, Floor Stokkermans, G. Janina S. Tessmann, Asia Wahedi, Laura Wöhrle for their help in data collection and/or development of materials. We are indebted to The Tetris Company, Inc., for issuing a special version of Tetris for the purpose of this project.

## Funding Information

None of the authors received any funding for the research reported here.

## Competing Interests

All authors declare no conflicts of interest.

## Data Accessibility

Study 1: All study materials are publicly available (<https://osf.io/n42he/>). All anonymized data of eligible participants are publicly available (<https://osf.io/n42he/>). The privacy statement in the research information letter in our informed consent procedure (<https://osf.io/v5z6a>) prohibits sharing all data from non-eligible participants, and raw datafiles and descriptions in intrusion diaries from eligible participants. All analysis scripts are publicly available (<https://osf.io/n42he/>).

Study 2: All study materials are publicly available (<https://osf.io/3mgmv/>). All anonymized data of eligible participants are publicly available (<https://osf.io/3mgmv/>). The privacy statement in the research information letter in our informed consent procedure (see <https://osf.io/w5psr> for the template shared with all study sites) prohibits all sharing of raw data files as well as publicly sharing the descriptions in intrusion diaries from eligible participants. The deidentified intrusion diaries may be made available on request, depending on ethics regulations at each individual study site. All analysis scripts are publicly available (<https://osf.io/3mgmv/>).

## Ethics Statement

All sites obtained ethics approval from local Ethics Review Boards (Study 1: UG: PSY-2021-S-0229,

PSY-2122-S-0050; SU: Approval 20-27. Study 2: ARU: Submitted: January 12, 2025 PDT. Accepted: January 26, 2025  
ETH2223-3804; FU: HREC 4070, UG / RU: PDT. Published: March 11, 2025 PDT.  
PSY-2122-S-0077; SU: Approval 20-27; UT: TSB\_RP345).  
Thus, this research complies with the Declaration of Helsinki (2023).



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

## References

- Agren, T., Hoppe, J. M., Singh, L., Holmes, E. A., & Rosén, J. (2021). The neural basis of tetris gameplay: Implicating the role of visuospatial processing. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*. <https://doi.org/10.1007/s12144-021-02081-z>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5TM* (5th ed.). American Psychiatric Publishing, Inc.
- Arnaudova, I., & Hagenaars, M. A. (2017). Lights ... action: Comparison of trauma films for use in the trauma film paradigm. *Behaviour Research and Therapy*, 93, 67–77. <https://doi.org/10.1016/j.BRAT.2017.02.007>
- Arnou, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., Thase, M. E., Kocsis, J. H., & Keller, M. B. (2003). The 16-Item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, 54(5), 573–583. [https://doi.org/10.1016/S0006-3223\(02\)01866-8](https://doi.org/10.1016/S0006-3223(02)01866-8)
- Asselbergs, J., Sijbrandij, M., Hoogendoorn, E., Cuijpers, P., Olie, L., Oved, K., Merckies, J., Plooijer, T., Eltink, S., & Riper, H. (2018). Development and testing of TraumaGameplay: An iterative experimental approach using the trauma film paradigm. *European Journal of Psychotraumatology*, 9(1). <https://doi.org/10.1080/20008198.2018.1424447>
- Badawi, A., Berle, D., Rogers, K., & Steel, Z. (2020). Do Cognitive Tasks Reduce Intrusive-Memory Frequency After Exposure to Analogue Trauma? An Experimental Replication. *Clinical Psychological Science*, 8(3), 569–583. <https://doi.org/10.1177/2167702620906148>
- Boals, A., Contractor, A. A., & Blumenthal, H. (2020). The utility of college student samples in research on trauma and posttraumatic stress disorder: A critical review. *Journal of Anxiety Disorders*, 73, 102235. <https://doi.org/10.1016/j.janxdis.2020.102235>
- Bourne, C., Frascaillho, F., Roth, A. D., & Holmes, E. A. (2010). Is it mere distraction? Peri-traumatic verbal tasks can increase analogue flashbacks but reduce voluntary memory performance. *Journal of Behavior Therapy and Experimental Psychiatry*, 41(3), 316–324. <https://doi.org/10.1016/j.jbtep.2010.03.001>
- Brennen, T., Blix, I., Nissen, A., Holmes, E. A., Skumlien, M., & Solberg, Ø. (2021). Investigating the frequency of intrusive memories after 24 hours using a visuospatial interference intervention: A follow-up and extension. *European Journal of Psychotraumatology*, 12(1), 1953788. <https://doi.org/10.1080/20008198.2021.1953788>
- Brühl, A., Heinrichs, N., Bernstein, E. E., & McNally, R. J. (2019). Preventive efforts in the aftermath of analogue trauma: The effects of tetris and exercise on intrusive images. *Journal of Behavior Therapy and Experimental Psychiatry*. <https://doi.org/10.1016/j.jbtep.2019.01.004>
- Brys, G., Hubert, M., & Struyf, A. (2004). A Robust Measure of Skewness. *Journal of Computational and Graphical Statistics*, 13(4), 996–1017. <https://doi.org/10.1198/106186004X12632>
- Chaltiel, D., & Hajage, D. (2023). *crosstable: Crosstables for Descriptive Analyses* (0.6.2) [Computer software]. <https://cran.r-project.org/web/packages/crosstable/index.html>
- Cundill, B., & Alexander, N. D. (2015). Sample size calculations for skewed distributions. *BMC Medical Research Methodology*, 15(1), Article 1. <https://doi.org/10.1186/s12874-015-0023-0>
- Deforges, C., Noël, Y., Ayers, S., Holmes, E. A., Sandoz, V., Avignon, V., Desseauve, D., Bourdin, J., Epiney, M., & Horsch, A. (2024). There was no call for immediate implementation of “Tetris” in clinical practice: Response to the commentary by Halvorsen et al. (2024). *Molecular Psychiatry*. <https://doi.org/10.1038/s41380-024-02766-4>
- Deforges, C., Sandoz, V., Noël, Y., Avignon, V., Desseauve, D., Bourdin, J., Vial, Y., Ayers, S., Holmes, E. A., Epiney, M., & Horsch, A. (2023). Single-session visuospatial task procedure to prevent childbirth-related posttraumatic stress disorder: A multicentre double-blind randomised controlled trial. *Molecular Psychiatry*. <https://doi.org/10.1038/s41380-023-02275-w>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Gromer, D. (2023). *apa: Format Outputs of Statistical Tests According to APA Guidelines* (0.3.4) [Computer software]. <https://cran.r-project.org/web/packages/apa/index.html>
- Hagenaars, M. A., Holmes, E. A., Klaassen, F., & Elzinga, B. (2017). Tetris and Word games lead to fewer intrusive memories when applied several days after analogue trauma. *European Journal of Psychotraumatology*, 8(sup1), 1386959. <https://doi.org/10.1080/20008198.2017.1386959>
- Hagenaars, M. A., van Minnen, A., Holmes, E. A., Brewin, C. R., & Hoogduin, K. A. L. (2008). The effect of hypnotically induced somatoform dissociation on the development of intrusions after an aversive film. *Cognition and Emotion*, 22(5), 944–963. <https://doi.org/10.1080/02699930701575151>
- Halvorsen, J. Ø., Wessel, I., & Cristea, I. A. (2024). Premature call for implementation of Tetris in clinical practice: A commentary on Deforges et al. (2023). *Molecular Psychiatry*, 1–2. <https://doi.org/10.1038/s41380-024-02642-1>

- Hemi, A., Sopp, M. R., Perel, A., Holmes, E. A., & Levy-Gigi, E. (2023). Cognitive flexibility moderates the efficacy of a visuospatial intervention following exposure to analog trauma. *Journal of Behavior Therapy and Experimental Psychiatry*, 81, 101858. <https://doi.org/10.1016/j.jbtep.2023.101858>
- Holmes, E. A., Brewin, C. R., & Hennessy, R. G. (2004). Trauma films, information processing, and intrusive memory development. *Journal of Experimental Psychology: General*, 133(1), 3. <https://doi.org/10.1037/0096-3445.133.1.3>
- Holmes, E. A., James, E. L., Coode-Bate, T., & Deeprose, C. (2009). Can Playing the Computer Game “Tetris” Reduce the Build-Up of Flashbacks for Trauma? A Proposal from Cognitive Science. *PLoS ONE*, 4(1), e4153. <https://doi.org/10.1371/journal.pone.0004153>
- Holmes, E. A., James, E. L., Kilford, E. J., & Deeprose, C. (2010). Key Steps in Developing a Cognitive Vaccine against Traumatic Flashbacks: Visuospatial Tetris versus Verbal Pub Quiz. *PLOS ONE*, 5(11), e13706. <https://doi.org/10.1371/journal.pone.0013706>
- Horsch, A., Vial, Y., Favrod, C., Harari, M. M., Blackwell, S. E., Watson, P., Iyadurai, L., Bonsall, M. B., & Holmes, E. A. (2017). Reducing intrusive traumatic memories after emergency caesarean section: A proof-of-principle randomized controlled study. *Behaviour Research and Therapy*, 94, 36–47. <https://doi.org/10.1016/j.brat.2017.03.018>
- Hubert, M., & Vandervieren, E. (2008). An adjusted boxplot for skewed distributions. *Computational Statistics & Data Analysis*, 52(12), 5186–5201. <https://doi.org/10.1016/j.csda.2007.11.008>
- Iyadurai, L., Blackwell, S. E., Meiser-Stedman, R., Watson, P. C., Bonsall, M. B., Geddes, J. R., Nobre, A. C., & Holmes, E. A. (2018). Preventing intrusive memories after trauma via a brief intervention involving Tetris computer game play in the emergency department: A proof-of-concept randomized controlled trial. *Molecular Psychiatry*, 23(3), 674–682. <https://doi.org/10.1038/mp.2017.23>
- Iyadurai, L., Highfield, J., Kanstrup, M., Markham, A., Ramineni, V., Guo, B., Jaki, T., Kingslake, J., Goodwin, G. M., Summers, C., Bonsall, M. B., & Holmes, E. A. (2023). Reducing intrusive memories after trauma via an imagery-competing task intervention in COVID-19 intensive care staff: A randomised controlled trial. *Translational Psychiatry*, 13(1), Article 1. <https://doi.org/10.1038/s41398-023-02578-0>
- James, E. L., Bonsall, M. B., Hoppitt, L., Tunbridge, E. M., Geddes, J. R., Milton, A. L., & Holmes, E. A. (2015). Computer game play reduces intrusive memories of experimental trauma via reconsolidation-update mechanisms. *Psychological Science*, 26(8). <https://doi.org/10.1177/0956797615583071>
- Kehyayan, A., Thiel, J. P., Unterberg, K., Salja, V., Meyer-Wehrmann, S., Holmes, E. A., Matura, J.-M., Dieris-Hirche, J., Timmesfeld, N., Herpertz, S., Axmacher, N., & Kessler, H. (2024). The effect of a visuospatial interference intervention on posttraumatic intrusions: A cross-over randomized controlled trial. *European Journal of Psychotraumatology*, 15(1), 2331402. <https://doi.org/10.1080/20008066.2024.2331402>
- Kessler, H., Holmes, E. A., Blackwell, S. E., Schmidt, A.-C., Schweer, J. M., Bücker, A., Herpertz, S., Axmacher, N., & Kehyayan, A. (2018). Reducing intrusive memories of trauma using a visuospatial interference intervention with inpatients with posttraumatic stress disorder (PTSD). *Journal of Consulting and Clinical Psychology*, 86(12), 1076–1090. <https://doi.org/10.1037/ccp0000340>
- Kessler, H., Schmidt, A.-C., James, E. L., Blackwell, S. E., von Rauchhaupt, M., Harren, K., Kehyayan, A., Clark, I. A., Sauvage, M., Herpertz, S., Axmacher, N., & Holmes, E. A. (2020). Visuospatial computer game play after memory reminder delivered three days after a traumatic film reduces the number of intrusive memories of the experimental trauma. *Journal of Behavior Therapy and Experimental Psychiatry*, 67. <https://doi.org/10.1016/j.jbtep.2019.01.006>
- Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Benjet, C., Bromet, E. J., Cardoso, G., Degenhardt, L., De Girolamo, G., Dinolova, R. V., Ferry, F., Florescu, S., Gureje, O., Haro, M., Huang, Y., Karam, E. G., Kawakami, N., Lee, S., Lepine, J.-P., Levinson, D., ... Koenen, K. C. (2017). Trauma and PTSD in the WHO World Mental Health Surveys. *European Journal of Psychotraumatology*, 8, 5. <https://doi.org/10.1080/20008198.2017.1353383>
- Krans, J., Näring, G., & Becker, E. S. (2009). Count out your intrusions: Effects of verbal encoding on intrusive memories. *Memory*, 17(8), 809–815. <https://doi.org/10.1080/09658210903130780>
- Lakens, D. (2022). *Improving Your Statistical Inferences* (v1.0.0) [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.6409077>
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances on Methods and Practices in Psychological Science*, 1(2), 259–269. <https://doi.org/10.1177/2515245918770963>
- Lange, K., Kühn, S., & Filevich, E. (2015). “Just Another Tool for Online Studies” (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE*, 10(6), e0130834. <https://doi.org/10.1371/journal.pone.0130834>
- Lau-Zhu, A., Henson, R. N., & Holmes, E. A. (2019). Intrusive memories and voluntary memory of a trauma film: Differential effects of a cognitive interference task after encoding. *Journal of Experimental Psychology: General*, 148(12), 2154–2180. <https://doi.org/10.1037/xge0000598>

- Lau-Zhu, A., Holmes, E. A., Butterfield, S., & Holmes, J. (2017). Selective association between Tetris game play and visuospatial working memory: A preliminary investigation. *Applied Cognitive Psychology*, 31(4), 438–445. <https://doi.org/10.1002/acp.3339>
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335–343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)
- Maechler, M., Sn, P. R., Sn, C. C., Cov, V. T., Ruckstuhl, M. S.-B., & Comedian, M. A. di P. (2023). *robustbase: Basic Robust Statistics* (0.99-0) [Computer software]. <https://cran.r-project.org/web/packages/robustbase/index.html>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314–324. <https://doi.org/10.3758/s13428-011-0168-7>
- Meyer, T., Brewin, C. R., King, J. A., Nijmeijer, D., Woud, M. L., & Becker, E. S. (2020). Arresting visuospatial stimulation is insufficient to disrupt analogue traumatic intrusions. *PLOS ONE*, 15(2), e0228416. <https://doi.org/10.1371/journal.pone.0228416>
- Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's Renaissance. *Annual Review of Psychology*, 69(1), 511–534. <https://doi.org/10.1146/annurev-psych-122216-011836>
- Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLOS Biology*, 18(3), e3000691. <https://doi.org/10.1371/journal.pbio.3000691>
- Page, S., & Coxon, M. (2017). Preventing post-traumatic intrusions using virtual reality. *Annual Review of CyberTherapy and Telemedicine*, 15, 129–134.
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing* [Computer software]. Foundation for Statistical Computing.
- Ramineni, V., Millroth, P., Iyadurai, L., Jaki, T., Kingslake, J., Highfield, J., Summers, C., Bonsall, M. B., & Holmes, E. A. (2023). Treating intrusive memories after trauma in healthcare workers: A Bayesian adaptive randomised trial developing an imagery-competing task intervention. *Molecular Psychiatry*, 28(7), Article 7. <https://doi.org/10.1038/s41380-023-02062-7>
- Robinson, D., Hayes, A., Couch, S., Software, P., Patil, I., Chiu, D., Gomez, M., Demeshev, B., Menne, D., Nutter, B., Johnston, L., Bolker, B., Briatte, F., Arnold, J., Gabry, J., Selzer, L., Simpson, G., ... Reinhart, A. (2023). *broom: Convert Statistical Objects into Tidy Tibbles* (1.0.5) [Computer software]. <https://cran.r-project.org/web/packages/broom/index.html>
- Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., Markowitz, J. C., Ninan, P. T., Kornstein, S., Manber, R., Thase, M. E., Kocsis, J. H., & Keller, M. B. (2003). The 16-Item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, 54(5), 573–583. [https://doi.org/10.1016/S0006-3223\(02\)01866-8](https://doi.org/10.1016/S0006-3223(02)01866-8)
- Schumi, J., & Wittes, J. T. (2011). Through the looking glass: Understanding non-inferiority. *Trials*, 12(1), 106. <https://doi.org/10.1186/1745-6215-12-106>
- Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096–1104. <https://doi.org/10.3758/BRM.42.4.1096>
- van Schie, K., van Veen, S. C., & Hagenaars, M. A. (2019). The effects of dual-tasks on intrusive memories following analogue trauma. *Behaviour Research and Therapy*, 120, 103448. <https://doi.org/10.1016/j.brat.2019.103448>
- Wilkinson, R. T., & Houghton, D. (1982). Field test of arousal: A portable reaction timer with data storage. *Human Factors*, 24(4), 487–493. <https://doi.org/10.1177/001872088202400409>
- Willroth, E. C., & Atherton, O. E. (2023). *Best Laid Plans: A Guide to Reporting Preregistration Deviations*. PsyArXiv. <https://doi.org/10.31234/osf.io/dwx69>

## Supplementary Materials

### Peer Review Communication

Download: [https://collabra.scholasticahq.com/article/130791-evidence-that-tetris-reduces-immediate-but-not-subsequent-daily-intrusions-of-a-trauma-film-a-multilab-replication-study/attachment/268107.docx?auth\\_token=LdbrX3L\\_dEhJgY\\_-qsED](https://collabra.scholasticahq.com/article/130791-evidence-that-tetris-reduces-immediate-but-not-subsequent-daily-intrusions-of-a-trauma-film-a-multilab-replication-study/attachment/268107.docx?auth_token=LdbrX3L_dEhJgY_-qsED)

---