Bottenheft *et al. BMC Psychology* (2025) 13:755
https://doi.org/10.1186/s40359-025-03006-2

BMC Psychology

RESEARCH Open Access

Experimental validation of a performance curve using a Tetris game

Charelle Bottenheft[1,2*], Eric L. Groen[1] and Jan B.F. van Erp[2,3]

Abstract

Background Theoretical models on cognitive performance predict that performance breaks down at higher task load levels, because of limiting cognitive resources. Experimental validation of these models is scarce, and they often don't accommodate effects of invested effort and/or stressors. We experimentally measured a performance-curve as a function of increasing task load using a Tetris game. In addition, we investigated if performance breakdown occurs at different task load levels in the presence of a stressor. We hypothesized that an external stressor would shift the performance-curve to the left (i.e. lower task load levels) while maintaining the same form.

Methods Twenty-one participants completed one training day and two test days (data collection: 2023). During training, an individual break-off level on the Tetris game was determined with a staircase method. Subsequently, on the two test days participants played nine predetermined task load levels on and around their break-off level. On one test day they were exposed to heat load, and on the other test day to room temperatures. Task load levels and test room temperature were counterbalanced.

Results The results show that a sigmoid model is successful in fitting the performance decline pattern at high task load levels. Heat load did not seem to affect the parameters of the sigmoid curves.

Conclusions This study provides experimental evidence for a performance-curve as a function of increasing task load, resembling the right side of the classic inverted U-curve. Contrary to our expectation, we did not find evidence for a stressor-induced shift of the individual performance-curve, presumably because the increase in body core temperature (T_C) was insufficient to affect cognitive performance. Tetris seems a suitable tool to reproduce the performance-curve, and to perform research into factors that may affect the relation between mental effort and performance.

Keywords Task load, Mental effort, Cognition, Stressors, Tetris, Performance

*Correspondence:
Charelle Bottenheft
c.bottenheft@utwente.nl; charelle.bottenheft@tno.nl
[1]TNO, Human Performance, Kampweg 55, Soesterberg 3679DE, The Netherlands
[2]Human Media Interaction, Computer Science, University of Twente, Twente, The Netherlands
[3]TNO, Human Machine Teaming, Kampweg 55, Soesterberg 3679DE, The Netherlands

Background

Several theoretical models have been developed for predicting human cognitive task performance [1–6]. These models try to answer the question to what extent humans can operate and handle a large amount of information. Sweller et al. [6] assumed that, as task demands increase, the mental effort required to maintain performance increases linearly, where mental effort is defined as the amount of cognitive resources assigned to the task. This relationship between task demands and mental effort forms the basis for the workload model of Veltman et al. [5]. In this model the relationship between performance and workload is represented by an inverted U-curve, as was previously introduced by Yerkes & Dodson [3]. With this performance curve, schematically shown in panel A of Fig. 1, Veltman et al. [5] distinguished between four situations; (1) low workload with low performance due to boredom, (2) normal workload with optimal performance (i.e., 100%), (3) high workload, requiring additional mental effort to maintain optimal performance; and (4) overload, where performance becomes suboptimal because cognitive resources are no longer sufficient to cope with the task demands. As depicted by the cross in Fig. 1 (panel A), relative performance starts to decrease when the task demands become too high (the high workload situation). Relative performance in this figure refers to task performance in relation to the desired individual (maximum) performance level.

In the original model of Veltman et al. [5] it was assumed that task demands on the x-axis comprise both the intrinsic load of the task and the (indirect) impact of stressors. To better predict the impact of these factors (task load and stressors), we previously proposed a modification to the model by explicitly distinguishing between task load and the impact of stressors [7] as follows: Task Demand = Task Load + Stressor Load. We hypothesize that a "negative" stressor competes with cognitive resources that are required for the task, similar to the concept proposed in the Multiple Resource Theory by Wickens [4]. In other words, coping with a negative stressor requires effort in addition to the effort to achieve the desired task performance. This is illustrated by the leftward shift of the performance and effort curves in panel B of Fig. 1. Thus, for the same level of task load, the amount of mental effort available for the task is less in the presence of a stressor, and relative performance will start to decrease already at lower task load levels.

Most literature on the performance curve seems to focus on the conceptual framework, while experimental validation of the (parameters of the) curve is lacking [3, 5, 8]. Experimental studies on the effects of stressors on task performance looked at a specific task load level rather than examining performance across a range of task load levels [7, 9]. In the current study we used a cognitive task with ordinal task load or difficulty levels to examine the effect of an external stressor on the performance at the right side of the curve. For the task we chose a Tetris game [10, 11]. The first aim of this study is to reproduce the right-hand side of the performance curve experimentally as a function of increasing task load. The second aim is to validate the predicted stressor-induced shift of the performance curve by inducing heat load. As explained above, we expected the performance curve to shift leftwards in the presence of this stressor.

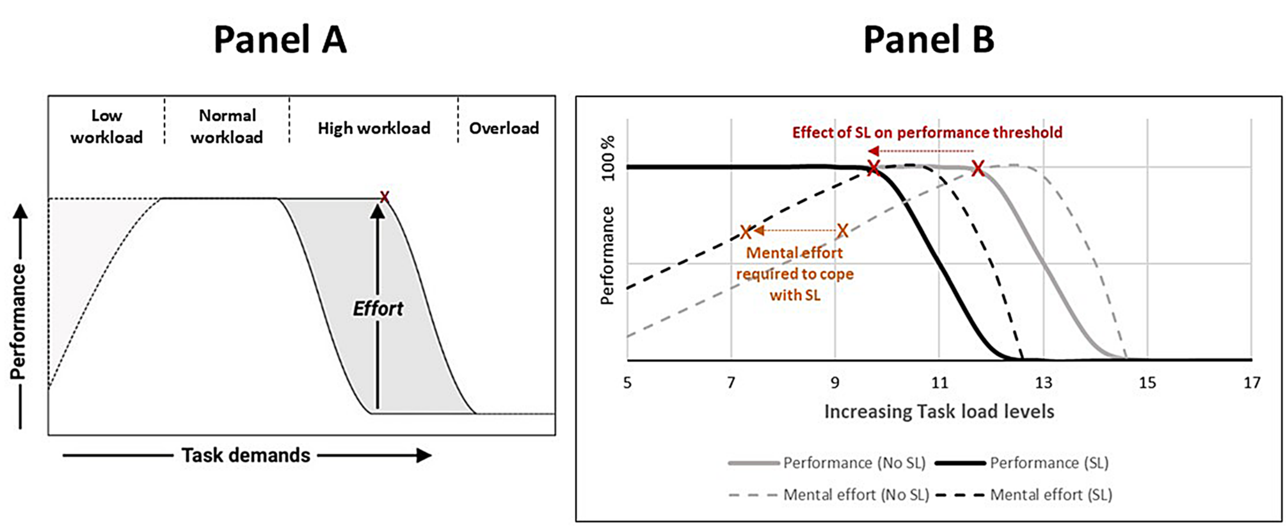

Fig. 1 Panel A: Hypothetical relation between task demands, performance, workload and mental effort (Veltman et al. [5]). Panel B: Expected performance curves for conditions without Stressor Load (No SL, grey) and with Stressor Load (SL, black) as function of increasing Task Load. Red crosses indicate the effect of SL on performance threshold and the orange crosses indicate the effect of mental effort required to cope with SL

Methods

Participants

A total of 21 participants completed this three-day study; Day 1 was a training day, Days 2 and 3 were experimental days. The participants were recruited through the organizational participant pool. Our study is conducted in accordance with the General Data Protection Regulation (GDPR) to ensure the privacy and protection of all participants' personal data. Approval for this study was granted by the Internal Review Board (reference number: 2022-050, year of data collection: 2023). All participants gave written informed consent. Exclusion criteria were: pregnancy, color blindness, cardiovascular disorders and psychiatric disorders. Ages ranged from 21 to 61 years old (M = 39.86, SD = 13.88) and nine of the participants were male.

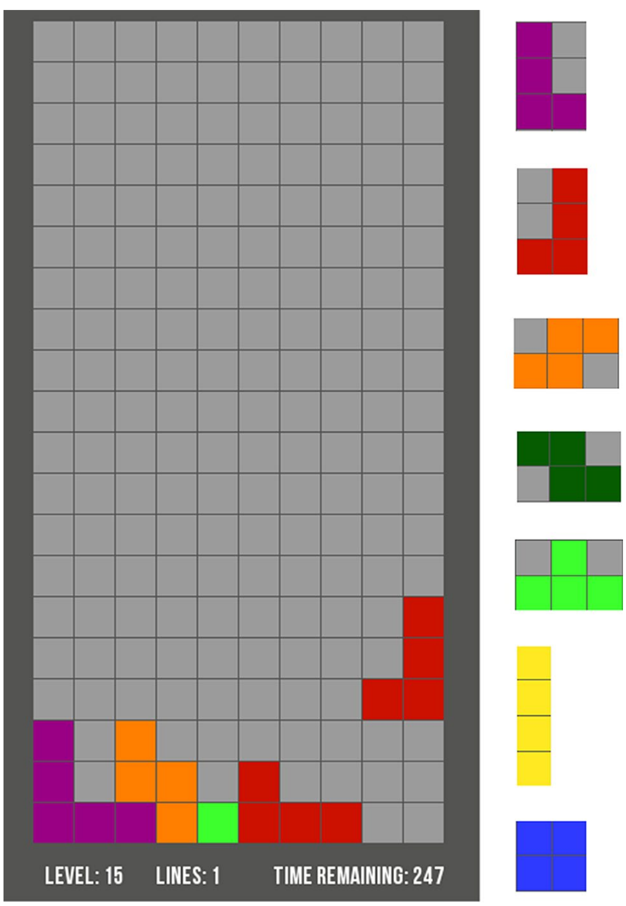

Fig. 2 Example of the Tetris game screen with 20 lines high and 10 blocks wide.
(source: programmed using a Python script (based on a script used by Schadll et al. [12])). *Note: On the right side, there is an overview of the seven different tetrominoes. The tetromino on the right side of the Tetris screen is falling. When it reaches the bottom in its current orientation, one row is successfully formed. From top to bottom, the tetrominoes are: L-tetromino (left side reflected), L-tetromino (right side reflected), T-tetromino (left side reflected), T-tetromino (right side reflected), skew tetromino, straight tetromino, and square.*

Our rationale for the sample size is based on a study by Chanel et al. (2011), who also used 20 participants. Similar studies have been conducted with sample sizes of 14 and 20 participants (Chanel et al., 2008; Patsis et al., 2013).

Materials

Tetris game and task load

Tetris is a computer game whereby blocks (called tetrominoes) fall from the top of the screen (Fig. 2). The falling speed depends on the level in the game (see next paragraph). By rotating the blocks and moving them horizontally using the arrow keys on the keyboard, the participant must try to form a horizontal row of blocks on the bottom of the screen. The Tetris screen in our study consisted of 20 lines and was 10 blocks wide. Each successfully formed row automatically disappeared and scored one point. Each level consisted of five minutes of playing. If the stack of blocks reached the top of the Tetris playing field within these five minutes, the game would be over. In that case, the participants had one chance to retry the same level. We used seven different tetrominoes: (1) straight tetromino, (2) square, (3) T-tetromino, (4) L-tetromino left side reflected (5) L-tetromino right side reflected (6) skew tetromino left side reflected and (7) skew tetromino right side reflected (see Fig. 2).

Chanel et al. [10] defined 25 playing levels with blocks falling with a speed of 1.070 lines per second at Level 1 and 27.027 lines per second at Level 25. The falling speed at the intermediate levels increased exponentially with the level. In our study we consider the falling speed directly related to the Task Load at (at least) an ordinal level. We only used the highest 20 levels that Chanel et al. [10] defined, to represent the right side of the performance curve in Fig. 1. For our study, we defined an intermediate level between each of the originally defined levels to create smaller increments in speed between levels. This corresponds to Levels 15.5 to 25 from Chanel et al. [10], which we conveniently renumbered from 1 to 20, implying that Level 1 in our study had a speed of 2.57 lines per second with blocks falling down one line every 0.39 s and Level 20 at a speed of 27 lines per second.

During the training day, participants practiced 5 min at Levels 1, 3, and 5 each, with the opportunity to repeat a level when the game was over before the five minutes of playing time. Additional guidance and instruction were given to ensure participants became proficient in the game. Then, from Level 5 the level was increased according to a staircase method to determine the individual break-off task load level; the level was increased by two levels when the previous level was completed but lowered by one level if the previous level was not achieved. The highest level achieved by a participant was used as their individual "break-off level". This is the same as the performance drop described in the introduction (see crosses in Fig. 1, Panels A and B). Participants played nine predetermined levels in a random order during both test days, to increase the likelihood that the data would fall within the steep part of the performance curve. The nine levels ranged from four levels below (normalized Task Load Levels −4, -3, -2, and −1) to four levels above (normalized Task Load Levels +1, +2, +3, and +4) their break-off level (normalized Task Load Level 0).

We computed two performance measures for each try of the Tetris game that have a finer-grained scores than the simple binary outcome of game-over or not after 5 min of playing time: (1) number of successfully formed rows divided by the speed of falling blocks, further referred to as Tetris score [11] and (2) elapsed time until game-over, that is, when uncleared lines reach the top of the playing field with a maximum of 300 s.

Thermal environment

We used heat load as an external stressor, which has been shown to effectively impact cognitive performance in earlier studies [7, 8, 13, 14]. The heat load was generated by radiant heat from an artificial sun (infrared halogen lamp, 13195X/98, 1000 W, 235 V REFL UNP) and by increasing the air temperature to 36 ˚C with a humidity of 40% in a climate chamber (Weiss Technik, Tiel, The Netherlands). Participants were sitting in the heated or not-heated climate chamber for 60 min while watching an episode of a nature documentary before playing the Tetris game. Table 1 shows the imposed air temperature, relative humidity and globe temperature for each condition while playing the Tetris game. The Wet Bulb Globe temperature (WBGT) outside increased to ~31 ˚C in the heat load condition.

We used several physiological and subjective measures for a manipulation check of heat load and to monitor the well-being of participants. Heart rate was monitored continuously using a chest strap and watch (OH1 and M430, Polar, Kemele, Finland). Mean skin temperature (T_{SK}) was measured at four positions according to Organization For Standardization ISO 9886 [15]: at the neck, scapula, left hand, and right shin. T_{SK} sensors (i-Button DS1922L, Maxim Integrated, USA) were fixed with Fixomull plasters, which are breathable and let sweat through. As an indication of body core temperature (T_C), gastrointestinal temperature was measured with an ingestible capsule (e-Celsius Performance, BodyCap, Caen, France, 17,7×8,9 mm) [16]. Participants ingested the capsule with water at least 60 min before playing the Tetris game. Stop criteria comprised of T_C exceeding 38.5 ˚C or failure of the monitoring system (T_C). In addition, participants indicated their perceived temperature sensation and thermal comfort during baseline and after each Tetris level [17]. Thermal comfort was rated using a five-point scale ranging from comfortable (0) to uncomfortable [4]. The temperature sensation was rated on a nine-point scale ranging from cold (-4) to very hot (+4).

Table 1 Mean air temperature (°C), relative humidity (%), Globe temperature (°C) and wet bulb Globe temperature (WBGT) outside (°C) while playing Tetris in both conditions

	Control	Heat load
Air temperature °C (dry bulb temperature)	21.96	37.54
Relative humidity (RH; %)	54.29	23.60
Globe temperature °C	21.91	42.69
WBGT outside °C	19.30	31.13

Subjective measures

Participants rated their perceived mental effort with the Rating Scale of Mental Effort (RSME) during baseline and after each Tetris level [18]. The RSME scale ranges from 0 to 150, with higher values reflecting higher workload. It has nine descriptors along the axis, for example, 'not effortful' at a value of 2 and 'rather effortful' at a value of 58. Furthermore, participants indicated the perceived difficulty for each Task Load level on a scale from 0 to 100. This difficulty rating is used as a manipulation check to verify increasing difficulty with Task Load levels. Various scales (five-point, nine-point, 0-150, and 0-100) are used for subjective measures because these are validated rating scales from different fields of expertise.

Design

The design was a within-subjects design with two independent variables: Task Load (*nine levels from −4 to +4 of the participant's break-off Task Load level*) and Stressor Load (*control condition* vs. *heat load condition*).

Participants came three times to the test location: once for a training day and twice for test days, with at least seven days between each visit. On one test day they were exposed to heat load (stressor condition) and on the other test day to room temperatures (control condition). The order of exposure was counterbalanced: ten participants were first exposed to heat load, and the other participants were first exposed to room temperature.

Data analysis

For statistical data analysis IBM SPSS (Statistical Package for the Social Sciences) version 26.0 was used [19]. All statistical tests were performed at a significance level of $\alpha = 0.05$. Tetris and physiological data were preprocessed using MATLAB R2019b [20]. To compare Tetris performance between participants, the nine predetermined Tetris Task Load levels were normalized according to the break-off level each participant achieved during the training session. This means that the break-off level was set to 0, and each level below or above this break-off level was adjusted accordingly, ranging from −4 to +4. In case of a retry, due to game over, performance scores were obtained by taking the mean of both attempts. Individual Tetris performance curves for both conditions were obtained by fitting an inverted sigmoid. The reason for using an inverted sigmoid is that the drop in performance with increasing task demands closely resembles the right side of an inverted U-shape. The curve captures the plateau and subsequent gradual decline in performance, providing a more precise mathematical representation of the observed performance drop. MATLAB Curve Fitting Toolbox was used with the following Eq. [21]:

$$Tetris\,performance\;curve\;\;(y) = top + \frac{bottom - top}{1 + e^{\frac{x - V50}{slope}}}$$

With top = upper plateau, $bottom$ = lower plateau, $V50$ = Task Load level at midpoint (this corresponded roughly with the individual break-off level), $slope$ = gradient at midpoint and x = relative Task Load level. In this equation, *V*50 and *slope* are the free parameters. For elapsed time the upper plateau was set to 300 s. and the lower plateau was set to 0 s. For Tetris score the upper plateau was set to 7.6 as this was the highest score achieved in our dataset and the lower plateau was set to 0. *V*50 had a minimum of -4, while the maximum was undefined because it is likely that the Task Load level at midpoint could be higher. Negative values for the *slope* are normal as the model uses an inverted slope. The same equation was used for the subjective RSME data, with an upper plateau set to 150 and a lower plateau set to 0.

To check whether the inverted sigmoid is an appropriate fit, it was compared to a simple linear fit with the following equation:

$$Tetris\,performance\;curve\;\;(y) = a + bx$$

With a = intercept (value of y when x = -4), b = slope, x = relative Task Load level.

R^2 coefficients for both the sigmoid and linear fit were calculated to determine how well the predictions approximate the real data points (goodness-of-fit). If R^2 values were negative, indicating a bad fit, data of individual parameters (*V*50 and slope) were excluded. R^2 coefficients of both fits were compared with a Wilcoxon Signed-Ranks test (first research aim).

A Shapiro-Wilk test was performed to test for normality. It showed that the distribution of R^2 coefficients for elapsed time, Tetris score and RSME ratings of both fits deviated significantly from normality, respectively for the sigmoid fits $W = 0.830$, $p < .001$, $W = 0.929$, $p = .021$, $W = 0.908$, $p = .005$, and respectively for the linear fits $W = 0.937$, $p = .037$, $W = 0.925$, $p = .016$, $W = 0.922$, $p < .012$. Thermal sensation and discomfort also deviated significantly from normality, respectively $W = 0.886$, $p < .001$ and $W = 0.847$, $p < .001$. Therefore, non-parametric tests were used because ranks and medians are more robust to non-normal distributions.

To determine the effect of Stressor Load (independent variable) on Tetris performance (dependent variable), individual parameters (*V*50 or intercept, and slope) obtained from the Tetris performance curve equations for each condition were compared with Wilcoxon Signed-Ranks tests (second research aim). The same was done for RSME ratings. Difficulty ratings in both conditions and during the various Task Load levels were visualized with boxplots to check for increasing Task Load.

Mean and maximum values of physiological data were calculated for the following subsequent time periods: (1) while watching the documentary and (2) when playing Tetris. Also, a weighted T_{SK} was calculated for these time periods with the following formula: $0.28{*}T_{neck} + 0.28{*}T_{scapula} + 0.16{*}T_{hand} + 0.28{*}T_{shin}$ (ISO 9886, 2004). These data were collected to measure the effectiveness of the heat load stressor to affect the participants' physiological strain.

Procedure

The training day started with instructions and signing an informed consent. After practicing the Tetris game, participants watched an episode of a nature documentary for one hour, followed by Tetris sessions to determine their break-off level.

To arrive hydrated at the test location during both test days, the participants were instructed to follow a hydration protocol at home. This protocol included [1] no exercise 24 h before arrival [2], drink ~ 5–7 ml x kg body weight water 4 h before arrival, and [3] drink another ~ 3–5 ml x kg body weight water when urine is still dark at arrival at the test location. The third step in this protocol was only applied if a participant indicated that he or she had not followed the first two steps of the protocol. It serves as a subjective prompt to drink more water. At the start of each test day, participants ingested the capsule with some water for core temperature monitoring. Participants were not allowed to drink or eat after ingestion of the pill for the duration of the experiment. After ingesting, sensors for heart rate and T_{SK} were attached. Next, participants were sitting in the heated or not heated climate chamber while watching an episode of a nature documentary before playing the Tetris game. Every 10 min the core temperature, the perceived temperature sensation and the thermal comfort were checked by the experiment leader. After this pre-heating phase, participants started to play the Tetris game at different levels in a random order.

Results

Data exclusion

The R^2 (inverted sigmoid model) for Tetris score was negative for two participants. After visual inspection of their performance curves, we excluded all Tetris data for these participants. A possible reason could be that these participants had not mastered the task and therefore their performance varied too much across the Task Load levels. Due to a technical error, Tetris data in the control condition was missing for one additional participant, resulting in a total of 18 complete Tetris datasets. Concerning the RSME, data for two participants in the control condition was removed due to negative ratings that occurred when they accidentally clicked below the scale on the computer screen. Besides this, the R^2 (inverted sigmoid model) for the RSME was negative for one participant in the control condition and for one participant in the heat load condition. Both were excluded from RSME analysis using the sigmoid model, leaving 19 participants. However, they were included in the RSME analysis using the linear model. The reason for not excluding these two participants in the linear model is that their values seemed appropriate based on visual inspection for a linear fit. RSME data in the control condition was missing for one additional participant, due to a technical error, resulting in a total of 18 complete RSME datasets for the sigmoid model and 20 for the linear model.

First research aim: curve fitting

The right-hand side of Fig. 3 shows examples of successful sigmoid fits for Tetris performance measures (elapsed time and Tetris score) and mental effort ratings (RSME). For most of the participants ($N = 13$) the sigmoid functions produced a high goodness-of-fit, indicated by a high R^2. For comparison, the left-hand side of Fig. 3 shows some examples of poor sigmoid fits (low R^2). Overall, the median R^2 of the sigmoid fit for both elapsed time and Tetris score is above 0.50 and for RSME 0.68, indicating good fits which explains more than 50% of the data (see Table 2). Extremely low (< 0.10) R^2 coefficients of the sigmoid fit for Tetris performance measures were found for a total of five participants.

See Table 2 for a comparison between the median R^2 of the sigmoid and linear fit for both performance measures as well as the mental effort rating. For both performance measures the R^2 of the sigmoid fit is significantly higher compared to the R^2 of the linear fit. Therefore, we chose the inverted sigmoid fit over the linear fit to determine the effect of Stressor Load on the Tetris performance curves. For the RSME rating no significant difference between R^2 for the sigmoid and linear fit was found (see Fig. 3A, B, and C for an example). For reasons of consistency, and because of the trend in the data, we also choose a sigmoid fit for the RSME ratings.

Second research aim: effect of stressor load on performance

Table 3 shows the results of the statistical analysis for differences between the heat-load and control condition for the two performance variables (elapsed time and Tetris score) and RSME. There were no effects on the fit parameters (*V*50 and *slope*) between conditions. With effect sizes between 0.03 and 0.22, an alpha level of 0.05, a total sample size of 18, divided into two groups, and two measurements, the power analysis reveals an overall low power between 0.06 and 0.42 (G*Power, version 3.1.9.7).

Manipulation checks

Difficulty ratings

Boxplots of difficulty ratings for each normalized Task Load level are shown in Fig. 4. Difficulty ratings increased from a median of 27 (range = 80) at Task Load level − 4 to a median of 75 (range = 73) at Task Load level + 4. After visual inspection, this confirms the increasing difficulty with Task Load levels.

Thermal strain

Table 4 shows descriptives for each physiological and subjective measure regarding thermal strain. Heart rate, T_{SK} and T_C were significantly affected by heat load, with higher mean heart rate, T_{SK} and T_C in the heat load condition compared to the control condition. Also, thermal sensation and thermal discomfort were significantly affected by heat load, with higher maximum ratings in the heat load condition compared to the control condition.

Discussion

The first research aim of this study was to collect experimental evidence for a performance curve as function of increasing task load using the Tetris game. In general, we found a performance decline at the right-hand side as predicted by the inverted U-curve shown in Fig. 1, panel B. For the majority of participants this performance decline could be fitted by a simple inverted sigmoid model. For five participants (out of 18), the fit to a sigmoid curve was not successful with low goodness-of-fit values ($R^2 < 0.10$). An explanation is that for these participants, their break-off level, which is the highest level achieved, was not determined correctly during the training day. If a participant did not reach an optimal performance level during training, their break-off level would fall outside the range of task load levels used during the test days. We can think of two reasons why this would occur. Firstly, there could be a training effect so that the participants performed better during the test days. Indeed, these participants showed no performance decline with increasing task load levels. To deal with this in the future, longer training will be required. Secondly,

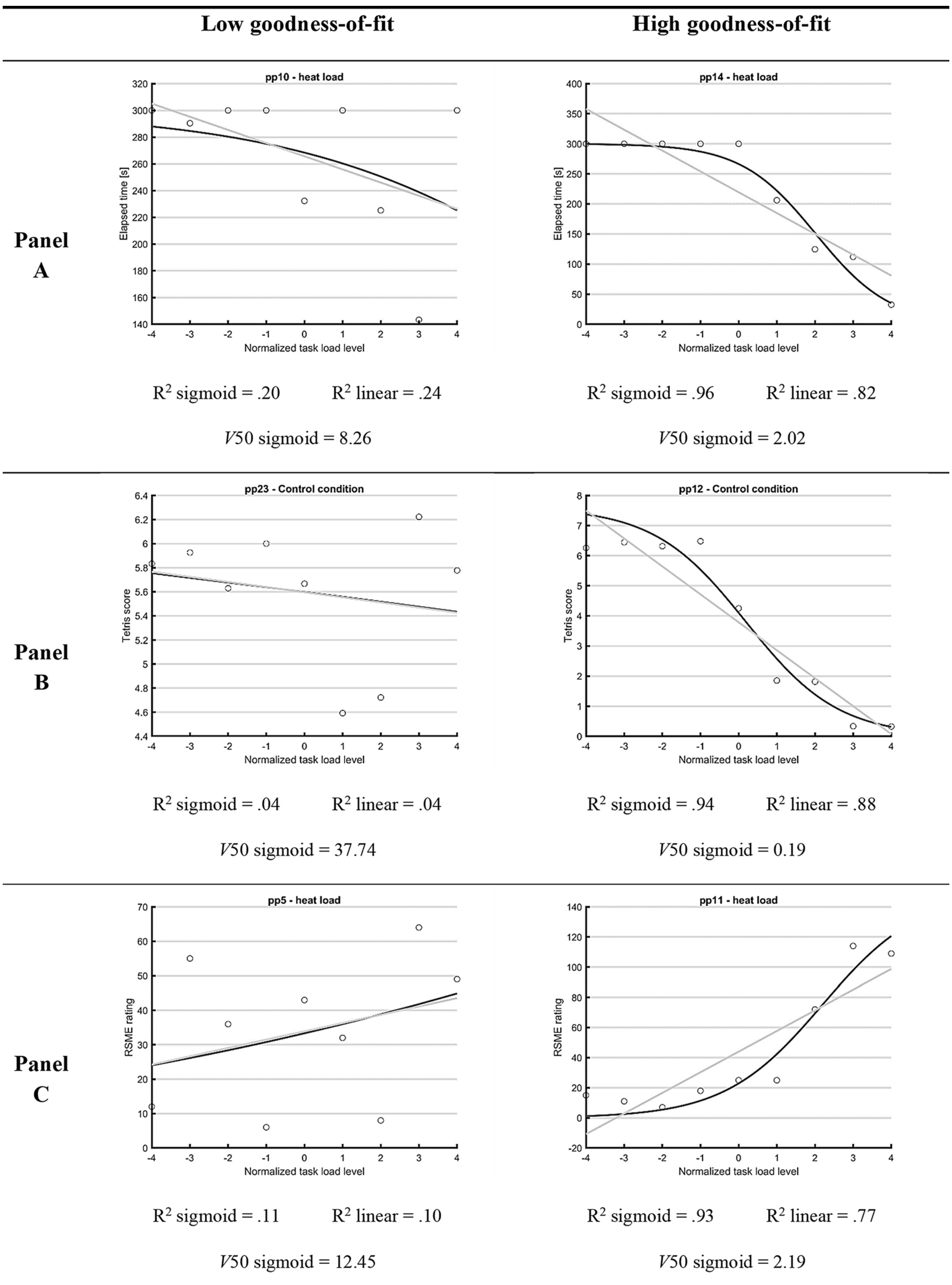

Fig. 3 Examples of individual sigmoid curves (dark line) and linear curves (light line) with a low R^2 (on the left side) and high R^2 (on the right side) for Tetris performance measures (Elapsed time = panel A and Tetris score = panel B) and mental effort rating (RSME; panel C). Examples are randomly chosen from both the heat load and control conditions

day-to-day variation in performance may have played a role. Although the sessions were at the same time of day, evaluating performance on a wider task load range (than the −4 to +4 used in this study) might be a solution. For example, maintain nine levels, but increase or decrease the task load level with two instead of one step. The disadvantage hereof may be that the lower task loads may show performance decline because of boredom as they may be below the cutoff value for the left side of the curve. This would reduce the goodness of fit.

Table 2 Comparison between sigmoid and linear fits for both performance measures (elapsed time, Tetris score) and mental effort rating (RMSE) with medians, range (Mdn ± range) and test statistics (Z and p-values) of the Wilcoxon signed-rank tests

		Sigmoid fit	**Linear fit**	**Test statistics**
Tetris performance	Elapsed time	0.79 ± 0.96	0.57 ± 0.85	Z = -4.466, *p* <.001**
	Tetris score	0.52 ± 0.93	0.46 ± 0.88	Z = -4.312, *p* <.001**
Mental effort	RSME	0.72 ± 0.91	0.68 ± 0.95	Z = -1.727, *p* = .084

Note. ** *p* <.001

Table 3 Parameters of the inverted sigmoid fits for both performance measures (elapsed time, Tetris score) and mental effort rating (RMSE) for both conditions with medians, range (Mdn ± range) and test statistics (Z and p-values) of the Wilcoxon Signed-Ranks tests

		Control condition	**Heat load condition**	**Test statistics**
Elapsed time	Task Load level at midpoint (*V50*)	3.29 ± 34.74	3.37 ± 18.20	Z = − 0.065, *p* = .948
	Slope	-1.55 ± 9.71	-1.47 ± 6.88	Z = − 0.936, *p* = .349
Tetris score	Task Load level at midpoint (*V50*)	1.06 ± 51.85	1.53 ± 33.81	Z = − 0.109, *p* = .913
	Slope	-2.77 ± 53.90	-3.07 ± 26.07	Z = − 0.544, *p* = .586
RSME	Task Load level at midpoint (*V50*)	2.40 ± 12.52	3.53 ± 16.52	Z = -1.293, *p* = .196
	Slope	-4.88 ± 11.64	-6.61 ± 11.82	Z = -1.448, *p* = .148

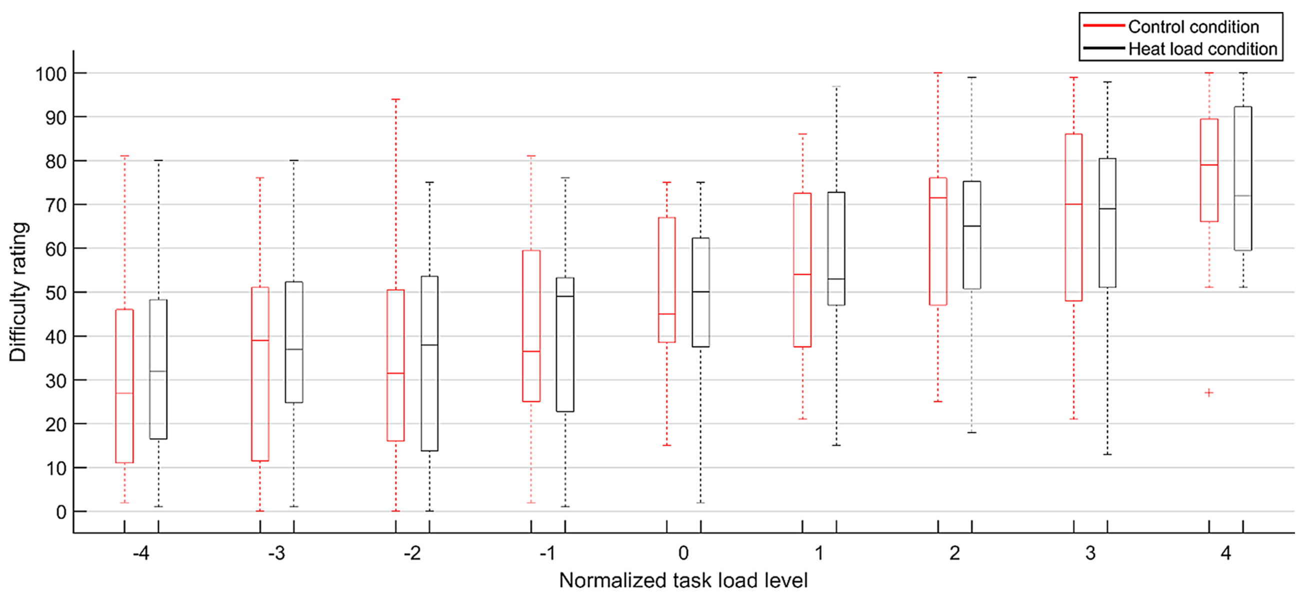

Fig. 4 Boxplots of difficulty ratings for each normalized Task Load level in both conditions. Red boxplots represent the median and range values for the control condition and black boxplots represent the median and range values for the heat load condition

Table 4 Mean ± SD for physiological measures and mdn ± range for subjective measures in the control and heat load condition

	Control condition	**Heat load condition**	**Test statistics**
Heart rate	73.36 ± 13.34	85.63 ± 12.04	t(19) = -3.07, *p* = .006*
T_{SK}	36.17 ± 1.14	40.45 ± 0.66	t(18) = -18.44, *p* <.001**
T_C	36.98 ± 0.22	37.38 ± 0.13	t(20) = -7.96, *p* <.001**
Thermal sensation	0.00 ± 2.00	3.00 ± 2.00	Z = -4.069, *p* <.001**
Thermal discomfort	0.00 ± 2.00	3.00 ± 4.00	Z = -3.893, *p* <.001**

* *p* <.05, ** *p* <.001

The mental effort ratings reflected a similar relationship between mental effort and increased task load. The results reflect the idea that, as task load increases, mental effort rises gradually and eventually plateaus when the task becomes too difficult and it is no longer possible to invest additional mental effort to keep performance at the same level. This suggests that when one cannot or does not want to invest extra effort their task performance will decline.

The second research aim of this study was to investigate whether an external stressor, that is heat load, would result in a shift of the performance curve to lower task load levels. The results do not seem to support this hypothesis. Heat load did not affect the mean parameters of the sigmoid performance curves. According to the physiological and subjective data the heat load did induce thermal strain and led to increased subjective heat load and discomfort. Although T_C was significantly higher in the heat load condition compared to the control condition and the heat load used in the current study proved to significantly affect multitasking performance and vigilance in a previous study [7], T_C remained below the range of 38.5–39.0 °C which is considered as a threshold for cognitive performance decline by Schmit et al. [22]. Thus, the heat load may not have been intense enough to elevate T_C significantly, so that the participants' performance remained unaffected. The literature suggests that heat load may have contrasting effects on cognitive performance due to various factors, like sensitivity of the cognitive task being used, intensity and duration of exposure [8, 14]. It is possible that the Tetris game requires cognitive functions, such as psychomotor skills, hand-eye coordination and manual dexterity that may be less sensitive to heat load than multitasking and vigilance we used previously and showed to be sensitive to the same heat load. Another exploratory factor can be the low achieved power for assessing the impact of heat load on Tetris performance. This highlights the need for a larger sample size and a priori sample size estimation to ensure adequate statistical power for detecting meaningful differences between conditions.

One of the strengths of this study is the inclusion of a staircase method for determining the break-off level for each participant. This approach allows for a more precise and individualized assessment of performance thresholds and reduces the likelihood of ceiling or floor effects. However, we observed a clear difference in proficiency among the participants because we used a wide range of difficulty. Some individuals struggled to reach level 5, while others achieved significantly higher levels. This distinction in proficiency suggests considerable variability in participant performance. Participants who are closer to the lower levels may not show significant decreases in performance, while participants who are closer to the higher levels might be more susceptible to the effects of stressors. However, this observation is speculative and would require formal statistical evaluation to confirm, which is not feasible given the current sample size. Another limiting factor that could have reinforced this distinction in proficiency is age and experience with playing Tetris or other computer games. Age can play a significant role in cognitive performance, with younger participants often having quicker reaction times and may adapt more easily to increasing difficulty levels [23]. Similarly, participants with prior experience in playing Tetris or other computer games may have an advantage over those who are less familiar with the game. In the context of this study, this means that age-and experience related differences could contribute to the variability in proficiency observed among participants. Future studies should therefore aim to better select individuals who are more similar in their optimal performance levels. By ensuring a more homogeneous group in terms of proficiency, we can improve the reliability and validity of the results.

Conclusions

This study provides experimental evidence for a performance curve which has often been conceptualized by an inverted U-shape. The results show that for high task load levels the decline of cognitive performance can be described by a sigmoid model. The Tetris game seems a suitable method to systematically vary Task Load. Contrary to our expectation, we did not find a shift of the performance curve in the presence of heat load. Presumably, the heat load was insufficient to induce the required increase of T_C to affect cognitive performance. The performance curve should be examined with other or more intense stressors to validate the impact of stressor on performance and effort.

Acknowledgements
The authors want to thank our team, in particular IS, DM and JB, for their contributions to the setup of the experiment and data acquisition.

Author contributions
CB, EG and JvE conceived and designed the study. CB performed the experiment. CB performed data analysis with input from all authors. All authors wrote the first drafts of the manuscript. All authors thoroughly reviewed and revised the manuscript and approved the submitted version.

Funding
This work was funded by the Dutch Ministry of Defense (V1917 '5thGenStressors' program).

Data availability
Data, methods, and materials are not fully disclosed and accessible because participants did not consent to their data being shared publicly.

Declarations

Ethics statement
The study involving human participants was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. All procedures were reviewed and approved by the TNO Internal Review Board (reference number: 2022-050). The participants provided their written informed consent to participate in this study.

Consent for publication
Not Applicable.

Competing interests
The authors declare no competing interests.

Received: 15 October 2024 / Accepted: 10 June 2025
Published online: 09 July 2025

References

1. Hancock PA, Vasmatzidis I. Effects of heat stress on cognitive performance: the current state of knowledge. Int J Hyperth. 2003. https://doi.org/10.1080/0265673021000054630.
2. Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev. 1956;63(2):81.
3. Yerkes RM, Dodson JD. The relationship of strength of stimulus to rapidity of habit formation. J Comp Neurol Psychol. 1908;18:459–82.
4. Wickens CD. Multiple resources and mental workload. Hum Factors. 2008. https://doi.org/10.1518/001872008X288394.
5. Veltman JA, Jansen C, Hockey GRJ, Gaillard AWK, Burov O. Differentiation of mental effort measures: consequences for adaptive automation. NATO science series I: life and behavioural sciences. Amsterdam: ISO; 2003. pp. 249–59.
6. Sweller J, Van Merrienboer JJG, Paas FGWC. Cognitive architecture and instructional design. Educational Psychol Rev. 1998. https://doi.org/10.1023/A:1022193728205.
7. Bottenheft C, Groen EL, Mol D, Valk PJL, Houben MMJ, Kingma BRM, van Erp JBF. Effects of heat load and hypobaric hypoxia on cognitive performance: a combined stressor approach. Ergonomics. 2023. https://doi.org/10.1080/00140139.2023.2190062.
8. Hancock PA, Warm JS. A dynamic model of stress and sustained attention. Hum Perform Extreme Environments: J Soc Hum Perform Extreme Environ. 2003. https://doi.org/10.7771/2327-2937.1024.
9. Martin K, Flood A, Pyne DB, Periard JD, Keegan R, Rattray B. The impact of cognitive, physical, and psychological stressors on subsequent cognitive performance. Hum Factors. 2024;66(1):71–87.
10. Chanel G, Rebetez C, Bétrancourt M, Pun T. Emotion assessment from physiological signals for adaptation of game difficulty. IEEE Trans Syst Man Cybernetics-Part A: Syst Hum. 2011;41(6):1052–63.
11. Patsis G, Sahli H, Verhelst W, De Troyer O. Evaluation of attention levels in a tetris game using a brain computer interface. In User Modeling, Adaptation, and Personalization: 21th International Conference, UMAP 2013, Rome, Italy, June 10–14, 2013 Proceedings 21. 2013. pp. 127–138).
12. Schadll S, Rodriguez-Raecke R, Heim L, Freiherr J. Playing Tetris lets you rate odors as less intense. Front Psychol. 2021;12:657188.
13. Gaoua N, Grantham J, Racinais S, El Massioui F. Sensory displeasure reduces complex cognitive performance in the heat. J Environ Psychol. 2012. https://doi.org/10.1016/j.jenvp.2012.01.002.
14. Martin K, McLeod E, Périard J, Rattray B, Keegan R, Pyne DB. The impact of environmental stress on cognitive performance: A systematic review. Hum Factors. 2019. https://doi.org/10.1177/0018720819839817.
15. Organization for Standardization. 2004b. Ergonomics— Evaluation of Thermal Strain by Physiological measurements. (ISO Standard No. 9886:2004). https://www.iso.org/standard/34110.html
16. Bongers CCWG, Daanen HAM, Bogerd CP, Hopman MTE, Eijsvogel TMH, Validity. Reliability, and inertia of four different temperature capsule systems. Med Sci Sports Exerc. 2018. https://doi.org/10.1249/MSS.0000000000001403.
17. Organization for Standardization. 2019. Ergonomics— Ergonomics of the physical environment — Subjective judgement scales for assessing physical environments. (ISO Standard No. 10551:2019). https://www.iso.org/standard/67186.html
18. Zijlstra FRH, Van Doorn L. The construction of a scale to measure subjective effort. Volume 43. Delft: Delft University of Technology; 1985. pp. 124–39.
19. IBM Corp. Released 2019. IBM SPSS statistics for windows, version 26.0. Armonk, NY: IBM Corp.
20. The MathWorks Inc. MATLAB version: 9.7.0. (R2019b), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com
21. The MathWorks Inc. Curve Fitting Toolbox version: 3.5.10. (R2019b), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com
22. Schmit C, Hausswirth C, Le Meur Y, Duffield R. Cognitive functioning and heat strain: performance responses and protective strategies. Sports Med. 2017;47:1289–302.
23. Wolkorte R, Kamphuis J, Zijdewind I. Increased reaction times and reduced response Preparation already starts at middle age. Front Aging Neurosci. 2014;6:79.

Publisher's note