Current Psychology (2026) 45:1067
https://doi.org/10.1007/s12144-026-09620-6

Validating an acute stress induction method for Spanish native speakers

Sergio Navas-León[1] · **Jon Andoni Duñabeitia**[1]

Received: 29 August 2025 / Accepted: 5 June 2026
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2026

Abstract

This study presents the first Spanish-language adaptation of an online stress induction protocol. The objective of this study was to adapt an existing protocol and provide initial evidence of the validity of this adaptation based on subjective responses. The protocol was based on time pressure and performance feedback to elicit stress. A total of 60 Spanish native speakers residing in Spain were randomly assigned to either a stress ($n=30$) or control condition ($n=30$). The experimental group completed general knowledge and arithmetic tasks under time pressure with feedback, while the control group completed easier versions of the same tasks without time constraints and feedback. Participants rated their state anxiety on a visual analogue scale and their emotional valence on a Likert-type scale. Participants in the experimental group reported significantly higher anxiety and lower emotional valence, indicating a more negative affective experience. These findings support the initial evidence of validity of a culturally adapted acute stress induction protocol for native Spanish speakers. The protocol may highlight its potential for scalable applications in both clinical and educational contexts. Different possible limitations linked to the relatively small sample size are also discussed.

Keywords Anxiety · Cultural adaptation · Emotions · Performance feedback · Spanish · Time pressure

Abbreviations

TSST	Trier Social Stress Test
VAS	Visual Analogue Scale

Acute stress induction paradigms are widely used to examine stress-related changes in psychological and behavioural functioning. However, the most established laboratory procedure (i.e., the Trier Social Stress Test, hereafter TSST) may pose practical constraints that limit scalability, particularly in online research. In this regard, a recent online stress-induction task has shown promising subjective effects (see Almazrouei et al., 2023), yet its cross-cultural validity remains largely untested. Extending its application to native Spanish speakers can help evaluate its cross-cultural validity and provide researchers with adapted tools for laboratory-based experiments requiring stress induction.

To fill the gap, the objective of the present study was twofold: first, to adapt an online stress induction protocol based on the TSST proposed by Almazrouei et al. (2023) for Spanish speakers living in Spain at the time of testing; and second, to provide initial evidence of its validity in terms of subjective and behavioural responses. Extending this protocol beyond English-speaking samples is important to evaluate its cross-cultural applicability, given evidence that stress reactivity may vary across cultural contexts (Miller & Kirschbaum, 2019; Santos-Carrasco & De La Casa, 2025).

Notably, in the present study general knowledge questions were selected from extensive, Spanish-normed item pools with empirically derived difficulty information, supporting both the ecological validity and the controlled manipulation of task difficulty (Buades-Sitjar et al., 2021, 2022). Thus, unlike Almazrouei et al. (2023), we relied on an extensive, validated item pool, which allowed a clearer manipulation of cognitive demand: general knowledge and arithmetic questions were systematically varied in difficulty

✉ Sergio Navas-León
snavas@nebrija.es

Jon Andoni Duñabeitia
jdunabeitia@nebrija.es

[1] Centro de Investigación Nebrija en Cognición (CINC), Department of Education, Universidad Nebrija, Madrid 28248, Spain

Published online: 05 June 2026

across conditions, with easy items in the control group and more challenging items in the experimental group. As methodological differences across TSST implementations can substantially influence acute stress responses, such a well-matched active control is essential to separate effects of psychosocial stress from those due to procedural variability (Goodman et al., 2017). Furthermore, to capture a broader range of stress responses, we assessed emotional valence using a Likert-type format. This addresses a limitation highlighted by Koncz et al. (2025), who, in a Hungarian-speaking sample, noted the absence of valence measures and recommended their inclusion, in line with evidence linking acute stress and emotional valence (Langer et al., 2023). Incorporating valence ratings thus enables a more comprehensive assessment of stress reactivity (Langer et al., 2023).

To achieve this objective, this study specifically aimed to: (1) assess between-group differences in self-reported state anxiety (VAS); and (2) assess between-group differences in perceived emotional valence. In this regard, we hypothesised that:

H1: Consistent with prior validations of comparable online acute stress tasks reporting heightened subjective state anxiety responses (e.g., Almazrouei et al., 2023; Koncz et al., 2025), participants in the stress condition would show higher state anxiety ratings compared with those in the control condition.

H2: Given evidence linking acute stress to more negative affective experience and altered emotional responding (Epel et al., 2018; Raio & Phelps, 2015), participants in the stress condition would show more negative emotional valence than in the control condition.

Literature review

Stress is an unavoidable element of daily life and plays a key role in shaping adaptive responses (Epel et al., 2018; Tuan & Loan, 2025). Acute stress, understood as a short-lived reaction to a recent and transient stressor, can overwhelm regulatory capacity and trigger a cascade of physiological, psychological and behavioural changes (Lu et al., 2021; Skoluda et al., 2015). Physiologically, it activates stress-response systems that increase heart rate, blood pressure and cortisol, thereby boosting energy and immune preparedness (Skoluda et al., 2015). At the psychological level, acute stress has been shown to increase perceived state anxiety and a sense of uncontrollability (Xin et al., 2017), and auditory stressors can evoke discrete emotions such as fear, sadness and anger (Epel et al., 2018; Raio & Phelps, 2015). From a cognitive perspective, acute stress can disrupt executive functions such as working memory and inhibition, while enhancing the consolidation of declarative memory but impairing working memory performance (Shields, 2020; Shields et al., 2016, 2017; Starcke et al., 2016). Importantly, stress responses vary markedly between individuals, with some showing maladaptive patterns and others greater resilience, influenced by factors such as sex, genetic predisposition, personality and mindset (Xin et al., 2017).

In this regard, a substantial body of work has examined the effects of acute laboratory-induced stress in humans (Dickerson & Kemeny, 2004). This research typically distinguishes between physical stressors (e.g., exposure to low temperatures) and psychological stressors (e.g., arithmetic under time pressure; Shields, 2020). A meta-analysis of 208 laboratory studies showed that the strongest stress responses occur when social-evaluative threat (being negatively judged by others) is combined with uncontrollability (inability to influence or avoid negative outcomes; Dickerson & Kemeny, 2004). The close link between stress and numerous medical conditions, poorer quality of life, and reduced life expectancy (Epel et al., 2018; Garfin et al., 2018) underscores the need to understand how people respond to specific stressors at the psychological level. Identifying the stressors that most strongly trigger emotional and cognitive responses may help clarify the pathways through which stress contributes to adverse health outcomes (Dickerson & Kemeny, 2004; O'Connor et al., 2021).

Considering this evidence, Kirschbaum et al. (1993) developed the TSST, a now well-established psychosocial stress induction procedure. In brief, the TSST requires participants to deliver an impromptu speech and perform a challenging mental arithmetic task in front of a neutral, non-reinforcing evaluation panel, who correct errors in a formal, unemotional manner. This combination of public performance, repeated correction, and lack of positive feedback reliably creates social-evaluative threat and uncontrollability, eliciting a robust stress response. It is evident that, over the years, a multitude of variations of the TSST have been developed (see Linares et al., 2020, for a review on the topic). These variations include, but are not limited to, the number of judges or the starting value in the arithmetic task. A systematic review confirmed that the TSST reliably activates the body's main stress response systems, producing consistent cortisol increases (Linares et al., 2020). Indeed, it is widely accepted that the TSST is the gold standard for inducing acute stress in laboratory settings (Almazrouei et al., 2023; Heyers et al., 2025; Linares et al., 2020). However, despite its robustness, the TSST presents notable practical limitations, most notably, its dependence on multiple trained experimenters and a highly controlled setting, which restrict its scalability and broader applicability (Almazrouei et al., 2023; Koncz et al., 2025).

In this context, and in line with the shift in behavioural science towards online research, digital platforms now allow rapid recruitment, broader sampling and the implementation of complex designs (Gagné & Franzen, 2023; Reips, 2021; Rodd, 2024; Tomczak et al., 2023). Building on the core TSST principles of perceived social-evaluative threat and uncontrollability, Almazrouei et al. (2023) recently developed and validated an online stress induction method. In their study, participants answered general knowledge questions and arithmetic problems under time pressure, while receiving evaluative feedback designed to elicit social-evaluative threat. In the control condition, the same tasks were completed without time limits or evaluative feedback. This manipulation proved effective: participants in the experimental group reported higher levels of state anxiety on a visual analogue scale than those in the control group.

While the original study by Almazrouei et al. (2023) demonstrated the efficacy of this stress induction method in English-speaking participants, extending its application to a sample of native Spanish speakers can help evaluate its cross-cultural validity. Multilevel perspectives on stress measurement emphasise that the sociocultural context may be a plausible source of variability (Crosswell & Lockwood, 2020; Epel et al., 2018). More specifically, the appraisal of psychosocial stressors, particularly social-evaluative threat and uncontrollability, may be influenced by relatively stable cultural norms and values. In this regard, Schwartz's (2006) theory of cultural value orientations conceptualises population-level values as slow-changing social norms and maps them onto three broad dimensions: autonomy vs. embeddedness, egalitarianism vs. hierarchy, and harmony vs. mastery.

Building on this framework, Miller and Kirschbaum's (2019) cross-national meta-analysis of TSST studies provided formal evidence of between-country heterogeneity in cortisol reactivity, with European studies showing larger average responses than North American studies ($d = 0.73$ vs. $d = 0.45$) and substantial variation across countries/regions; descriptively, Spain (along with some Latin regions) fell among those with comparatively lower TSST responses. They further reported that country-level cultural values, especially the harmony–mastery dimension, were reliably associated with TSST responses (higher harmony/lower mastery predicting larger responses, and mastery predicting attenuated responses), noting that Spain clusters more towards harmony values whereas English-speaking countries cluster more towards mastery values. In complementary analyses, they also observed that higher national prevalence of internalising disorders was associated with blunted TSST cortisol responses and discussed Spain among the countries combining lower cortisol responses with higher major depression prevalence relative to some other European countries, while emphasising that these links are probabilistic and potentially confounded.

Beyond these macro-cultural accounts, several more proximal mechanisms may contribute to cross-cultural variability in TSST paradigms. Cultural "display rules" may influence how participants communicate subjective state anxiety and emotional valence (Matsumoto et al., 2008), and language may shape emotion conceptualisation and labelling, such that translation and linguistic context can subtly affect how stress-related feelings are categorised and reported (Lindquist et al., 2015).

Method

The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments after obtaining the approval from the institutional review board of the Universidad Nebrija (code: UNNE-2025-048). At the beginning of the experiment, written informed consent was obtained from all participants, who received a small monetary compensation a few days after participation. The protocol was implemented in Gorilla Experiment Builder (Anwyl-Irvine et al., 2020) to support replication, transparency, and future adaptation. The material is available for consultation via the following link: https://app.gorilla.sc/openmaterials/1086018.

Methodological design and rationale

This study employed a between-subjects experimental design (stress vs. control) to provide an initial validation of the Spanish-language adaptation of an online acute stress induction protocol (Almazrouei et al., 2023; Koncz et al., 2025). A between-subjects approach was considered the most appropriate design for this purpose because repeated exposure to the same manipulation within participants can introduce carryover and learning effects (Charness et al., 2012; Greenwald, 1976) as well as demand characteristics that may alter responding (Orne, 1962). Moreover, repeated exposure to comparable psychosocial stressors has been shown to produce habituation in stress responding, which could attenuate effects across administrations (James et al., 2023). By allocating participants to one condition only, we sought to preserve the effectiveness of the manipulation and obtain a clearer estimate of condition-related differences in subjective responses, while also maintaining consistency with prior validation studies to facilitate direct comparison of findings across samples and cultural contexts (Almazrouei et al., 2023; Koncz et al., 2025).

Accordingly, between-group comparisons in subjective responses provide a direct test of whether the adapted protocol elicits acute stress-related changes consistent with the

hypotheses outlined above (Dickerson & Kemeny, 2004). From a theoretical point of view, the protocol was designed to operationalise core theoretical mechanisms of psychosocial stress induction, namely uncontrollability and social-evaluative threat, which are central triggers of acute stress responses (see Dickerson & Kemeny, 2004). Specifically, time pressure was implemented to increase perceived uncontrollability by constraining participants' ability to regulate pace and performance, thereby heightening the sense that outcomes could not be fully controlled (see Szalma et al., 2008, for a meta-analysis on the topic). In parallel, performance feedback and comparative evaluation (e.g., "**WRONG**" messages and performance-related comments) were used to elicit social-evaluative threat by signalling potential negative judgement and unfavourable comparison with others, mirroring evaluative components central to TSST-inspired procedures (Dickerson & Kemeny, 2004).

Participants

The sample consisted of 60 participants aged between 18 and 40 years (M=29.68, SD=5.17). Regarding sex, 63.33% identified as male (n=38) and 36.67% as female (n=22). In terms of employment status or educational engagement, 43.33% of participants reported being in full-time education (n=26), while 26.67% were unemployed but actively seeking work (n=16). Additionally, 11.67% were working part-time (n=7), 6.67% indicated they were about to start a new job within the next month (n=4), and another 6.67% provided invalid or missing responses (n=4). Finally, 3.33% reported other employment or educational situations (n=2). Participants were randomly assigned to either the stress group (n=30; 50%) or the control group (n=30; 50%). No participants withdrew from the study. Data were collected via the Prolific platform (http://prolific.com) in a single session. Recruitment was restricted to individuals currently living in Spain at the time of participation. All participants were native Spanish speakers. The sample was largely homogeneous in terms of origin, as the vast majority were not only living in Spain and native Spanish speakers, but also born in Spain and holding Spanish nationality. Only 3 out of the 60 participants were born in Latin American countries, even though they were living in Spain at the time of the study[1].

Recruitment was carried out in February 2025, drawing participants from a pool of individuals registered for online research. The study included participants with normal or corrected-to-normal vision who were 18 years of age or older. Individuals were excluded if they reported a history of significant brain injury, neurological disorders, or ocular/visual conditions, as well as any history of mental illness. Those currently using psychotropic medication, such as antidepressants, consuming recreational drugs, or lacking sufficient proficiency in Spanish were also not eligible to participate.

The target sample size (N=60; 30 participants per group) was determined a priori based on feasibility considerations and to allow detection of at least medium-to-large effects. To characterise the sensitivity of this design, we conducted an a priori sensitivity analysis in G*Power 3.1 (Heinrich Heine University Düsseldorf, Germany; http://www.gpower.hhu.de/en.html). For a two-tailed independent-samples t test with α=0.05, power $(1-\beta)=0.80$, and $n_1=n_2=30$, the minimum detectable effect size is Cohen's d=0.74. Thus, the study was adequately powered to detect medium-to-large between-group differences.

Instruments

Socio-demographic information Participants provided information regarding their age, sex-at-birth, education level and employment status.

Acute stress induction task The task consisted of four blocks, each containing six general knowledge and six arithmetic questions. The general knowledge questions were selected based on their difficulty level from previously published and normed materials (Buades-Sitjar et al., 2021, 2022). The experimental version involved time pressure, error-based feedback, and block-level performance feedback. The control version had no time constraints or feedback. Task accuracy, which was not a primary outcome of the study, was defined as the number of correct responses (i.e., total correct answers) across the task. Item difficulty was manipulated across conditions as in Almazrouei et al. (2023): the control group received easy items (e.g., general knowledge question: "*What is another name for the Mona Lisa?*"; arithmetic question: "*35 − 20*"), whereas the experimental group completed more challenging questions, in line with the intended variation in cognitive demand (e.g., general knowledge question: "*What is the highest peak in Europe?*"; arithmetic question: "*25 ÷ 5 × 3 ÷ 5*").

State anxiety State anxiety was operationally defined as participants' momentary, self-reported intensity of subjective anxious arousal (feeling stressed/anxious) during the task. A visual analogue scale (VAS) ranging from 0 ("*Not stressed at all*") to 100 ("*Extremely stressed*") was used. The initial slider position was randomised to reduce bias. VAS responses can be used as an index of state anxiety and show adequate psychometric properties (reliability, validity, and sensitivity to change) (Abend et al., 2014; Lesage, Berjot, & Deschamps, 2012).

Emotional Valence Emotional valence was operationally defined as the affective comfort–discomfort experienced

[1] Three participants were born in Latin American countries; however, no anomalies or outliers were detected in their responses. Analyses excluding these participants yielded the same pattern of results.

during the task. Participants rated their emotional experience using a Likert-type scale ranging from 1 ("*Very uncomfortable*") to 7 ("*Very comfortable*"). Valence ratings are commonly used in TSST-based paradigms to capture the affective component of the stress experience, with TSST-related conditions eliciting more negative valence than controls (Freund et al., 2023).

Procedure

Utilising Gorilla Experiment Builder (Anwyl-Irvine et al., 2020), we adapted the procedure developed by Almazrouei et al. (2023) for use in the Spanish language. Prior to participation, all subjects provided written consent in accordance with the stipulated guidelines. Subsequently, demographic data were also collected, encompassing age, sex, and education level. Participants were then randomly assigned via Prolific (https://prolific.com) to either the experimental or the control condition. Subsequently, the participants completed a form to be completed and a set of written instructions detailing the task at hand. In accordance with the original protocol, participants were assigned the task of providing responses to a series of general knowledge and arithmetic questions, selected at random. Participants were blind to the specific aim of the study, namely the induction of stress. The consent form and task instructions provided only general procedural information.

The experimental group was shown a warning indicating that their performance would be monitored. During the task, incorrect responses triggered a red "**WRONG**" message, correct answers elicited a neutral grey "**OK**," and timeouts resulted in a red "**TIME OUT**!" message. At the conclusion of each block, participants received either a neutral or a negative feedback message depending on their performance. Specifically, those who scored three or fewer correct answers were shown a negative message that compared their performance to that of other participants. This heightened the social-evaluative threat component of stress. Conversely, participants who correctly answered four or more questions were presented with a neutral message that did not make any reference to comparative performance. Participants in the control group completed the same number and type of questions but without time limits or feedback.

Each participant completed four blocks of trials. Within each block, participants completed six trials of general knowledge, followed immediately by six arithmetic trials. This resulted in a fixed order presentation within each block, with a total of 48 trials completed by each participant. To ensure data quality and to circumvent both inadvertent and mechanised responses, a series of quality control measures were implemented. These comprised a commitment statement and three attention checks administered at various points throughout the experiment. Even though participants who failed more than two attention checks were to be excluded, all participants met this criterion and were retained in the analysis. Upon completion of the task, participants completed a VAS to assess perceived stress. The following question was posed to the participants: "*Looking back, how stressed did you feel while answering the arithmetic and general knowledge questions?*" Participants rated their stress on a scale ranging from 0 ("*Not stressed at all*") to 100 ("*Extremely stressed*"). To minimise anchoring effects, the initial position of the slider was randomised for each participant. Subsequently, participants were invited to rate the emotional valence of the task using a Likert-type scale in response to the question: "*How did this task make you feel*?" using a scale of response options ranging from 1 ("*Very uncomfortable*") to 7 ("*Very comfortable*"). The mean completion time for the task was 5.76 min (SD = 0.96) for the control group and 8.15 min (SD = 1.34) for the experimental group, calculated from session timestamps recorded by Gorilla Experiment Builder.

Data analysis

Independent samples t-tests were conducted to assess between-group differences in three continuous variables: accuracy, state anxiety and emotional valence. These analyses were performed using JASP (v. 0.19.3) (Love et al., 2019). Assumptions of normality were examined both visually (i.e., histograms and Q-Q plots) and using the Shapiro-Wilk test. Potential outliers were defined a priori as observations lying more than 1.5 interquartile ranges above the third quartile or below the first quartile in boxplots; their exclusion did not change the overall pattern of results. All tests were two-tailed with an alpha level set at 0.05. For all between-group comparisons, Cohen's d was computed as (control – experimental), such that negative values indicate higher scores in the experimental group and positive values indicate higher scores in the control group. All analyses were planned a priori.

Results

Performance on acute stress task

Accuracy, which was not a primary outcome of the study, was analysed at the participant level as the total number of correct responses as a sanity check of the correct functioning of the task. An independent samples *t*-test using Welch's correction revealed a significant difference between groups, $t(37.60) = 20.51$, $p < .001$. Participants in the control condition

(M=45.53, SD=2.50) performed substantially better than those in the experimental condition (M=19.73, SD=6.42). The assumption of homogeneity of variances was violated, $F(1, 58)=20.46$, $p < .001$, and therefore Welch's t-test is reported. Normality assumptions were not violated, $W=0.96$, $p = .058$. The effect size was large (Cohen's $d=5.29$, 95% CI [3.99, 6.58]), although this estimate should be interpreted cautiously, given the difference in item difficulty between conditions. A non-parametric Mann–Whitney test confirmed the robustness of the effect, $W=900$, $p < .001$.

State anxiety and emotional valence

State anxiety levels were significantly higher in the experimental group (M=77.50, SD=19.64) than in the control group (M=41.53, SD=29.11). Normality assumptions were not violated, Shapiro–Wilk $W=0.976$, $p = .291$, although homogeneity of variances was not met, Brown–Forsythe $F(1, 58)=8.46$, $p = .005$. Accordingly, Welch's t-test was used, revealing a statistically significant difference, $t(50.87) = -5.61$, $p < .001$, Cohen's $d = -1.45$, 95% CI [–2.02, −0.86]. This result was corroborated by a non-parametric Mann–Whitney U test, $U=139.50$, $p < .001$.

Similarly, emotional valence ratings were significantly lower in the experimental group (M=2.57, SD=1.33) than in the control group (M=4.40, SD=1.71). The Shapiro–Wilk test indicated deviations from normality, $W=0.936$, $p = .004$, whereas the assumption of equal variances was not violated, Brown–Forsythe $F(1, 58)=1.94$, $p = .169$. A Student's t-test confirmed a significant difference, $t(58)=4.63$, $p < .001$, Cohen's $d=1.20$, 95% CI [0.64, 1.74]. This result was also corroborated by a non-parametric Mann–Whitney U test, $U=721.50$, $p < .001$ (Table 1).

As shown in Fig. 1, participants exposed to the experimental condition experienced significantly greater state anxiety and negative emotional valence than those in the control group, with large and consistent effect sizes across both measures.

Table 1 Descriptive statistics for state anxiety and emotional valence by group

Variable	Group	n	M	SD
State anxiety (0–100)	Control	30	41.53	29.11
	Experimental	30	77.50	19.64
Emotional valence (1–7)	Control	30	4.40	1.71
	Experimental	30	2.57	1.33

State anxiety was measured on a 0–100 scale; emotional valence was measured on a 1–7 scale. M=mean; SD=standard deviation

Discussion

The objective of this study was to adapt a recently developed TSST-inspired stress-induction method for Spanish native speakers and to provide initial evidence of its validity in terms of subjective responses, as previously described for English-speaking individuals (Almazrouei et al., 2023). We hypothesised that participants in the stress condition would report higher state anxiety and more negative emotional valence than those in the control condition. Both hypotheses were supported.

The findings are consistent with the intended stress manipulation. Participants in the stress condition reported

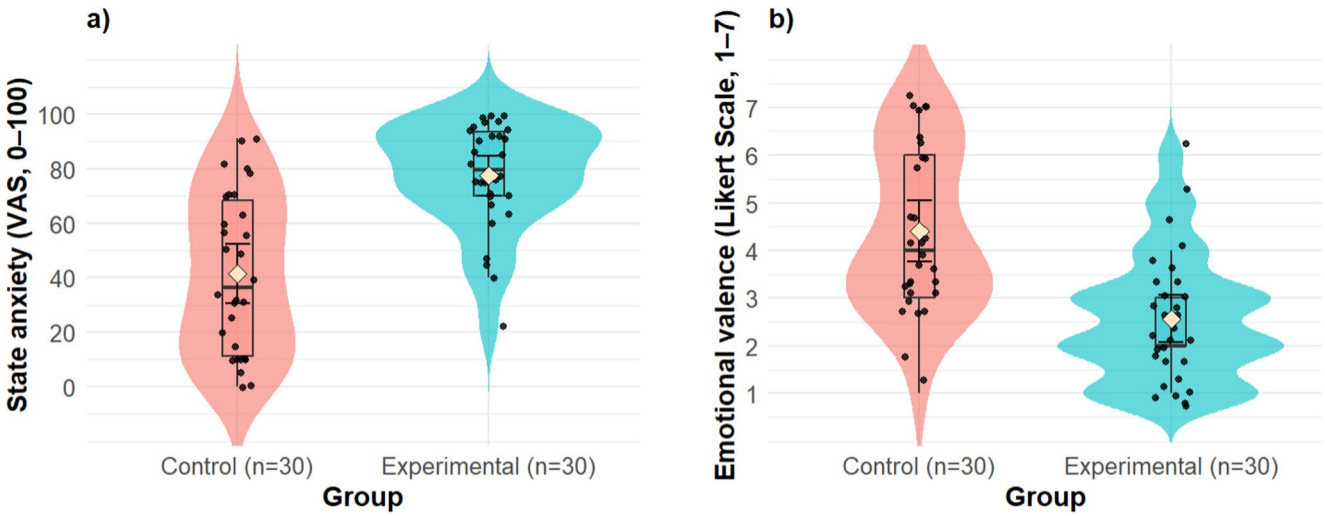


Fig. 1 Group differences in anxiety and emotional valence. *Note*. (**a**) Distribution of anxiety scores by group. (**b**) Distribution of emotional valence scores by group. Higher valence values reflect a more positive or comfortable emotional response to the task. Each dot represents an individual participant

significantly higher state anxiety and more negative emotional valence than those in the control condition, replicating and extending previous work using related online protocols (Almazrouei et al., 2023; Koncz et al., 2025). These effects are theoretically compatible with a manipulation designed to elicit perceived uncontrollability and social-evaluative threat, two core mechanisms highlighted in psychosocial stress models (Dickerson & Kemeny, 2004).

An alternative explanation is that the observed pattern may reflect, at least in part, differences in item difficulty between conditions, as the control group received easy items while the stress group completed more challenging questions. This design choice should be interpreted in light of the original protocol, in which lower-accuracy items were assigned to the stress condition to increase the likelihood of failure and negative feedback, thereby strengthening two central components of the manipulation: perceived uncontrollability and social-evaluative threat (Almazrouei et al., 2023). Nevertheless, the results should be interpreted with caution, as they cannot be attributed exclusively to stress-related disruption and may also reflect the increased cognitive load of the experimental group. However, the stress manipulation remains a plausible account, given that the performance decrement co-occurred with marked differences in state anxiety and emotional valence, consistent with prior validations using time pressure and evaluative feedback (Almazrouei et al., 2023; Koncz et al., 2025), and because the protocol was designed to elicit uncontrollability and social-evaluative threat, which are key stress triggers (Dickerson & Kemeny, 2004).

In line with this interpretation, several converging findings emerged. Firstly, the subjective responses were modulated by the experimental condition. Participants in the stress group reported significantly higher state anxiety levels ($d = -1.45$), replicating previous research using this protocol (i.e., Almazrouei et al., 2023; Koncz et al., 2025). Secondly, and crucially, participants in the experimental group also reported significantly more negative emotional valence ($d = 1.20$), indicating that the stress condition not only induced anxious arousal but also a marked reduction in perceived comfort. The stress induction procedure was conducted without the presence of an experimenter, yet it yielded significant subjective outcomes. This finding is consistent with the results of recent studies (see Almazrouei et al., 2023; Koncz et al., 2025), and it underscores the efficacy of this protocol in complementing conventional laboratory-based protocols in the domain of stress research (Koncz et al., 2025; Rodd, 2024; Tomczak et al., 2023).

From an applied perspective, our findings may suggest that this protocol could elicit acute stress in an online setting. Although we did not assess therapeutic or academic outcomes, future work may explore whether similar tasks could potentially be embedded within digital emotion-regulation interventions (e.g., ACT/DBT-informed programmes; Blackledge & Hayes, 2001; Harvey et al., 2019; Wind et al., 2020) or could be examined in studies of test state anxiety and evaluative contexts in education (Criollo et al., 2024; Marques et al., 2025; Mavridis & Tsiatsos, 2017; von der Embse et al., 2018). Methodologically, the study contributes culturally adapted materials for Spanish native speakers, including general knowledge stimuli drawn from a large validated normative database, thereby enhancing ecological and psychometric adequacy (Buades-Sitjar et al., 2021). It also extends prior work by incorporating emotional valence as a complementary subjective indicator of the task experience, in line with recommendations to adopt a more multidimensional assessment of stress-related responses (Almazrouei et al., 2023; Koncz et al., 2025; Langer et al., 2023). Finally, the protocol was implemented in Gorilla Experiment Builder, supporting accessibility, transparency, replication, and future adaptation within an open-science framework (Anwyl-Irvine et al., 2020; Patall et al., 2021).

Notwithstanding the strengths, it is imperative to acknowledge the limitations of the present study. Firstly, the sample size was modest ($n = 60$). Whilst the current study has revealed sufficient power to detect medium to large effects, future research should recruit larger and more diverse samples using stratified random sampling. This is in view of the evidence that acute stress responses are moderated by demographic group, being more pronounced in females and older individuals (Kudielka & Kirschbaum, 2005). Secondly, the generalisability of the findings may be influenced by sample characteristics, including the restricted age range, and the fact that all native Spanish speaking participants were living in Spain. Future research should examine whether similar effects are observed in more diverse populations, other Spanish-speaking countries and other variants of the Spanish language. In addition, its applicability to other age groups, settings, or clinical populations remains to be established, as these factors may influence responses to the protocol. Thirdly, the study relied exclusively on self-report measures, without incorporating physiological indicators (i.e. salivary cortisol, heart rate variability, or skin conductance). This limitation impedes the capacity to obtain a more comprehensive picture of acute stress reactivity. Nonetheless, Rodrigo-Herrero et al. (2026) recently used this same protocol in a Spanish sample and incorporated physiological measures based on vocal and acoustic markers of stress, providing complementary evidence for the procedure. Fourthly, the study concentrated on immediate post-task responses, without exploring acute stress recovery or the potential persistence of effects over time. This represents a limitation in understanding the temporal unfolding of acute

stress. As Epel et al. (2018) have observed, there is a paucity of longitudinal studies in this area, and such research would be valuable for capturing how acute stress responses evolve (and resolve). Finally, while the utilisation of the Prolific platform offers clear advantages in terms of accessibility, it may also introduce self-selection bias, as participants tend to be more experienced and attentive in online research contexts (Anwyl-Irvine et al., 2020; Hauser & Schwarz, 2016). This could influence how they engage with the task and may reduce their sensitivity to the acute stress protocol.

Conclusions

In summary, the present study provides preliminary evidence that this TSST-inspired online stress-induction protocol can be culturally adapted for Spanish native speakers and may elicit subjective responses consistent with the intended manipulation. Participants in the stress condition reported higher state anxiety and more negative emotional valence than those in the control condition. These findings support the potential value of the protocol for future experimental research, while also underscoring the need for further validation using larger and more diverse samples, physiological indicators, convergent measures, and longitudinal follow-up assessments.

Acknowledgements Not applicable.

Author contributions SNL: Data curation, Formal analysis, Investigation, Writing - Original draft, Writing - Review & Editing, Methodology; **JA**: Conceptualization, Investigation, Funding acquisition, Supervision, Writing - Review & Editing, Resources, Methodology, Project administration, Software.

Funding This publication is supported by the research project "Accesibilidad lingüística en la web: investigación psicolingüística para la implementación de la Lectura Fácil" funded by Fundación Ramón Areces and by the project grant PID2024-161331NB-I00 funded by the Spanish Agencia Estatal de Investigación MICIU/AEI/https://doi.org/10.13039/501100011033, where Jon Andoni Duñabeitia is the Principal Investigator in both projects; this funding covered a €5 participant compensation.

Data availability The datasets used and/or analyzed during the current study are freely available at the following link: https://doi.org/10.5281/zenodo.1585026.

Declarations

Ethical approval This study was approved by the Ethics Committee of the Universidad Nebrija (Code: UNNE-2025-048). Data collection in this study followed the ethical standards laid down in the Declaration of Helsinki of 1964 and its subsequent amendments.

Informed consent Informed consent was obtained from all individual participants included in the study.

Conflict of interest The authors have no conflict to declare

Competing Interests The author(s) declare no competing interests.

References

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use?. *Psychological Bulletin*, *83*(2), 314. https://doi.org/10.1037/0033-2909.83.2.314

Almazrouei, M. A., Morgan, R. M., & Dror, I. E. (2023). A method to induce stress in human subjects in online research environments. *Behavior Research Methods*, *55*(5), 2575–2582. https://doi.org/10.3758/s13428-022-01915-3

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Blackledge, J. T., & Hayes, S. C. (2001). Emotion regulation in acceptance and commitment therapy. *Journal of Clinical Psychology*, *57*(2), 243–255. https://doi.org/10.1002/1097-4679(200102)57:2%3C;243::aid-jclp9%3E;3.0.co;2-x

Buades-Sitjar, F., Boada, R., Guasch, M., Ferré, P., Hinojosa, J. A., Brysbaert, M., & Duñabeitia, J. A. (2021). The thousand-question Spanish general knowledge database. *Psicológica*, *2021*, 109–119. https://doi.org/10.2478/psicolj-2021-0006

Buades-Sitjar, F., Boada, R., Guasch, M., Ferré, P., Hinojosa, J. A., & Duñabeitia, J. A. (2022). The predictors of general knowledge: Data from a Spanish megastudy. *Behavior Research Methods*, *54*, 898–909. https://doi.org/10.3758/s13428-021-01669-4

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of economic behavior & organization*, *81*(1), 1–8. https://doi.org/10.1016/j.jebo.2011.08.009

Criollo, S., Guerrero-Arias, A., Buenaño Fernández, D., Jaramillo-Alcazar, Á., & Luján-Mora, S. (2024). Using mixed reality (MR) as an emerging technology for improving higher education: Analysis of mental workload. *Emerging Science Journal*, *8*, 410–424. https://doi.org/10.28991/ESJ-2024-SIED1-024

Crosswell, A. D., & Lockwood, K. G. (2020). Best practices for stress measurement: How to measure psychological stress in health research. *Health Psychology Open*, *7*(2), 2055102920933072. https://doi.org/10.1177/2055102920933072

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Epel, E. S., Crosswell, A. D., Mayer, S. E., Prather, A. A., Slavich, G. M., Puterman, E., & Mendes, W. B. (2018). More than a feeling: A unified view of stress measurement for population science. *Frontiers in Neuroendocrinology*, *49*(December 2017), 146–169. https://doi.org/10.1016/j.yfrne.2018.03.001

Gagné, N., & Franzen, L. (2023). How to Run Behavioural Experiments Online: Best Practice Suggestions for Cognitive Psychology and Neuroscience. *Swiss Psychology Open*, *3*(1), 1. https://doi.org/10.5334/spo.34

Garfin, D. R., Thompson, R. R., & Holman, E. A. (2018). Acute stress and subsequent health outcomes: A systematic review. *Journal of Psychosomatic Research*, *112*(May), 107–113. https://doi.org/10.1016/j.jpsychores.2018.05.017

Goodman, W. K., Janson, J., & Wolf, J. M. (2017). Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test. *Psychoneuroendocrinology*, *80*, 26–35. https://doi.org/10.1016/j.psyneuen.2017.02.030

Harvey, L. J., Hunt, C., & White, F. A. (2019). Dialectical behaviour therapy for emotion regulation difficulties: A systematic review. *Behaviour Change*, *36*(3), 143–164. https://doi.org/10.1017/bec.2019.9

Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, *48*(1), 400–407. https://doi.org/10.3758/s13428-015-0578-z

Heyers, K., Pfeifer, L. S., Merz, C. J., Stockhorst, U., Güntürkün, O., Wolf, O. T., & Ocklenburg, S. (2025). TSST-OL: Comparison between online and laboratory application and effects on empathy. *Psychoneuroendocrinology*, *171*(April 2024)), 107211. https://doi.org/10.1016/j.psyneuen.2024.107211

James, K. A., Stromin, J. I., Steenkamp, N., & Combrinck, M. I. (2023). Understanding the relationships between physiological and psychosocial stress, cortisol and cognition. *Frontiers in endocrinology*, *14*, 1085950. https://doi.org/10.3389/fendo.2023.1085950

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1-2), 76–81. https://doi.org/10.1159/000119004

Koncz, A., Köteles, F., Aranyossy, B., & Horváth, Á. (2025). Acute psychological stress does not influence joint position reproduction performance in the elbow joint. *PloS One*, *20*(3), e0319061. https://doi.org/10.1371/journal.pone.0319061

Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological Psychology*, *69*(1), 113–132. https://doi.org/10.1016/j.biopsycho.2004.11.009

Langer, K., Jentsch, V. L., & Wolf, O. T. (2023). Rapid effects of acute stress on cognitive emotion regulation. *Psychoneuroendocrinology*, *151*(November 2022), 106054. https://doi.org/10.1016/j.psyneuen.2023.106054

Linares, N. F. N., Charron, V., Ouimet, A. J., Labelle, P. R., & Plamondon, H. (2020). Neurobiology of Stress A systematic review of the Trier Social Stress Test methodology: Issues in promoting study comparison and replicable research. *Neurobiology of Stress*, *13*, 100235. https://doi.org/10.1016/j.ynstr.2020.100235

Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in psychology*, *6*, 121301. https://doi.org/10.3389/fpsyg.2015.00444

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., & Wagenmakers, E. J. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, *88*(1). https://doi.org/10.18637/jss.v088.i02

Lu, S., Wei, F., & Li, G. (2021). The evolution of the concept of stress and the framework of the stress system. *Cell Stress*, *5*(6), 76–85. https://doi.org/10.15698/CST2021.06.250

Marques, A., Anastasiadou, M., & Henriques, R. (2025). The Impact of Team-Based Learning on Anxiety Among Graduate Students in a Data Science Master's Program. *Emerging Science Journal*, *9*, 1–18. https://doi.org/10.28991/ESJ-2025-SIED1-01

Matsumoto, D., Yoo, S. H., & Nakagawa, S. (2008). Culture, emotion regulation, and adjustment. *Journal of personality and social psychology*, *94*(6), 925. https://doi.org/10.1037/0022-3514.94.6.925

Mavridis, A., & Tsiatsos, T. (2017). Game-based assessment: Investigating the impact on test anxiety and exam performance. *Journal of Computer Assisted Learning*, *33*(2), 137–150. https://doi.org/10.1111/jcal.12170

Miller, R., & Kirschbaum, C. (2019). Cultures under stress: A cross-national meta-analysis of cortisol responses to the Trier Social Stress Test and their association with anxiety-related value orientations and internalizing mental disorders. *Psychoneuroendocrinology*, *105*, 147–154. https://doi.org/10.1016/j.psyneuen.2018.12.236

O'Connor, D. B., Thayer, J. F., & Vedhara, K. (2021). Stress and Health: A Review of Psychobiological Processes. *Annual Review of Psychology*, *72*, 663–688. https://doi.org/10.1146/annurev-psych-062520-122331

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 17(11), 776–783. https://doi.org/10.1037/h0043424

Patall, E. A. (2021). Implications of the open science era for educational psychology research syntheses. *Educational Psychologist*, *56*(2), 142–160. https://doi.org/10.1080/00461520.2021.1897009

Szalma, J. A., Hancock, P. A., & Quinn, S. (2008, September). A meta-analysis of the effect of time pressure on human performance. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 52, No. 19, pp. 1513–1516). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/154193120805201944

Raio, C. M., & Phelps, E. A. (2015). The influence of acute stress on the regulation of conditioned fear. *Neurobiology of Stress*, *1*(1), 134–146. https://doi.org/10.1016/j.ynstr.2014.11.004

Reips, U. D. (2021). Web-Based Research in Psychology. *Zeitschrift Für Psychologie*, *229*(4), 198–213. https://doi.org/10.1027/2151-2604/a000475

Rodd, J. M. (2024). Moving experimental psychology online: How to obtain high quality data when we can't see our participants. *Journal of Memory and Language*, *134*. https://doi.org/10.1016/j.jml.2023.104472

Rodrigo-Herrero, I., Navas-León, S., & Duñabeitia, J. A. (2026). Singing reveals vocal biomarkers of stress and cognitive load. *Discover Psychology*. https://doi.org/10.1007/s44202-026-00679-7

Von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of affective disorders*, *227*, 483–493. https://doi.org/10.1016/j.jad.2017.11.048

S Shields, G. (2020). Stress and cognition: A user's guide to designing and interpreting studies. *Psychoneuroendocrinology*, *112*(May 2019), 104475. https://doi.org/10.1016/j.psyneuen.2019.104475

Santos-Carrasco, D., & De La Casa, L. G. (2025). Spanish Validation of the Maastricht Acute Stress Test (MAST): A Cost-effective Stress Induction Protocol. *The Spanish Journal of Psychology*, *28*, e1. https://doi.org/10.1017/SJP.2025.4

Schwartz, S. (2006). A theory of cultural value orientations: Explication and applications. *Comparative sociology*, *5*(2-3), 137–182.

Shields, G. S., Sazma, M. A., & Yonelinas, A. P. (2016). The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. *Neuroscience and Biobehavioral Reviews*, *68*, 651–668. https://doi.org/10.1016/j.neubiorev.2016.06.038

Shields, G. S., Sazma, M. A., McCullough, A. M., & Yonelinas, A. P. (2017). The effects of acute stress on episodic memory: A meta-analysis and integrative review. *Psychological Bulletin*, *143*(6), 636–675. https://doi.org/10.1037/bul0000100

Skoluda, N., Strahler, J., Schlotz, W., Niederberger, L., Marques, S., Fischer, S., Thoma, M. V., Spoerri, C., Ehlert, U., & Nater, U. M. (2015). Intra-individual psychological and physiological responses to acute laboratory stressors of different intensity. *Psychoneuroendocrinology*, *51*, 227–236. https://doi.org/10.1016/j.psyneuen.2014.10.002

Starcke, K., Wiesen, C., Trotzke, P., & Brand, M. (2016). Effects of acute laboratory stress on executive functions. *Frontiers in Psychology*, *7*(MAR), 1–8. https://doi.org/10.3389/fpsyg.2016.00461

Tomczak, J., Gordon, A., Adams, J., Pickering, J. S., Hodges, N., & Evershed, J. K. (2023). What over 1,000,000 participants tell us about online research protocols. *Frontiers in Human Neuroscience*, *17*(July), 1–8. https://doi.org/10.3389/fnhum.2023.1228365

Tuan, T. A., & Loan, D. T. T. (2025). Predicting Stress During Sleep from Biosensor Data: An Optimized Machine Learning Framework. *Journal of Human Earth and Future*, *6*(3), 640–653. https://doi.org/10.28991/HEF-2025-06-03-010

Wind, T. R., Rijkeboer, M., Andersson, G., & Riper, H. (2020). The COVID-19 pandemic: The 'black swan' for mental health care and a turning point for e-health. *Internet Interventions*, *20*, 100317. https://doi.org/10.1016/j.invent.2020.100317

Xin, Y., Wu, J., Yao, Z., Guan, Q., Aleman, A., & Luo, Y. (2017). The relationship between personality and the response to acute psychological stress. *Scientific Reports*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-17053-2

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.