# Time pressure in translation: Psychological and physiological measures

Yu Weng,[1] Binghan Zheng,[1] and Yanping Dong[2]

[1] Durham University | [2] Zhejiang University

Translators may experience significant psychological and physiological responses to time pressure. This study examines such responses with the aim of identifying valid indicators of time pressure in written translation. Forty-five postgraduates participated in the study, translating three comparable English texts into Chinese under three time conditions (*Short*, *Standard*, and *Free*). A positive relation between time stringency and the arousal level detected by a set of self-reporting and biomarker measures was hypothesised. The hypothesis was corroborated by results derived from participants' self-reporting on stress and anxiety, and the biomarkers of heart rate, blood pressure, and pupil dilation, but not by skin temperature, galvanic skin response (GSR), and heart rate variability (HRV). Thus, the measures that confirm the hypothesis are considered successful indicators of time pressure in translation. In addition, an inverted 'U-shaped' pattern was observed in the relation between time stringency and the arousal level indexed by GSR and HRV. These findings may facilitate research and training in translation and other cognitively demanding language-processing activities.

**Keywords:** time pressure, stress measurement, written translation, psychometric instrument, physiological stress response

## 1. Introduction

The effects of time pressure on task performance have attracted attention from various disciplines in recent decades. Existing empirical translation[1] studies have investigated the influence of time pressure on translation product quality (Ghobadi, Madadi, and Najafian 2017), translators' linguistic behaviours (Jensen and Jakobsen 2000; Hansen 2006; Alves and Liparini Campos 2009), and patterns of visual attention distribution (Sharmin et al. 2008). The psychological and physiological impacts of time pressure on translators, however, have rarely been investigated empirically. Time pressure as a form of stress could evoke particular psychological experiences and trigger the 'fight-or-flight' response, leading to various physiological effects (McCarty 2007). Investigating these psychological and physiological impacts of time pressure could facilitate the identification of valid time-pressure indicators in translation experimentation, develop our understanding of the multifaceted translation process as a whole, and contribute to better monitoring of translators' physical and mental wellbeing.

Translation is a higher-order cognitive activity (Angelone 2010), where the interplay between time pressure and intensive cognitive processing may lead to an intricate complex of responses reflected especially in the physiological arousal state. Translation is also seen as a process of problem-solving and decision-making activities which are easily influenced by stress (e.g., Wemm and Wulfert 2017). Thus, it is of interest to explore how time pressure, as a form of stress, modulates various psychological and physiological responses, in order to identify reliable indicators of stress in translation activities. This could facilitate research on

[1] The default meaning of 'translation' in this article is written translation, while other translation modalities are specified where necessary.

other stressors in translation, and further contribute to investigations of stress coping in translation practice and training.

### **1.1** Methodological considerations

Time pressure is the difference between the amount of time available and the amount of time required to execute a task (Rastegary and Landy 1993). This essentially renders it a form of psychological stress which, as Lazarus (1976) suggests, occurs when demands placed on individuals exceed their resources. Therefore, the detection of time pressure is mainly based on the measurement of various responses to stress. The methodological necessity of time-pressure measurement in relevant empirical studies investigating, for example, its impact on behaviours or performance, has long been emphasised. Without the sufficient and valid measurement of state change, it is difficult to determine "whether there are different time-pressured states, how we can distinguish these states, how we can induce these states experimentally, and whether they change the nature of judgement and decision making" (Maule and Hockey 1993, 84). Yet empirical translation research focusing on the behavioural effects of time pressure has thus far paid little attention to such fundamental methodological issues, possibly as a consequence of the relatively limited research methods involved in previous studies.

In practice, choosing feasible and appropriate methods in translation-process experimentation should take into consideration criteria such as noise-resistance, non-invasiveness, temporal resolution, and affordability (Seeber 2013). According to Bayer-Hohenwarter (2009), there are three categories of time-pressure measurement approaches that are applicable to translation-related experiments: subjective, pragmatic, and physiological measures. Examining self-reported (subjective) stress or anxiety levels under certain circumstances is the most straightforward method. In particular, (a state of) anxiety can be defined as a state of hypervigilance in anticipation of a threat that can be triggered by acute stress (Daviu et al. 2019). Thus, while stress is a particular psychological consequence of time pressure, anxiety is one's psychological reaction to stress. This legitimates the use of anxiety and stress as measurements of time pressure. Furthermore, in a stressful situation, our body releases a surge of stress hormones (adrenaline and cortisol) into the blood, which boosts the body's alertness and activates the sympathetic nervous system. As a result, the blood vessels begin to narrow, leading to an increase in blood pressure, an accelerated heart rate, and reduced Heart Rate Variability (HRV) which is the variation in the time interval between consecutive heartbeats (Jennings 2007; Sherwood and Carels 2007; Kyriakou et al. 2019). Meanwhile, stress may also enlarge the pupil size, lower the peripheral (fingertip) skin temperature, and stimulate Galvanic Skin Response (GSR) which is the continuous dynamic variations of the electrical properties of the skin (Partala and Surakka 2003; Kyriakou et al. 2019). The sensitivity of these biomarkers to the activation of the sympathetic nervous system therefore justifies their use as physiological indicators of stress or time pressure.

Bayer-Hohenwarter (2009) argues that physiological data may yield the most reliable results in detecting or measuring time pressure because of their objective nature. However, the interest in the degree of physiological activation, or arousal innervated by the sympathetic nervous system, comes from its function of preparing the body for either physical or intellectual actions (Pijeira-Díaz et al. 2018). In other words, emotional, cognitive, and motor drivers could all induce sympathetic arousal, and assessing alterations in sympathetic arousal may reveal a comprehensive spectrum of different states associated with, for example, emotion, cognition, and attention (Critchley 2002). Changes in arousal thus provide information not only about emotional states such as stress, which is generally indicated by a state of high sympathetic activation, but also about the intensity of cognitive processing, such as effort exertion during the course of task execution. Hence, it is necessary to identify arousal measures that are more sensitive to the stress-related drivers when we determine time-pressure indicators in translation.

### **1.2** Existing translation and interpreting research on time pressure or stress

As mentioned, investigation of the psychological and physiological effects of time pressure or stress on translators is scarce. To the best of our knowledge, only one recent study, Baghi and Khoshsaligheh (2019), investigated stress (but not time pressure in particular) experienced by translators in written and sight translation. They examined heart rate and blood pressure as physiological indicators of stress and found that sight translation was a more stressful modality than written translation for translators, as evidenced by a significant rise in both heart rate and blood pressure.

In contrast to studies on translation, the psychological and physiological effects of stress have received more attention in interpreting research, where a wide range of methods have been applied. Psychometric instruments and questionnaires have been used to assess, for example, conference interpreters' job stress (Cooper, Davies, and Tung 1982), student interpreters' stress levels and their coping strategies (Kao and Craigie 2013), and student interpreters' anxiety and its impact on their learning outcomes (Chiang 2009, 2010). Biomarkers such as heart rate, blood pressure, GSR, and salivary cortisol have also been frequently employed as physiological indicators of stress to investigate a range of questions, such as how interpreters' stress and effort vary under different conditions during a day's work (Klonowicz 1994); the influence of prolonged turns on the quality of the output (Moser-Mercer, Künzli, and Korac 1998); how novices and expert interpreters' stress levels differ during simultaneous interpreting (Kurz 2002, 2003); and whether the speaker's rate of delivery influences the level of stress experienced by interpreting trainees in a simultaneous interpreting task (Korpal 2016). Some studies on interpreting have triangulated psychometric instruments and physiological measures. For instance, Moser-Mercer (2005) used the salivary cortisol test in combination with the State-Trait Anxiety Inventory (STAI) to determine interpreters' stress level during remote interpreting. Physiological measures, such as GSR, combined with self-report measures, have also been used to assess emotional reactivity in bilingual emotional language processing (Jankowiak and Korpal 2018), and empathy (Korpal and Jasielska 2019). Such applications demonstrate a core feature of physiological measures: they are sensitive to the intensity rather than the type of arousal. When triangulated with subjective measures, they could be indicative of the valence and the intensity of the state. It is thus important to further explore how arousal, indexed by different biomarkers, varies across different time-pressure conditions and, more importantly, whether there are stable and reliable measures that reflect time pressure in translation activities.

### **1.3** The present study

The review in Sections 1.1 and 1.2 shows that the psychological and physiological effects of time pressure on translators have rarely been explored and that there is a need to determine indicators of time pressure in translation activities. The present study aims to identify valid time-pressure indicators by examining variation in translators' psychological experiences of anxiety and stress, and physiological responses indexed by several biomarkers (outlined in Section 2.4), under different time conditions. The hypothesis formulated for this purpose is that all the tested psychological and physiological measures will show significant effects and vary in accordance with the stringency of the time conditions. In other words, the more stringent the time condition is, the higher the arousal level, as indexed by all the proposed measures, will be.

The study makes use of a within-subject design. Each participant translated three comparable texts from English into Chinese, with each translation task rotated across three time conditions (*Short*, *Standard*, and *Free*). Variations in all the proposed psychological and physiological measures were examined to determine the effects of the three time conditions. Indicators of time pressure in translation may thus be ascertained if a positive relation emerges between time stringency and the arousal level indexed by these measures.

## 2. Methods

### **2.1** Participants

Forty-five postgraduates (thirty-nine master's students in a one-year programme and six first- or second-year doctoral candidates) majoring in Translation Studies were recruited as participants on a voluntary basis. Participants included forty females and five males with an average age of 24.4 years (range = 21–34, *SD* = 2.58). All were native Mandarin Chinese speakers with English as their second language, and had an average IELTS score of 7.4 (range = 7–8, *SD* = 0.38). All were touch-typists and had normal or corrected-to-normal vision. To minimise any negative influences on data quality, the participants were required not to drink alcohol in the twenty-four hours before the experiment, and they all reported no known diseases. A consent form ensuring anonymity and confidentiality was signed by each participant, and they were rewarded with a supermarket gift card for their participation in the experiment.

**2.2** Materials

Three source texts were selected and excerpted from *The Economist*, an international weekly newspaper focusing on current affairs, international business, politics, and technology. Each text comprised, on average, 11 sentences and 201 words (range = 196–207, *SD* = 5.51). The texts shared a similar topic: the trade war between China and the US. They were made comparable in their level of difficulty, determined by objective measures as well as subjective evaluation. The objective measures of text complexity used were based on Jensen (2009), and include readability indices (the first six items in Figure 1) and word frequency (the last item in Figure 1). Each word in each text was examined in the British National Corpus, and words which appeared fewer than 1000 times (i.e., the frequency value) in the corpus were marked as 'Frequency 1000'. The word frequency index in this study thus refers to: [(Frequency 1000 ÷ the number of total unrepeated words of the text) × 100%]. Figure 1 shows that the three texts were comparable across all these indices.

---------------------------

INSERT FIG. 1 HERE

---------------------------

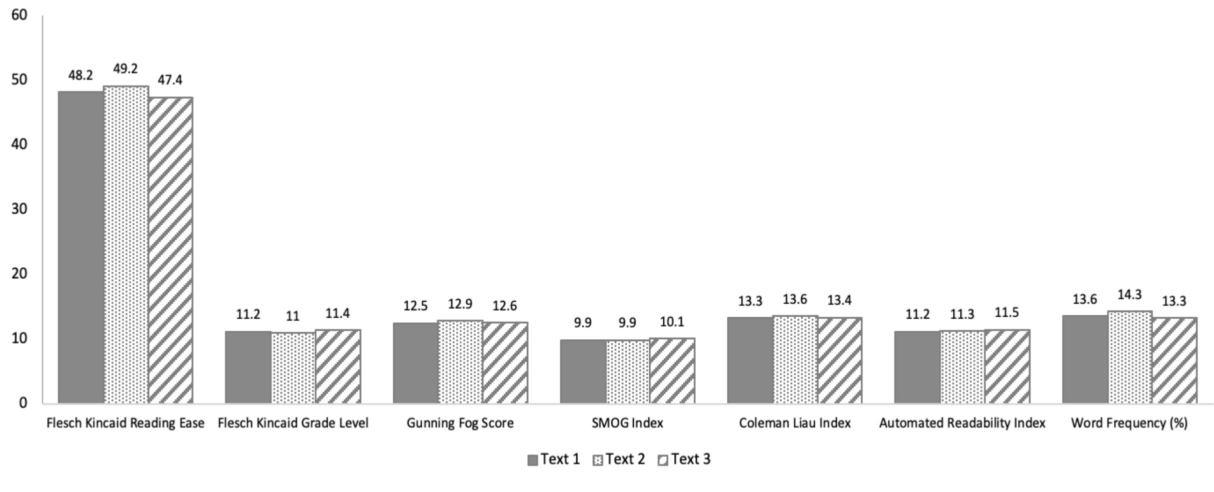


**Figure 1.** Scores of the objective measures of text difficulty

The subjective evaluation was conducted in two steps. First, eight external experts evaluated each text for 'Comprehensibility' and 'Translatability' and gave a score using a twenty-one-point scale[2], with 'very low' on the left and 'very high' on the right. Second, a pre-test involving another thirteen participants (master's students in Translation Studies) was conducted. After translating the texts, each participant retrospectively assessed the difficulty level of the text on the twenty-one-point scale. A repeated measures ANOVA shows that there is no significant difference among the three texts (Table 1). Overall, the objective measures as well as the subjective evaluation show that the three texts are comparable in (translation) difficulty.

---------------------------

INSERT TABLE 1 HERE

---------------------------

**Table 1.** Statistics on the subjective evaluation of text difficulty level

| | **Comprehensibility** | | | **Translatability** | | | **Difficulty of translation** | | |
|---|---|---|---|---|---|---|---|---|---|
| **Texts** | **Mean (*SD*)** | **Range** | ***N*** | **Mean (*SD*)** | **Range** | ***N*** | **Mean (*SD*)** | **Range** | ***N*** |
| **1** | 50.00 (10.69) | 40–60 | 8 | 52.50 (10.35) | 40–60 | 8 | 55.00 (17.08) | 30–80 | 13 |
| **2** | 40.00 (0.00) | 40–40 | 8 | 45.00 (14.14) | 20–60 | 8 | 55.38 (14.21) | 30–80 | 13 |
| **3** | 47.50 (21.21) | 40–100 | 8 | 57.50 (19.82) | 40–100 | 8 | 58.85 (10.03) | 40–70 | 13 |
| **ANOVA** | $F(2,14) = 1.36$; $\boldsymbol{p = .289}$ | | | $F(2, 14) = 1.43$, $\boldsymbol{p = .272}$ | | | $F(2, 24) = 0.71$, $\boldsymbol{p = .501}$ | | |

*SD* = standard deviation; *N* = number of participants; Range = participants' response range

## **2.3** Time constraints

[2] The difficulty assessment scales designed for the experts (comprehensibility and translatability) and students (difficulty of translation) adopted the scale of NASA Task Load Index (NASA TLX) (Hart and Staveland 1988) (see details in Section 2.4), and thus its score calculation method (raw scores) was applied: [(the number of points a participant marked - 1) × 5].

Three time conditions were designed: *Short*, *Standard*, and *Free*. In the pre-test discussed in Section 2.2, the thirteen participants were asked to translate the texts without a time constraint, and the average time interval needed for the task, 20 minutes and 25 seconds (which is also the median value), was adopted as the *Standard* translation timeframe in the present study. The timeframe for the *Short* session was determined as quartile one of the dataset, which is 16 minutes and 15 seconds. The *Free* session had no time constraint.

## **2.4** Apparatus

The perceived stress level was measured by a self-report question on stress[3] adapted from the NASA Task Load Index (NASA TLX) (Hart and Staveland 1988), a subjective workload assessment tool which has been widely used for measuring the mental workload of a task across different dimensions. Each participant was asked to assess their stress levels on the standard twenty-one-point scale, with descriptions ranging from 'very low' to 'very high'. The State-Trait Anxiety Inventory (STAI) (Spielberger et al. 1983) was employed to measure the participants' anxiety levels. STAI comprises two components: the State Anxiety Scale (S-Anxiety) which evaluates the situational anxiety level, and the Trait Anxiety Scale (T-Anxiety) which evaluates a person's stable anxiety baseline. The S-Anxiety form was completed after each session (i.e., translation of one text), and the T-Anxiety form was completed after the whole experiment as is recommended by the instrument's instructions.

---

[3] The question "How stressed were you during this task?" is adapted from the original NASA TLX question on frustration. A further three questions (on mental demand, temporal demand, and effort) from the same questionnaire were also administered, but this article focuses on the results from the stress question. The NASA TLX's score calculation method (raw scores) was consistently adopted (see Footnote 2).

The physiological measures of arousal tested in the present study are heart rate, blood pressure, pupil dilation, skin temperature, GSR, and HRV. Table 2 details these measures and their indices. The heart rate, skin temperature, HRV, and GSR data were continuously collected with a frequency of 4 Hz by a non-invasive device Empatica E4 Wristband (Empatica Srl, Milan, Italy), which does not affect the translation or typing process. The pupil size data was recorded by a Tobii TX300 eye-tracker (Tobii Technology, Stockholm, Sweden) with a sampling rate of 300 Hz. After each task, blood pressure was measured immediately using an Omron M7 Intelli IT blood pressure monitor (Omron Healthcare, Kyoto, Japan).

---------------------------

INSERT TABLE 2 HERE

---------------------------

**Table 2.** Summary of physiological measures

| Measure | Index |
|---|---|
| Heart rate | The mean value of heart rate |
| Blood pressure | Systolic blood pressure<br>Diastolic blood pressure<br>Mean arterial pressure* |
| Pupil dilation | The mean value of pupil dilation |
| Skin temperature | The mean value of skin temperature |
| Galvanic Skin Response (GSR) | The mean value of the GSR raw data (GSR Mean)<br>The mean amplitude of Skin Conductance Response (SCR Amplitude)<br>The frequency of Skin Conductance Response per minute (SCR Frequency) |
| Heart Rate Variability (HRV) | Root mean square of successive differences between the normal-to-normal heartbeat intervals (RMSSD)<br>Standard deviation of the normal-to-normal heartbeat intervals (SDNN)<br>Percentage of successive normal-to-normal heartbeat intervals that differ by more than 50 milliseconds (pNN50) |

* Mean arterial pressure is calculated using the formula [(systolic blood pressure + 2×diastolic blood pressure) ÷ 3]. The mean value of other measures (e.g., heart rate, pupil dilation, skin temperature and GSR) refers to the average of the continuously recorded data during a translation task session.

## **2.5** Procedure

The experiment was conducted at an eye-tracking lab at Durham University in the UK; and the study was approved by the research ethics committee of Durham University. After the consent form was signed, each participant was asked to fill in a basic information form before the experiment. Thereafter, the participants had a warm-up session to familiarise them with the interfaces and devices in the experimental setting. The participants were briefed before each task using an instruction: for the *Short* session, they were told that the time given for the task was less than is normally required; for the *Standard* session, they were told that the time given was the average time normally needed; and for the *Free* session, they were told that they could take as long as they needed to complete the task. In addition, the participants were informed that their translation product would be assessed by an expert and they would receive their feedback at a later date.

When the experiment started, the source text was presented at the top of the screen followed by the target-text editing box on the lower part of the screen. A countdown timer (for the *Standard* and *Short* sessions) or a stopwatch (for the *Free* session) was displayed at the bottom of the screen to remind the participants about the time remaining or time used. The purpose of visualising the elapse of time and giving the instructions was to intensify the participants' subjective perception of time pressure during the process of translation. The participants were not allowed to consult any online or offline resources during the translation tasks. However, they were provided with a glossary list of seven words in the source text which were identified in the pre-test as most likely to be unfamiliar to the participants. The participants were free to consult resources before each task regarding the words on this glossary list. The combinations of task order, text, and time condition were counterbalanced to minimise fatigue or order effects on the testing variable of Time Condition.

**2.6** Data processing and analysis

The GSR indices of SCR Amplitude and SCR Frequency were obtained via a MATLAB-based software, Ledalab (version 3.4.9; http://www.ledalab.de/). Continuous decomposition analysis (Benedek and Kaernbach 2010) was used for the extraction of SCRs, and an amplitude threshold of 0.01 μS was adopted. The indices of HRV (i.e., RMSSD, SDNN, and pNN50) were extracted using the software Kubios (Tarvainen et al. 2014). Processed using Tobii Pro Studio (Tobii Technology 2016), the pupil size data (left and right eye) for both eyes were extracted with the validity label set to 'high confidence',[4] and the data for both eyes were then averaged.

In order to show the whole picture, we present the results of all the measures below based on the analyses of the complete dataset (forty-five participants * three tasks). However, data on some measures that need to be continuously recorded (i.e., heart rate, pupil dilation, GSR, and HRV) could be partially lost due to movements of the participants during typing. Thus, these measures were also subjected to a data-screening process to ensure data quality. The results from the trimmed dataset are provided in Appendix A (along with the trimming criteria) and presented briefly in Section 3.4.

The statistical analysis was conducted using linear mixed-effects regression (LMER) modelling via the lmer4 package (Bates et al. 2015) of the statistical software R version 3.6.3 (R Core Team 2018). Separate LMER models were built for each target measure with the fixed effect of Time Condition (*Short*, *Standard*, and *Free*). In the analysis of S-Anxiety, we also considered T-Anxiety as a secondary predictor in the model. Two levels (*High* and *Low*)

[4] Tobii Pro Studio labels the confidence level that each eye has been correctly identified. The values range from 0 (high confidence) to 4 (eye not found).

were used in the categorisation of the participants' T-Anxiety based on whether it was higher than (>) or lower than/equal to (≤) the average value of the sampling group. Individual differences were considered by including each participant as a random effect in the models. *P*-values of the fixed effects in each model were obtained via Satterthwaite's approximation in the lmerTest package (Kuznetsova, Brockhoff, and Christensen 2017). The repeated measures correlation tests were conducted via the rmcorr package (Bakdash and Marusich 2017).

## 3. Results

Table 3 summarises several aspects of the task completion status in the three time conditions. For the eight participants (out of forty-five) who did not complete the task in the *Short* session, the average proportion of source-text words that they translated was 75.13%; for the two participants who did not complete the task in the *Standard* session, the average proportion of source-text words that they translated was 86.96%. This outline of translation product quantity generally demonstrates that as the time constraint became more relaxed, the participants produced more target text. The average amount of time used in the *Free* condition was 23 minutes and 10 seconds, where thirty-two of the forty-five participants (71.11%) took a longer interval than the timeframe of the *Standard* session. This means that by removing the deadline, the task time was prolonged for most participants. The results of the psychological and physiological measures analysed are presented in the following sections.

---------------------------

INSERT TABLE 3 HERE

---------------------------

**Table 3.** Task completeness in the three time conditions

| | Short | Standard | Free |
|---|---|---|---|
| Average target-text word count (Chinese characters) | 352 | 374 | 380 |
| Target-text word count range | 199–426 | 267–442 | 294–479 |
| Number of participants who did not complete the translation task | 8 | 2 | 0 |
| Task completion rate* | 82.22% | 95.56% | 100% |

*Task completion rate = [(number of participants who completed the task ÷ total number of participants) × 100%]

**3.1** Psychological measures

The Type III ANOVA table of the LMER models with Satterthwaite's method shows that the overall effect of Time Condition is significant in affecting the participants' self-reported stress ($F(2, 88) = 19.57$, $p < .001$) and S-Anxiety ($F(2, 88) = 14.08$, $p < .001$). Table 4 details how the estimates vary across different time conditions. Stress and S-Anxiety display a similar tendency: as the time condition becomes less stringent, participants' Stress and S-Anxiety levels scale down correspondingly. This tendency corroborates our hypothesis.

---------------------------

INSERT TABLE 4 HERE

---------------------------

**Table 4.** Model summary of the effect of Time Condition on Stress and S-Anxiety

| **Measure** | **Time Condition** | **Estimate** | **Std. error** | ***df*** | ***t*-value** | ***p*-value** |
|---|---|---|---|---|---|---|
| Stress | (Intercept) | 54.22 | 3.42 | 68.91 | 15.86 | < .001 |
| | Standard | -12.33 | 2.75 | 88.00 | -4.49 | **< .001** |
| | Free | -16.56 | 2.75 | 88.00 | -6.02 | **< .001** |
| S-Anxiety | (Intercept) | 41.73 | 1.75 | 65.05 | 23.84 | < .001 |
| | Standard | -4.04 | 1.32 | 88.00 | -3.07 | **.003** |
| | Free | -6.96 | 1.32 | 88.00 | -5.28 | **< .001** |

An additional analysis for S-Anxiety was conducted involving the T-Anxiety level (*High* and *Low*) as a second predictor. The results show that both Time Condition ($F(2, 86) = 12.62$, $p < .001$) and T-Anxiety ($F(1, 43) = 4.24$, $p = .045$) significantly influenced the variation in S-Anxiety, but there is no interaction between these two predictors ($F(2, 86) = 0.99$, $p = .374$). It is found that, irrespective of Time Condition, participants with a higher T-Anxiety generally tend to have higher S-Anxiety scores than those with a lower T-Anxiety. However, T-Anxiety did not modulate the effect of Time Condition on the variation of S-Anxiety.

**3.2** Physiological measures

The Type III ANOVA table (with Satterthwaite's method) of the LMER models for all the physiological measures (see Table 5) shows that Time Condition plays a significant role in modulating the participants' heart rate, systolic blood pressure, mean arterial pressure, pupil dilation, and the mean value of GSR. A marginal effect of Time Condition on diastolic blood pressure ($p = .052$) is observed, and the effect is insignificant for all the other physiological measures.

---------------------------

INSERT TABLE 5 HERE

---------------------------

**Table 5.** Type III ANOVA table of the LMER models of all the physiological measures

| Measure | Sum square | Mean square | Num. *df* | Den. *df* | *F*-value | *p*-value |
|---|---|---|---|---|---|---|
| Heart rate | 293.96 | 146.98 | 2 | 88 | 7.61 | **< .001** |
| BP: systolic blood pressure | 567.24 | 283.62 | 2 | 88 | 3.21 | **.045** |
| BP: diastolic blood pressure | 105.17 | 52.59 | 2 | 88 | 3.05 | .052 |
| BP: mean arterial pressure | 205.25 | 102.62 | 2 | 88 | 4.56 | **.013** |

| | | | | | | |
|---|---|---|---|---|---|---|
| Pupil dilation | 0.07 | 0.03 | 2 | 88 | 7.52 | **< .001** |
| Skin temperature | 0.23 | 0.11 | 2 | 88 | 0.30 | .742 |
| GSR: GSR Mean | 11.13 | 5.57 | 2 | 88 | 3.16 | **.047** |
| GSR: SCR Amplitude | 0.00 | 0.00 | 2 | 88 | 1.15 | .320 |
| GSR: SCR Frequency | 271.47 | 135.73 | 2 | 88 | 1.88 | .159 |
| HRV: RMSSD | 587.44 | 293.72 | 2 | 88 | 1.45 | .240 |
| HRV: SDNN | 255.23 | 127.61 | 2 | 88 | 1.45 | .240 |
| HRV: pNN50 | 337.83 | 168.92 | 2 | 88 | 1.49 | .230 |

BP = Blood pressure; Num. *df* = Numerator degrees of freedom; Den. *df* = denominator degrees of freedom

Table 6 details how the estimates vary across different time conditions. In general, heart rate, systolic blood pressure, mean arterial pressure, and pupil dilation show an obvious positive relation with time stringency in accordance with our initial hypothesis. For diastolic blood pressure, it is found that the highest estimate occurs in the *Short* condition (76.58 mmHg) and the lowest in the *Standard* (74.58 mmHg) rather than the *Free* condition (74.87 mmHg). Nevertheless, there is an obvious drop from the *Short* condition to the less stringent time conditions. The models of heart rate, systolic blood pressure, and pupil dilation show that the difference in estimates between the *Short* and *Free* conditions is significant, while the difference between the *Short* and *Standard* conditions is insignificant. By contrast, the decreases in mean arterial pressure from the *Short* to *Standard* and from the *Short* to *Free* conditions are all statistically significant.

For the measure of skin temperature, the peak is observed in the *Standard* condition, signalling a slight drop in arousal, while the estimates of the *Short* and *Free* conditions are relatively comparable (Table 6); the differences are, however, not significant. Interestingly, although only the mean value of GSR captures an overall significant effect of Time Condition (Table 5), all the indices of GSR and HRV display an essentially similar tendency. With higher GSR and lower HRV values signalling a more activated arousal state, the results of these two measures show that the arousal levels peak in the *Standard* condition and are lower in the *Short* and *Free* conditions. This tendency, with the *Standard* condition showing an evidently more activated arousal level than the other two conditions, is consistent across the indices of GSR and HRV, and features an inverted 'U-shaped' relation between time stringency and arousal.

---------------------------

INSERT TABLE 6 HERE

---------------------------

**Table 6.** Model summary of the effect of Time Condition on all the physiological measures

| Measure | Time Condition | Estimate | Std. error | *df* | *t*-value | *p*-value |
|---|---|---|---|---|---|---|
| Heart rate | (Intercept) | 85.15 | 1.60 | 55.30 | 53.14 | < .001 |
| | Standard | -1.72 | 0.93 | 88.00 | -1.85 | .067 |
| | Free | -3.61 | 0.93 | 88.00 | -3.90 | **< .001** |
| BP: systolic blood pressure | (Intercept) | 108.93 | 1.97 | 88.77 | 55.33 | < .001 |
| | Standard | -3.18 | 1.98 | 88.00 | -1.60 | .112 |
| | Free | -4.96 | 1.98 | 88.00 | -2.50 | **.014** |
| BP: diastolic blood pressure | (Intercept) | 76.58 | 1.18 | 64.46 | 65.07 | < .001 |
| | Standard | -2.00 | 0.87 | 88.00 | -2.29 | **.025** |
| | Free | -1.71 | 0.87 | 88.00 | -1.96 | .054 |
| BP: mean arterial pressure | (Intercept) | 87.36 | 1.26 | 68.05 | 69.31 | < .001 |
| | Standard | -2.39 | 1.00 | 88.00 | -2.39 | **.019** |
| | Free | -2.79 | 1.00 | 88.00 | -2.79 | **.006** |
| Pupil dilation | (Intercept) | 3.06 | 0.04 | 48.74 | 85.41 | < .001 |
| | Standard | -0.03 | 0.01 | 88.00 | -1.95 | .054 |
| | Free | -0.05 | 0.01 | 88.00 | -3.88 | **< .001** |
| Skin temperature | (Intercept) | 32.70 | 0.25 | 52.77 | 130.48 | < .001 |
| | Standard | 0.10 | 0.13 | 88.00 | 0.76 | .452 |
| | Free | 0.03 | 0.13 | 88.00 | 0.23 | .816 |
| GSR: GSR Mean | (Intercept) | 1.92 | 0.41 | 60.34 | 4.65 | < .001 |
| | Standard | 0.52 | 0.28 | 88.00 | 1.86 | .066 |
| | Free | -0.15 | 0.28 | 88.00 | -0.53 | .599 |
| GSR: SCR Amplitude | (Intercept) | 0.07 | 0.01 | 64.66 | 5.46 | < .001 |
| | Standard | 0.01 | 0.01 | 88.00 | 1.29 | .202 |
| | Free | 0.00 | 0.01 | 88.00 | -0.06 | .955 |
| GSR: SCR Frequency | (Intercept) | 17.45 | 2.59 | 61.24 | 6.75 | < .001 |
| | Standard | 3.37 | 1.79 | 88.00 | 1.88 | .063 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Free | 0.97 | 1.79 | 88.00 | 0.54 | .590 |
| HRV: RMSSD | (Intercept) | 68.84 | 3.43 | 74.86 | 20.08 | < .001 |
| | Standard | -2.75 | 3.00 | 88.00 | -0.92 | .362 |
| | Free | 2.36 | 3.00 | 88.00 | 0.79 | .433 |
| HRV: SDNN | (Intercept) | 47.76 | 2.25 | 75.21 | 21.20 | < .001 |
| | Standard | -1.24 | 1.98 | 88.00 | -0.63 | .533 |
| | Free | 2.09 | 1.98 | 88.00 | 1.06 | .293 |
| HRV: pNN50 | (Intercept) | 37.42 | 2.31 | 84.65 | 16.20 | < .001 |
| | Standard | -3.57 | 2.24 | 88.00 | -1.59 | .115 |
| | Free | -0.49 | 2.24 | 88.00 | -0.22 | .828 |

Figure 2 depicts the different impacts of Time Condition on twelve psychological and physiological measures/indices, with the first six (upper part of Figure 2) featuring a positive relation and the second six (lower part of Figure 2) an inverted ‘U-shaped’ relation between time stringency and arousal. Specifically, the variations in psychological measures of stress and anxiety, as well as in the biomarkers of heart rate, systolic blood pressure, mean arterial pressure, and pupil dilation, are modulated by Time Condition and feature a clear ‘step-down’ pattern: as the time constraint becomes less stringent, the psychological and physiological effects represented by these measures are alleviated. This pattern corroborates our initial hypothesis. The arousal level indexed by GSR and HRV, as mentioned, displays a remarkable inverted ‘U-shaped’ relation with time stringency, although only one of the GSR indices (GSR Mean) captures an overall statistically significant effect of Time Condition for this tendency. The visualised tendencies of diastolic blood pressure and skin temperature are not displayed as they neither capture any statistical significance nor fall in the two patterns identified above.

---------------------------

INSERT FIG. 2 HERE

---------------------------

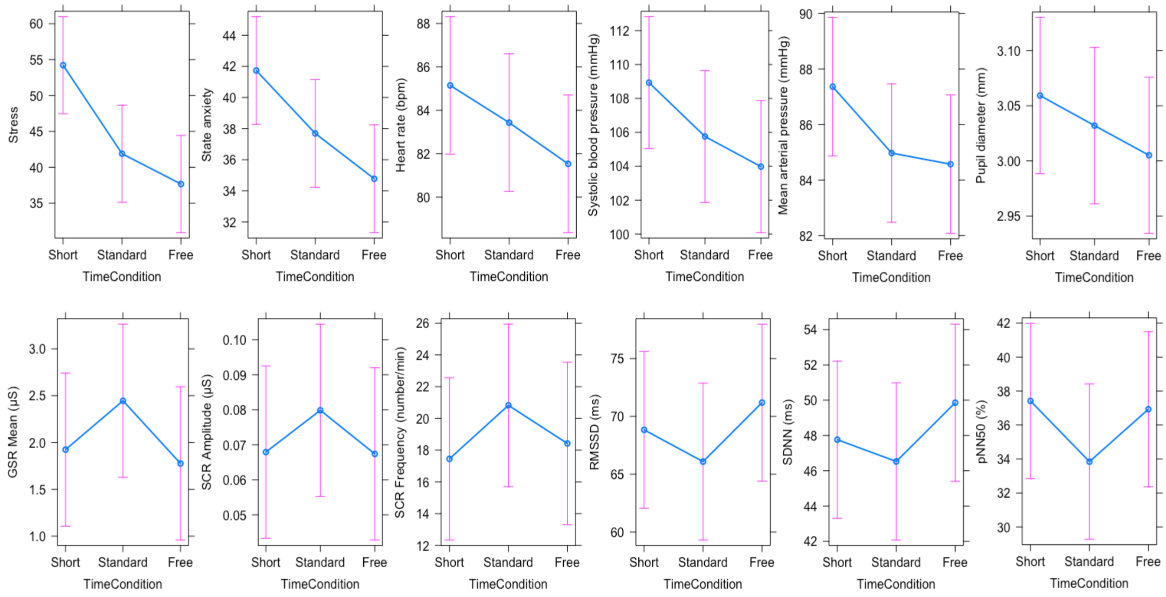


**Figure 2.** Variation in the twelve psychological and physiological measures and indices across different time conditions

**3.3** Correlations among the psychological and physiological measures

In order to further explore the relations among the tested measures, repeated measures correlation tests were conducted based on the two types of relation exhibited with time stringency (see Table 7 and Table 8).

---------------------------

INSERT TABLE 7 HERE

---------------------------

**Table 7.** Pairwise repeated measures correlation coefficient among the psychological and physiological measures that display a positive relation with time stringency

| | Anxiety | Heart rate | SBP | MAP | Pupil dilation |
|---|---|---|---|---|---|
| **Stress** | $\mathbf{r_{rm} = .42, p < .001}$ | $r_{rm} = .09, p = .372$ | $\mathbf{r_{rm} = .21, p = .047}$ | $r_{rm} = .20, p = .055$ | $r_{rm} = .18, p = .095$ |
| **Anxiety** | | $r_{rm} = .13, p = .206$ | $\mathbf{r_{rm} = .37, p < .001}$ | $\mathbf{r_{rm} = .36, p < .001}$ | $r_{rm} = .17, p = .101$ |
| **Heart rate** | | | $r_{rm} = .11, p = .300$ | $r_{rm} = .00, p = .987$ | $\mathbf{r_{rm} = .52, p < .001}$ |
| **SBP** | | | | $\mathbf{r_{rm} = .84, p < .001}$ | $r_{rm} = .17, p = .098$ |
| **MAP** | | | | | $r_{rm} = .19, p = .065$ |

Anxiety = S-Anxiety; SBP = systolic blood pressure; MAP = mean arterial pressure

Table 7 shows that each measure/index is mildly or moderately correlated with at least one of the other measures/indices in a positive relationship. Specifically, the two measures within the psychological category (i.e., Stress and S-Anxiety) show a moderate positive correlation with each other. Across the categories, the physiological measures of blood pressure (i.e., systolic blood pressure and mean arterial pressure) display a mild positive correlation with either one or both of the psychological measures. In addition, although heart rate and pupil dilation are not correlated with either Stress or S-Anxiety, these two measures show a moderate positive correlation with each other. Overall, this outline of the internal correlations among these measures further consolidates their positive relation with time stringency.

---------------------------

INSERT TABLE 8 HERE

---------------------------

**Table 8.** Pairwise repeated measures correlation coefficient among the GSR and HRV indices

| | SCR Amplitude | SCR Frequency | RMSSD | SDNN | pNN50 |
|---|---|---|---|---|---|
| **GSR Mean** | **$r_{rm}$ = .76, $p$ < .001** | **$r_{rm}$ = .51, $p$ < .001** | $r_{rm}$ = -.18, $p$ = .089 | $r_{rm}$ = -.18, $p$ = .095 | $r_{rm}$ = -.16, $p$ = .128 |
| **SCR Amplitude** | | **$r_{rm}$ = .65, $p$ < .001** | $r_{rm}$ = -.12, $p$ = .241 | $r_{rm}$ = -.11, $p$ = .318 | $r_{rm}$ = -.09, $p$ = .420 |
| **SCR Frequency** | | | $r_{rm}$ = .061, $p$ = .565 | $r_{rm}$ = .02, $p$ = .833 | $r_{rm}$ = .10, $p$ = .345 |
| **RMSSD** | | | | **$r_{rm}$ = .93, $p$ < .001** | **$r_{rm}$ = .88, $p$ < .001** |
| **SDNN** | | | | | **$r_{rm}$ = .78, $p$ < .001** |

The GSR indices are GSR Mean, SCR Amplitude, and SCR Frequency; the HRV indices are RMSSD, SDNN, and pNN50.

The results shown in Table 8 clearly demonstrate that all three indices of GSR are strongly correlated with each other and that all three indices of HRV are also strongly correlated with each other. However, no strong correlation coefficient is captured between indices across the two measures.

### **3.4** Results from the trimmed dataset

The analysis of the trimmed dataset (Appendix A, Table A1) shows no difference in tendency for heart rate and pupil dilation compared to the results with the complete dataset presented in Section 3.2: a significant positive relation is detected between time stringency and these two measures. The relation between time stringency and the arousal reflected by the GSR and HRV indices is also consistent with the findings for the full dataset, except that all three indices of HRV (i.e., RMSSD, SDNN, and pNN50) demonstrate a statistically significant effect of Time Condition. In other words, time stringency has a more significant impact on modulating HRV in the better trimmed dataset, with the exact tendency shown in the findings for the complete dataset.

## **4. Discussion**

This study aimed to identify valid psychological and physiological indicators of time pressure in translation. We examined how three time conditions (*Short*, *Standard*, and *Free*) modulate participants' self-reported stress and anxiety levels, as well as a variety of biomarkers (heart rate, blood pressure, pupil dilation, skin temperature, GSR, and HRV). The analyses in Section 3 all point to two distinctive patterns: a positive, and an inverted 'U-shaped', relation between time stringency and the detected arousal level. The psychological measures (on stress and anxiety) and some of the physiological measures (heart rate, systolic blood pressure, mean arterial pressure, and pupil dilation) corroborate the hypothesis regarding the positive relation with time stringency. Thus, these psychological and physiological measures are considered valid time-pressure indicators in this study. Specific attention has also been paid to the inverted 'U-shaped' relation between time stringency and arousal, which is visible with the physiological measures of GSR and HRV.

**4.1** Psychological effects

As expected, when the participants were translating under the more stringent time constraints, there was an obvious rise in their self-reported stress levels. We found that the stringency of the time constraint significantly affected the S-Anxiety experienced by the participants, which is in line with the tendency captured by the self-reported stress levels. The positive correlation between stress and S-Anxiety levels also confirms the consensus in the literature that anxiety is a psychological consequence of stress: a high level of stress leads to a more anxious state. In examining the interplay of T-Anxiety with Time Condition in the analysis, we found that T-Anxiety did not modulate the role that Time Condition played in affecting S-Anxiety. This indicates that in the present study the effect of time pressure is sufficiently robust to be independent of the functioning of participants' T-Anxiety. In other words, participants with both lower and higher T-Anxiety tend to have higher S-Anxiety when they are confronted with a more stringent time constraint.

Methodologically, this result substantiates assumptions that an intrinsic part of the emotional experience is the subjective feeling associated with it (Minkel and Phillips 2015). In regard to detecting time pressure in a translation experiment, the self-reported question on stress and the STAI have been proven to be valid psychometric instruments. However, despite their adequacy in detecting time pressure when used together in the present study, Slavich, Taylor, and Picard (2019) suggest the inherent limitation of subjective measures:

they can easily be affected by a "cognitive bias and social desirability" (409). Thus, it is understandable that participants tend to focus their consciousness on perceptible changes in either their stress level or anxious state in response to the overtly manipulated time constraints.

Apart from its methodological significance in translation research, the striking relationship found between time stringency and stress, and between time stringency and anxiety, have didactic implications for translation practice and translator training. On the one hand, the more translators are aware of the influence of stressors such as tight deadlines in their daily work, the better they might consciously manage them. On the other hand, stress and anxiety do not have an entirely negative impact on task performance: a number of studies have shown that negative emotions such as anxiety can enhance accuracy and coherence in translation (Rojo and Ramos Caro 2016). Thus, as a convenient means of triggering such emotions, time pressure could be utilised appropriately in translation practice and training to facilitate translators' performance.

### **4.2** Physiological effects

The physiological measures generally exhibit two types of relation between arousal and time stringency: a positive and an inverted 'U-shaped' relation. We have found that as the time constraint becomes increasingly stringent, participants tend to have elevated blood pressure, a faster heart rate, and dilated pupils, and these measures are considered reliable indicators of time pressure in the present study.

The findings on heart rate and blood pressure are partly in line with Korpal's (2016) study in which heart rate reflected stress levels in interpreting activities but systolic and diastolic blood pressure did not. Korpal (2016) indicates that this may be caused by the fact that hypertension has been recognised as a marker of chronic stress rather than the momentary experience of stress. Diastolic blood pressure in the present study only captures a marginal effect of time pressure, showing that working under the *Short* time condition is much more stressful than the less stringent conditions. More importantly, we find that the other two blood pressure indices, systolic blood pressure and mean arterial pressure, turn out to be reliable markers of time pressure which capture the expected positive relation with time stringency and show significant positive correlations with the psychological measures.

Pupil dilation has frequently been mentioned as an indicator of cognitive load or mental effort during a cognitively demanding task (Paas et al. 2003; Seeber 2013). In this study, however, pupil dilation has been proven to be a valid marker of time pressure or stress in translation activities. The significant positive correlation found between heart rate and pupil dilation further confirms this finding. Future studies should thus interpret the underlying origin of dilated pupils with caution (i.e., consider whether the stress-related factors prevail over the factor of intensive cognitive processing during the task). Skin temperature failed to demonstrate any significant effect of time stringency in this study. A possible reason for this could be that this measure is easily influenced by the physical environment in direct contact with the skin’s surface, which makes it insufficiently sensitive to detect the variation of arousal induced by the change of time conditions.

Another notable pattern is the inverted ‘U-shaped’ relation between time stringency and the arousal state reflected by the GSR and HRV indices. The results show that the arousal indexed by these two measures was highest in the moderately time-constrained task. Although only the mean value of GSR showed an overall statistical significance for this pattern, the consistent tendency that emerged for all the indices of the two measures is compelling. The result obtained from the trimmed dataset with a stronger effect of time stringency on the HRV indices further substantiates this relation.

As mentioned, apart from the stress-related origin, changes in cognitive functioning can also be reflected by the physiological arousal. Considering that GSR and HRV have been widely used as measures of cognitive load or mental effort in a variety of behavioural studies (e.g., Luque-Casado et al. 2016; Nourbakhsh et al. 2017), we suggest that these two measures might have mirrored more of the intensity of cognitive processing in the present study. In order to confirm this, we further examined the participants' responses to another question of the NASA TLX questionnaire, on effort: "How hard did you have to work to accomplish your level of performance?" (The statistical analysis of this question can be found in Appendix B.) The variation in the scores on this question also displays an inverted 'U-shaped' relation with time stringency, and the overall effect is statistically significant (Appendix B, Table B1). This finding indicates that participants might have exerted more effort in the *Standard* session under a moderate time pressure; and this more intensive cognitive processing may be evidenced in the increase in physiological arousal indexed by GSR and HRV in the *Standard* condition. As a result, it is inferred that these two physiological measures might be more sensitive to variation in cognitive-processing intensity under stressed or time-pressured conditions, rather than the mere stressed states. In this regard, further investigations could be conducted involving other types of cognitive load measures to provide more in-depth validations.

## 5. Conclusions

This study examined the variation in psychological (stress and anxiety) and physiological (heart rate, blood pressure, pupil dilation, skin temperature, GSR, and HRV) responses of translators modulated by different time-constraint conditions in translation activities. It aimed to identify valid measures of time pressure during the translation process. The results reveal that increased time stringency raises perceived stress and anxiety levels and induces elevated blood pressure, a faster heartbeat, and dilated pupils. The arousal state indexed by GSR and HRV, however, demonstrates a remarkable inverted ‘U-shaped’ relation with time stringency. We suggest that these two measures might be more indicative of the intensity of cognitive processing in the context of time-pressured translation activities. Overall, the results imply the different potential of these measures, especially the biomarkers, which are associated with either the stress-driven factors such as time pressure or factors related more to cognitive intensity during language processing.

The present study may provide methodological support for future research on diverse stressors in translation or other cognitively demanding language-processing activities. However, we are mindful of two limitations in this study. First, the controlled laboratory setting may inevitably raise some concerns about ecological validity. For example, although the participants could consult any external resources regarding the unfamiliar words provided before the experiment, they were not allowed to access resources during the experiment. This constraint may alter their decision-making behaviours compared to their routine work when resources are at hand. Second, baseline assessment of the resting condition for the physiological measures is not included. The lack of comparison between the baseline data and data from the test conditions may weaken the subsequent analysis of changes across different tasks. Thus, including the ability to consult resources and participants’ baseline data (especially for the physiological measures) would be beneficial in future research.

**Acknowledgements**

The work was supported by the National Social Science Foundation of China (No. 20BYY014). Yu Weng is funded by the China Scholarship Council (CSC).

## References


Alves, Fabio, and Tânia Liparini Campos. 2009. "Translation Technology in Time: Investigating the Impact of Translation Memory Systems and Time Pressure on Types of Internal and External Support." In *Behind the Mind: Methods, Models and Results in Translation Process Research*, edited by Susanne Göpferich, Arnt Lykke Jakobsen, and Inger M. Mees, 191–218. Frederiksberg: Samfundslitteratur.

Angelone, Erik. 2010. "Uncertainty, Uncertainty Management and Metacognitive Problem Solving in the Translation Task." In *Translation and Cognition*, edited by Gregory M. Shreve and Erik Angelone, 17–40. Amsterdam: John Benjamins.

Baghi, Hoorieh, and Masood Khoshsaligheh. 2019. "Stress in Written and Sight Translation in Training Setting." *Hikma* 18 (2): 237–255.

Bakdash, Jonathan Z., and Laura R. Marusich. 2017. "Repeated Measures Correlation." *Frontiers in Psychology* 8: 1–13.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using lmer4." *Journal of Statistical Software* 67 (1): 1–48.

Bayer-Hohenwarter, Gerrit. 2009. "Methodological Reflections on the Experimental Design of Time-Pressure Studies." *Across Languages and Cultures* 10 (2): 193–206.

Benedek, Mathias, and Christian Kaernbach. 2010. "A Continuous Measure of Phasic Electrodermal Activity." *Journal of Neuroscience Methods* 190 (1): 80–91.

Chiang, Yung-Nan. 2009. "Foreign Language Anxiety in Taiwanese Student Interpreters." *Meta* 54 (3): 605–621.

Chiang, Yung-Nan. 2010. "Foreign Language Anxiety and Student Interpreters' Learning Outcomes: Implications for the Theory and Measurement of Interpretation Learning Anxiety." *Meta* 55 (3): 589–601.

Cooper, Cary L., Rachel Davies, and Rosalie L. Tung. 1982. “Interpreting Stress: Sources of Job Stress among Conference Interpreters.” *Multilingua* 1 (2): 97–107.

Critchley, Hugo D. 2002. “Review: Electrodermal Responses: What Happens in the Brain.” *The Neuroscientist* 8 (2): 132–142.

Daviu, Nuria, Michael R. Bruchas, Bita Moghaddam, Carmen Sandi, and Anna Beyeler. 2019. “Neurobiological Links between Stress and Anxiety.” In *Stress Neurobiology Workshop 2018*, edited by Lawrence Reagan, Richard Hunter, and Matthew N. Hill, special issue of *Neurobiology of Stress* 11: 1–9.

Fink, George, ed. 2007. *Encyclopedia of Stress.* 2nd ed. Amsterdam: Academic Press.

Ghobadi, Mehdi, Golnaz Madadi, and Bahareh Najafian. 2017. “A Study of the Effects of Time Pressure on Translation Quantity and Quality.” *International Journal of Comparative Literature and Translation Studies* 5 (2): 7–13.

Hansen, Gyde. 2006. “Time Pressure in Translation Teaching and Translation Studies.” In *Interdisciplinarité en Traduction: Proceedings of the IInd International Colloquium on Translation organised by Yildiz Technical University*, vol. II, edited by Sündüz Öztürk Kasar, 71–80. Istanbul: Les Éditions Isis.

Hart, Sandra G., and Lowell E. Staveland. 1988. “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research.” In *Advances in Psychology*, vol. 52, edited by Peter A. Hancock and Najmedin Meshkati, 139-183. Amsterdam: North Holland.

Jankowiak, Katarzyna, and Paweł Korpal. 2018. “On Modality Effects in Bilingual Emotional Language Processing: Evidence from Galvanic Skin Response.” *Journal of Psycholinguistic Research* 47 (3): 663–677.

Jennings, J. Richard. 2007. “Heart Rate.” In Fink 2007, 274–277.

Jensen, Astrid, and Arnt Lykke Jakobsen. 2000. "Translating under Time Pressure: An Empirical Investigation of Problem-Solving Activity and Translation Strategies by Non-professional and Professional Translators." In *Translation in Context: Selected Contributions from the EST Congress, Granada 1998*, edited by Andrew Chesterman, Natividad Gallardo San Salvador, and Yves Gambier, 105–116. Amsterdam: John Benjamins.

Jensen, Kristian Tangsgaard Hvelplund. 2009. "Indicators of Text Complexity." In *Behind the Mind: Methods, Models and Results in Translation Process Research*, edited by Susanne Göpferich, Arnt Lykke Jakobsen, and Inger M. Mees, 61–80. Frederiksberg: Samfundslitteratur.

Kao, Po-Chi, and Philip Craigie. 2013. "Evaluating Student Interpreters' Stress and Coping Strategies." *Social Behavior and Personality* 41 (6): 1035–1043.

Klonowicz, Tatiana. 1994. "Putting One's Heart into Simultaneous Interpretation." In *Bridging the Gap: Empirical Research in Simultaneous Interpretation*, edited by Sylvie Lambert and Barbara Moser-Mercer, 213–224. Amsterdam: John Benjamins.

Korpal, Paweł. 2016. "Interpreting as a Stressful Activity: Physiological Measures of Stress in Simultaneous Interpreting." In *Language Processing in Translation*, edited by Bogusława Whyatt, special issue of *Poznań Studies in Contemporary Linguistics* 52 (2): 297–316.

Korpal, Paweł, and Aleksandra Jasielska. 2019. "Investigating Interpreters' Empathy: Are Emotions in Simultaneous Interpreting Contagious?" *Target* 31 (1): 2–24.

Kurz, Ingrid. 2002. "Physiological Stress Responses during Media and Conference Interpreting." In *Interpreting in the 21st Century: Challenges and Opportunities*, edited by Giuliana Garzone and Maurizio Viezzi, 195–202. Amsterdam: John Benjamins.

Kurz, Ingrid. 2003. "Physiological Stress During Simultaneous Interpreting: A Comparison of Experts and Novices." *The Interpreters' Newsletter* 12: 51–67.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune Haubo Bojesen Christensen. 2017. "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82 (13): 1–26.

Kyriakou, Kalliopi, Bernd Resch, Günther Sagl, Andreas Petutschnig, Christian Werner, David Niederseer, Michael Liedlgruber, Frank H. Wilhelm, Tess Osborne, and Jessica Pykett. 2019. "Detecting Moments of Stress from Measurements of Wearable Physiological Sensors." *Sensors* 19 (17): 1–26.

Lazarus, Richard S. 1976. *Patterns of Adjustment*. New York: McGraw-Hill.

Luque-Casado, Antonio, José C. Perales, David Cárdenas, and Daniel Sanabria. 2016. "Heart Rate Variability and Cognitive Processing: The Autonomic Response to Task Demands." *Biological Psychology* 113: 83–90.

Maule, A. John, and G. Robert J. Hockey. 1993. "State, Stress, and Time Pressure." In *Time Pressure and Stress in Human Judgment and Decision Making*, edited by Ola Svenson and A. John Maule, 83–101. New York: Plenum Press.

McCarty, Richard. 2007. "Fight-or-Flight Response." In Fink 2007, 62–64.

Minkel, Jared, and Samantha Phillips. 2015. "Methodological Considerations When Integrating Experimental Manipulations of Sleep and Emotion." In *Sleep and Affect*, edited by Kimberly A. Babson and Matthew T. Feldner, 119–138. London: Academic Press.

Moser-Mercer, Barbara. 2005. "Remote Interpreting: The Crucial Role of Presence." *Bulletin VALS-ASLA (Swiss Association of Applied Linguistics)* 81: 73–97.

Moser-Mercer, Barbara, Alexander Künzli, and Marina Korac. 1998. "Prolonged Turns in Interpreting: Effects on Quality, Physiological and Psychological Stress (Pilot Study)." *Interpreting* 3 (1): 47–64.

Nourbakhsh, Nargess, Fang Chen, Yang Wang, and Rafael A. Calvo. 2017. "Detecting Users' Cognitive Load by Galvanic Skin Response with Affective Interference." *ACM Transactions on Interactive Intelligent Systems* 7 (3): 1–20.

Paas, Fred, Juhani E. Tuovinen, Huib Tabbers, and Pascal W. M. Van Gerven. 2003. "Cognitive Load Measurement as a Means to Advance Cognitive Load Theory." *Educational Psychologist* 38 (1): 63–71.

Partala, Timo, and Veikko Surakka. 2003. "Pupil Size Variation as an Indication of Affective Processing." *International Journal of Human-Computer Studies* 59 (1–2): 185–198.

Pijeira-Díaz, Héctor J., Hendrik Drachsler, Paul A. Kirschner, and Sanna Järvelä. 2018. "Profiling Sympathetic Arousal in a Physics Course: How Active Are Students?" *Journal of Computer Assisted Learning* 34 (4): 397–408.

R Core Team. 2018. "R: A Language and Environment for Statistical Computing". R Foundation for Statistical Computing, Vienna, Austria. Accessed December 1, 2021. https://www.R-project.org/.

Rastegary, Haleh, and Frank J. Landy. 1993. "The Interactions among Time Urgency, Uncertainty, and Time Pressure." In *Time Pressure and Stress in Human Judgment and Decision Making*, edited by Ola Svenson and A. John Maule, 217–239. New York: Plenum Press.

Rojo, Ana, and Marina Ramos Caro. 2016. "Can Emotion Stir Translation Skill? Defining the Impact of Positive and Negative Emotions on Translation Performance." In *Reembedding Translation Process Research*, edited by Ricardo Muñoz Martín, 107–130. Amsterdam: John Benjamins.

Seeber, Kilian G. 2013. "Cognitive Load in Simultaneous Interpreting: Measures and Methods." *Target* 25 (1): 18–32.

Sharmin, Selina, Oleg Špakov, Kari-Jouko Räihä, and Arnt Lykke Jakobsen. 2008. "Where and for How Long Do Translators Look at the Screen while Translating?" In *Looking at Eyes: Eye-tracking Studies of Reading and Translation Processing*, edited by Susanne Göpferich, Arnt Lykke Jakobsen and Inger M. Mees, 31–51. Frederiksberg: Samfundslitteratur.

Sherwood, Andrew, and Robert A. Carels. 2007. "Blood Pressure." In Fink 2007, 335–342.

Slavich, George M., Sara Taylor, and Rosalind W. Picard. 2019. "Stress Measurement Using Speech: Recent Advancements, Validation Issues, and Ethical and Privacy Considerations." *Stress* 22 (4): 408–413.

Spielberger, Charles D., Richard Gorsuch, Robert E. Lushene, Peter R. Vagg, and Gerard A. Jacobs. 1983. *Manual for the State-Trait Anxiety Inventory (STAI)*. Palo Alto, CA: Consulting Psychologists Press.

Tarvainen, Mika P., Juha-Pekka Niskanen, Jukka A. Lipponen, Perttu O. Ranta-aho, and Pasi A. Karjalainen. 2014. "Kubios HRV – Heart Rate Variability Analysis Software." *Computer Methods and Programs in Biomedicine* 113 (1): 210–220.

Tobii Technology. 2016. "Tobii Studio User's Manual." Accessed December 1, 2021. https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf/?v=3.4.5.

Wemm, Stephanie E., and Edelgard Wulfert. 2017. "Effects of Acute Stress on Decision Making." *Applied Psychophysiology and Biofeedback* 42 (1): 1–12.

**Address for correspondence**

Binghan Zheng

School of Modern Languages and Cultures

Durham University

Elvet Riverside

DH1 3JT DURHAM

United Kingdom

binghan.zheng@durham.ac.uk

ORCID: 0000-0001-5302-4709

**Co-author information**

Yu Weng

School of Modern Languages and Cultures

Durham University

Yanping Dong

Department of Linguistics, School of International Studies

Zhejiang University

## Appendix A: Further data analysis

As mentioned in Section 2.6, the continuously recorded data (i.e., heart rate, HRV, pupil size, and GSR) were screened to ensure better data quality. The data screening criteria are:

1) **Heart rate and HRV data:** The length of the recorded heart rate and HRV data of each session was extracted, and its percentage in the total task time was calculated as follows: [(length of the recorded HRV data ÷ total task time) × 100%]. Sessions were discarded when this percentage was lower than 1 standard deviation (20.08%) below the mean value (43.54%) of the whole dataset. Thus, 23 out of 135 sessions (17%) were discarded.

2) **Pupil data:** The Tobii Pro Studio software gives the value of Gaze Sample Percentage [(the number of correctly identified gaze data sample ÷ the number of sampling attempts) × 100%] as an indicator of the overall quality of an eye-tracking data file. An eye-tracking data file was discarded when its Gaze Sample Percentage was lower than 1 standard deviation (10.97%) below the mean value (76.81%) of the whole dataset. Thus, 22 out of 135 sessions (16.3%) were discarded.

3) **GSR data:** Sessions with a frequency of skin conductance response less than 1/minute (5 out of 135 sessions, 3.7%) were discarded, which is based on the criterion mentioned by Pijeira-Díaz et al. (2018).

The results of the statistical analysis on the trimmed dataset are presented in Table A1.

---------------------------

INSERT TABLE A1 HERE

---------------------------

**Table A1.** Model summary of the effect of Time Condition on heart rate, pupil dilation, GSR and HRV indices (in the trimmed dataset) and type III ANOVA test (with Satterthwaite's method) of the models

| Measure | Time Condition | Estimate | Std. error | *df* | *t*-value | *p*-value | ANOVA |
|---|---|---|---|---|---|---|---|
| Heart rate | (Intercept) | 85.28 | 1.69 | 58.12 | 50.44 | < .001 | *F*(2, 69.88) = 5.57, *p* = **.006** |
| | Standard | -2.24 | 1.10 | 69.62 | -2.04 | **.045** | |
| | Free | -3.69 | 1.11 | 70.03 | -3.33 | **.001** | |
| Pupil dilation | (Intercept) | 3.06 | 0.04 | 47.71 | 84.38 | < .001 | *F*(2, 67.58) = 8.96, *p* **< .001** |
| | Standard | -0.04 | 0.01 | 67.54 | -2.71 | **.009** | |
| | Free | -0.06 | 0.01 | 67.57 | -4.16 | **< .001** | |
| GSR: GSR Mean | (Intercept) | 1.93 | 0.42 | 63.39 | 4.59 | < .001 | *F*(2, 83.65) = 3.20, *p* = **.046** |
| | Standard | 0.56 | 0.30 | 83.88 | 1.86 | .067 | |
| | Free | -0.15 | 0.29 | 83.60 | -0.51 | .613 | |
| GSR: SCR Amplitude | (Intercept) | 0.07 | 0.01 | 68.29 | 5.37 | < .001 | *F*(2, 83.80) = 1.19, *p* = .311 |
| | Standard | 0.01 | 0.01 | 84.09 | 1.28 | .203 | |
| | Free | 0.00 | 0.01 | 83.74 | -0.07 | .946 | |
| GSR: SCR Frequency | (Intercept) | 17.62 | 2.62 | 64.41 | 6.73 | < .001 | *F*(2, 83.41) = 1.83, *p* = .167 |
| | Standard | 3.49 | 1.92 | 83.66 | 1.82 | .073 | |
| | Free | 0.80 | 1.89 | 83.35 | 0.43 | .671 | |
| HRV: RMSSD | (Intercept) | 74.22 | 3.45 | 72.33 | 21.50 | < .001 | *F*(2, 71.02) = 3.48, *p* = **.036** |
| | Standard | -7.55 | 2.93 | 70.64 | -2.58 | **.012** | |
| | Free | -2.75 | 2.95 | 71.33 | -0.93 | .355 | |
| HRV: SDNN | (Intercept) | 51.14 | 2.22 | 68.50 | 23.04 | < .001 | *F*(2, 70.61) = 3.29, *p* = **.043** |
| | Standard | -4.44 | 1.78 | 70.26 | -2.50 | **.015** | |
| | Free | -1.52 | 1.80 | 70.88 | -0.84 | .401 | |
| HRV: pNN50 | (Intercept) | 41.05 | 2.41 | 82.28 | 17.05 | < .001 | *F*(2, 71.87) = 4.62, *p* = **.013** |
| | Standard | -6.96 | 2.31 | 71.44 | -3.02 | **.004** | |
| | Free | -3.03 | 2.32 | 72.32 | -1.31 | .196 | |

## Appendix B: The effect of Time Condition on the self-reported effort level

---------------------------

INSERT TABLE B1 HERE

---------------------------

**Table B1.** Model summary of the effect of Time Condition on effort level and type III ANOVA test (with Satterthwaite's method) of the model

| Measure | Time Condition | Estimate | Std. error | *df* | *t*-value | *p*-value | ANOVA |
|---|---|---|---|---|---|---|---|
| Effort | (Intercept) | 57.56 | 2.50 | 74.40 | 23.02 | < .001 | *F*(2, 88) = 4.46, *p* = **.014** |
| | Standard | 5.44 | 2.17 | 88.00 | 2.51 | **.014** | |
| | Free | -0.33 | 2.17 | 88.00 | -0.15 | .879 | |