Physiological Parameter Response to Variation of Mental Workload

Adrian Cornelius Marinescu, Sarah Sharples, Alastair Campbell Ritchie, University of Nottingham, Nottingham, United Kingdom, Tomas Sánchez López, Michael McDowell, Airbus Group Innovations UK, Newport, United Kingdom, and Hervé P. Morvan, University of Nottingham, Nottingham, United Kingdom

Objective: To examine the relationship between experienced mental workload and physiological response by noninvasive monitoring of physiological parameters.

Background: Previous studies have examined how individual physiological measures respond to changes in mental demand and subjective reports of workload. This study explores the response of multiple physiological parameters and quantifies their added value when estimating the level of demand.

Method: The study presented was conducted in laboratory conditions and required participants to perform a visual-motor task that imposed varying levels of demand. The data collected consisted of physiological measurements (heart interbeat intervals, breathing rate, pupil diameter, facial thermography), subjective ratings of workload (Instantaneous Self-Assessment Workload Scale [ISA] and NASA-Task Load Index), and the performance.

Results: Facial thermography and pupil diameter were demonstrated to be good candidates for noninvasive workload measurements: For seven out of 10 participants, pupil diameter showed a strong correlation (R values between .61 and .79 at a significance value of .01) with mean ISA normalized values. Facial thermography measures added on average 47.7% to the amount of variability in task performance explained by a regression model. As with the ISA ratings, the relationship between the physiological measures and performance showed strong interparticipant differences, with some individuals demonstrating a much stronger relationship between workload and performance measures than others.

Conclusion: The results presented in this paper demonstrate that physiological and pupil diameter can be used for noninvasive real-time measurement of workload.

Application: The methods presented in this article, with current technological capabilities, are better suited for workplaces where the person is seated, offering the possibility of being applied to pilots and air traffic controllers.

Keywords: mental workload, human performance, facial thermography, pupil diameter, physiological measures

INTRODUCTION

Since the 1980s, passenger air traffic has doubled every 15 years, and it is expected to double again by 2034, with 70% of the traffic relying on the extant network (Airbus, 2015). Near-future air transport challenges, such as increased air traffic, the need for more efficient routes, or the introduction of free flight, raise new issues of relevance to human factors. The pilot of the future will have to operate in a more congested airspace, aided by more complex technology. One human factors notion that has potential to support the management of increased demand against available cognitive capabilities is workload. This study explores techniques for measuring the mental workload experienced by participants by using noninvasive and minimally intrusive physiological measurements. These measurements have tremendous potential to aid the real-time understanding of human workload but present a number of challenges in the design and implementation of methodologies (see, e.g., Parasuraman & Mehta, 2015; Sharples & Megaw, 2015).

Mental workload has been suggested to have a strong relationship with human performance, the current consensus being that both excessively high and excessively low levels of mental workload influence performance negatively (Sharples & Megaw, 2015; Young, Brookhuis, Wickens, & Hancock, 2014). Traditionally, some methods of workload assessment have been difficult to implement in situ in a real work environment due to being invasive (e.g., interrupting tasks or requiring uncomfortable equipment to be worn). Advances in physiological sensors and data-analytic techniques mean that tools such as facial thermography (Ora & Duffy, 2007) are now realistic candidates for noninvasive capture of workload in real time. Lehrer
et al. (2010) concluded in a flight simulator study that the minimum R-R intervals (time interval between heartbeats extracted from electrocardiogram [ECG] data) in a task significantly discriminated between high- and combined moderate- and low-load tasks. Eye movement activity was used by Ahlstrom and Friedman-Berg (2006) in an air traffic control study, and they concluded that blink duration, blink frequency, mean saccade distance, and pupil diameter can provide a sensitive measure of mental workload. They have established that an increase in experienced mental workload level (subjectively measured on a scaled from 1 to 10 using the Air Traffic Workload Input Technique) is correlated with an increase in pupil diameter (Ahlstrom & Friedman-Berg, 2006). For such candidate measures to be deployed, new knowledge is required to establish the validity, reliability, and sensitivity of such tools.

Authors of previous studies have explored whether it is possible to infer mental workload by using facial thermography. These studies have shown a high correlation of workload with the decreasing temperature of the skin covering the tip of the nose. Ora and Duffy (2007) first used a simulator driving task together with a mental arithmetic loading task to increase the mental workload while measuring nose and forehead temperature, followed by performing a study in a real-car driving situation. They demonstrated that there is a strong correlation between the change in nose surface temperature and the subjective ratings for mental workload, whereas the forehead temperature remained relatively constant (Ora & Duffy, 2007). Another study in a ship simulator showed that nasal temperature and heart rate variability are good indices for effective navigation and also connected the measures to the variation of mental workload (Murai, Hayashi, Okazaki, & Stone, 2008). The reason for the nose temperature drop identified by Ora and Duffy (2007) is the vasoconstriction response of the autonomic nervous system to mental stress or negative emotion, mediated primarily by the sympathetic nervous system. Thermal imaging of the forehead, nose, eyes, cheeks, and chin during a cognitive stress test was able to classify mental workload into three levels with 81% accuracy (Stemberger, Allison, & Schnell, 2010). However, these studies did not establish the “added value” of facial thermography as a physiological tool over other techniques, such as heart rate or pupil diameter/eye movements (both of which can require the use of more intrusive and personally worn monitoring equipment).

In the study presented in this paper we explore the changes in the physiological parameters that occur as the level of mental workload varies and examine whether a combination of these parameters could be used for estimating the level of mental workload. The study uses a task that has varying level of demand with the aim of eliciting different levels of experienced workload, which are then captured by subjective and physiological measures.

The hypotheses of this study are as follows:

**Hypothesis 1:** There will be a measurable difference in subjective workload between the two levels of task difficulty.

**Hypothesis 2:** The subjective ratings of workload will be associated with changes in physiological measures.

**Hypothesis 3:** Multiple physiological measures can be used in combination to analyze workload.

**METHOD**

This research complied with the American Psychological Association Code of Ethics and was approved by the Faculty of Engineering Research Ethics Committee at University of Nottingham. Informed consent was obtained from each participant. Participants were presented with an information sheet and consent form. They verified that they were over 18 years old and had no preexisting heart-related condition and no skin conditions or allergies that could prevent them from wearing the heart rate chest strap.

**Participants**

Fourteen students and staff from the University of Nottingham took part in the study (11 men and three women; mean age = 28.3 years, $SD = 4.9$, range = 21–38). The participants were recruited via e-mail and were compensated with a £20 Amazon voucher for their time.
The data from four participants were discarded due to data-recording problems and difficulties in tracking the facial features. Data from the remaining 10 participants are presented here.

**Apparatus**

The Zephyr BioHarness 3 chest strap was used for measuring posture, heart, and breathing activity. The device outputs raw ECG data at a sampling rate of 1000 Hz and also a processed version of the raw signal including the R-R intervals, heart rate, and breathing rate (Medtronic, Annapolis, MD, USA).

For eye tracking, the RED 250 eye tracker was used in stand-alone configuration, measuring pupil diameter and gaze data at 60 Hz (Sensomotoric Instruments, Teltow, Germany).

The FLIR SC7000 thermal infrared camera with a spectral range of 3 to 5 µm was employed for monitoring the facial thermal features of the participants. The resolution of the camera is 640 × 512 pixels and was used at a sampling frequency of 50 Hz. The camera offers a noise equivalent differential temperature of less than 25 mK (FLIR Systems, Wilsonville, OR, USA).

The near-infrared light used by the eye tracker for illumination has a wavelength of 870 nm (Sensomotoric Instruments, 2011, p. 186), which is outside the 3- to 5-µm (FLIR, 2012, p. 2) spectral range of the thermal camera; therefore it does not influence the measurements of the thermal camera.

To perform the task, the participants were seated about 1.5 m away from a 55-inch (1,397-mm) LCD flat-screen display; the task covered a rectangular area of 652 mm × 718 mm (height × width), and the rest of the area was black. Although no data were recorded with regard to the light intensity in the room, it was kept as constant as possible by keeping the light off and having the blinds closed. The background of the screen during the task was black, and it was expected that as the number of balls on-screen increased, the light intensity coming from the screen would increase as well, inducing a pupil contraction response (Winn, Whitaker, Elliott, & Phillips, 1994). In fact, results demonstrated the opposite was observed, meaning that the most likely dominant factor inducing the dilation of the pupil was the task difficulty. Had the light intensity from the screen been constant, the observed effect may have been even larger.

**Materials**

In order to explore the relationship between mental workload, variation of performance, and objective physiological parameters, a specific computer-based task was designed to impose different levels of mental demand on the participant.

The task consisted of a computer game with three stages of two levels of difficulty, in total lasting 29 min; each stage consisted of 13 sub-stages (45 s each) of varying difficulty, a task paradigm previously used in our research group (Sharples, Edwards, & Balfe, 2012). Table 1 describes the task stages in terms of targets, difficulty level, and number of substages.

During each of the stages, the participant was presented with moving colored balls on a black background. The movement of the balls gives the impression that they are falling from the top of the screen. At the beginning of each of the three stages, the participant was told which were the target balls; the task was to aim at the target balls using a joystick and shoot using a button on the joystick before the balls reached the yellow line and dragged it down. During Stages 1 and 3 of Level 1 difficulty, the target balls were red (Figure 2, left), and during Stage 2 of Level 2 difficulty (Figure 2, right), the color of the balls no longer represented an identifier of the balls to be targeted. Instead, the ones having odd numbers written on them represented the target, introducing an additional cognitive element with the intent of increasing mental demand. Each of the stages was made up of 13 substages, each presenting the participant with a set number of target balls on the screen at any time; when a target ball was shot, the game generated another one. The number of balls per substage was varied as presented in Figure 1 in order to control the level of demand.

The position of the joystick was indicated by a red circular cursor that turned green once it was within range of the target balls, and the participant could make a successful shot (Figure 2) using the front button on the joystick. At the beginning of the stage, the horizontal yellow
line was at the top of the screen; when a target ball reached the yellow line, it was dragged down. The participants were told that they had to fight the balls from dragging the yellow line down by shooting at them. Whenever a target ball was shot, the yellow line went up by a small increment, and whenever the participant missed a shot, the yellow line went by the same increment. The main reasons for using the horizontal yellow line were

- to prevent participants from focusing on the balls that are high on the screen and abandoning the ones that are lower and will soon disappear off the screen, in this way subjecting all participants to the same number of targets at one time;
- to give participants a simple goal to fight toward—keeping the yellow line high up on the screen; and
- to obtain a continuous measure of performance in terms of how high on the screen they were able to maintain the yellow line at any moment.

After each substage, lasting 45 s, the participant was prompted by a voice in the task for his or her subjective assessment of mental workload, saying, “Level, please.” The task was not frozen while asking for the workload level; the participant just had to say a number from 1 to 5. At the end of each stage, the task was paused and the participants were shown the task score they achieved in comparison with the other participants as a means of increasing motivation.

Sample screen recordings of the task can be found at the following links:

- Stage 1 sample: https://www.youtube.com/watch?v=7a4MaTZ5PzE
- Stage 2 sample: https://www.youtube.com/watch?v=FNwAnWgM024

### Design

The independent variable that was manipulated during the study was the task difficulty (i.e., imposed demand). The dependent variables were the physiological measures, the subjective assessment of the perceived level of mental workload, and the task performance.

The Instantaneous Self-Assessment Workload Scale (ISA; Brennan, 1992) was used once every 45 s to collect subjective data about the mental workload.
level of perceived mental workload. The ISA was developed primarily as a subjective measure of mental workload for air traffic controllers, and it involves the participants self-rating their workload on a scale from 1 (low) to 5 (high). The main reason for using the ISA scale throughout the task was the low level of intrusion, as the participant would verbally rate the perceived level of mental workload when prompted by an auditory message (“Level, please”).

At the end of each of the three task stages, the participant filled in a NASA-Task Load Index (NASA-TLX; Hart & Staveland, 1988) questionnaire for a subjective assessment of workload. The reason for using NASA-TLX was to get a more detailed retrospective multidimensional subjective assessment of each of the three stages to determine whether the manipulation of imposed demand through task difficulty had resulted in a perceived experience of increased workload.

Procedure

Participants were invited to read the information sheet describing the details of the study and then fill in a consent form. They were then asked to play a training version of the stimulus task until they became familiar with the rules and the controls. After the training was finished, each participant was invited to attach the Zephyr sensor around his or her chest in a private space; the thermal and visual cameras were then aligned to match the height of each participant. Before starting the actual task, the eye tracker was calibrated. When the participant was ready, he or she played Stage 1 of the stimulus task, which lasted for almost 10 min, at the end of which the participant’s score was shown in comparison to the participants before. During the game play, the participant rated the level of mental workload on the ISA once every 45 s. After the first stage was over, the participant filled in the NASA-TLX questionnaire. Stages 2 (higher-demand Level 2) and 3 (original-demand Level 1) of the task then followed. Before starting each of the stages, the eye tracker was recalibrated, and after finishing each of the stages, the participant was shown his or her score and filled in a NASA-TLX questionnaire. After Stage 3 ended and the questionnaire had been completed, the participant was invited to remove the Zephyr sensor in a private space. Participants were then offered a £20 voucher as a reward for their time.

RESULTS

The results are presented in several stages. First, the results of the inferential tests to examine the impact of the manipulation of the task demand on the measures of workload and performance are presented. The aim of these tests is to confirm that the demand manipulation affected workload and performance in the man-

Figure 2. Left: Level 1 difficulty stage (Stage 1). Right: Level 2 difficulty stage (Stage 2).
ner anticipated. The second analysis examines the relationship between the different measures of workload, using bivariate correlations and reporting both correlation significance and the coefficient of determination to indicate effect size. The final analysis uses multiple linear regression to determine the percentage of variability in task performance explained by the physiological measures and the relative contribution of each of the measures.

Subjective and Performance Data

A one-way ANOVA, $F(1, 28) = 4.56, p = .041$, $\eta^2 = .14$, confirmed that there was a difference between the two levels of difficulty in terms of the NASA-TLX Mental Demand scale, confirming that Stage 2 (odd-numbered balls as targets) was perceived to be more mentally demanding than Stages 1 and 3 (red balls as targets); however, the effect size is small, group differences explaining about 14% of the variance.

One of the disadvantages of using the ISA technique is subjectivity in interpretation of the absolute meaning of numbers on the rating scale and thus the limited absolute validity that can be inferred from the ratings. However, it can be assumed that the relative validity of the ratings is robust, and therefore in order to compare the results across the participants, the data were normalized to a common scale ranging between 0 and 1.

Figure 3 shows the mean performance score for all participants (better performance in the task results in a higher score) at substage scale, plotted against the mean normalized ISA rating for all participants. There is a negative correlation between the two mean scores: The Pearson correlation coefficient is $R(37) = -0.74$ with $p < .01$, showing that as the mean subjective level of mental workload increased, the mean task performance decreased.

Whereas Figure 3 shows the mean performance and level of mental workload, Table 2 shows the individual correlations with performance of both the mean ISA normalized and each participant’s rating. It can be observed that for the individual (non-normalized) ISA ratings, three of the participants did not have significant correlations to the .05 level and the $R^2$ value is smaller in general compared with the mean ISA normalized correlation. (Note that no family-wise corrections, such as Bonferroni, were applied as tests were conducted on independent [participant-based] data sets, but it should be acknowledged that as normal when multiple tests are conducted, one in 20 will be significant by chance if a $p < .05$ level of significance is adopted.) Overall these data demonstrate a clear association between performance and subjective workload.

Physiological Data

The physiological data collected consisted of heart R-R interbeat intervals, breathing rate, pupil diameter, and facial skin temperature measured by thermography. All physiological data reported are the mean of the readings taken during the 45-s duration of each of the substages.

Due to the fact that physiological data depend so much on the physiology of each of the participants and also on the reaction each participant has to the stimulus task, the correlations of each of the physiological signals with the ISA subjective ratings (both mean normalized and individual values) will be presented in tabular form for each of the participants individually, together with strong- and weak-correlation example plots. This presentation helps us understand whether any association between physiology and subjective ratings applies across a population or whether there are different levels of strength of relationships between different predictive variables in different populations.

Table 3 shows the correlation of the R-R interbeat intervals with both the mean normalized ISA values and the individual ISA ratings; correlations with $p$ values smaller than .05 are bolded. For three of the participants (1, 6, and 9), the R-R values were significantly correlated with both the mean normalized ISA and to their individual ISA ratings. A negative moderate correlation was found for Participants 1 and 6, whereas Participant 9 showed a weak correlation with the subjective ISA ratings. The R-R values for Participants 1 and 6 showed a moderate negative correlation with their individual ISA ratings but not a significant correlation with the mean normalized values. Participant 4 was the only participant to show a positive significant correlation between R-R and mean ISA.
Physiological Response to Mental Workload

37

Figure 3. Mean Instantaneous Self-Assessment Workload Scale ratings versus mean score.

Table 2: Mean and Normalized ISA Ratings Correlated With Performance

| Participant | Mean ISA Normalized | Individual ISA |
|-------------|---------------------|----------------|
|             | R(37)   | R²     | p   | R(37)   | R²     | p   |
| 1           | -.632   | .401   | <.01| -.300   | .090   | .0632|
| 2           | -.652   | .426   | <.01| -.574   | .330   | <.01|
| 3           | -.648   | .420   | <.01| -.620   | .385   | <.01|
| 4           | -.706   | .499   | <.01| -.421   | .178   | <.01|
| 5           | -.729   | .532   | <.01| -.551   | .304   | <.01|
| 6           | -.659   | .434   | <.01| -.434   | .188   | <.01|
| 7           | -.759   | .576   | <.01| -.229   | .053   | .15 |
| 8           | -.783   | .613   | <.01| -.754   | .569   | <.01|
| 9           | -.681   | .465   | <.01| -.038   | .001   | .81 |
| 10          | -.742   | .551   | <.01| -.769   | .592   | <.01|

Note. ISA = Instantaneous Self-Assessment Workload Scale.

Figure 4 shows the R-R measure for Participant 1 plotted against mean ISA normalized and individual ISA, representing an example of strong correlation, whereas Figure 5 shows the same measures for Participant 10, representing the weakest correlation.

Table 4 shows the correlations of pupil diameter with both the mean normalized ISA values and the individual ISA ratings; pupil diameter data from all participants except for 7 and 10 have moderate to strong positive correlations with the mean ISA normalized. Participants 7 and 10 show a weak positive correlation with the individual ISA ratings. Only Participants 1 and 9 do not show a significant correlation with the individual ISA ratings. For most participants, a clear increase in pupil diameter was observed with the increase of workload. Figure 6 shows the pupil diameter measure for Participant 9 plotted against mean ISA normalized and individual ISA, representing an example of strong correlation, whereas Figure 7 shows the same
Table 5 shows the correlations of breathing rate with both the mean normalized ISA values and the individual ISA ratings; only the breathing rate data for Participant 7 showed a moderate positive correlation with the mean normalized ISA values and a weak correlation with the individual ISA values. Participant 1 showed a moderate positive correlation between breathing rate and individual ISA ratings. Figure 8 shows the breathing rate measure for Participant 7 plotted against mean ISA normalized and individual ISA, representing an example of strong correlation, whereas Figure 9 shows the same measures for Participant 2, representing the weakest correlation.

In order to extract the thermal data from the images, a feature-tracking algorithm was deployed, splitting the face into regions of interest. For each frame, the temperature was extracted from inside the circular points, from along the lines and from inside some of the triangular areas without using markers, making the technique less intrusive (Figure 10). Features from below the nose were not tracked due to the difficulty imposed by facial hair in some of the participants. The nose and forehead can therefore be considered as ideal sites for skin temperature measurement, as they would normally be unobscured, which might present a challenge in a real-life application as well.

Table 6 shows the correlations of the average temperature inside point P (nose tip) with both the mean normalized ISA values and the individual ISA ratings; only Participants 1 and 9 showed strong and moderate negative correlations at the .01 level for the mean ISA normalized. Participants 2, 7, and 10 showed weak negative correlations, significant at the .05 level, with the mean ISA normalized values. Participant 6 was the only one to show a weak positive correlation with the mean ISA normalized values. Participants 4 and 7 showed stronger correlations with the individual (non-normalized) ISA ratings.

Table 7 shows the correlations of the average temperature inside point V with both the mean normalized ISA values and the individual ISA ratings; only Participant 1 showed a strong negative correlation at the .01 level for the mean ISA normalized, whereas Participants 2 and 4 showed a weak negative correlation, significant at the .05 level, with the mean ISA normalized. Participants 1, 4, and 7 showed moderate to weak negative correlations with the individual ISA values. Figure 11 shows the temperature in points P and V for Participant 1 plotted against mean ISA normalized and individual ISA, representing an example of strong correlation, whereas Figure 12 shows the same measures for Participant 8, representing the weakest correlation.

Table 3: R-R Intervals Correlated With Subjective ISA Reports

| Participant | Mean ISA Normalized | Individual ISA |
|-------------|---------------------|----------------|
|             | R(37) | R²   | p   | R(37) | R²   | p   |
| 1           | -.696 | .484 | <.01 | -.535 | .286 | <.01 |
| 2           | -.197 | .039 | .22  | .016  | 0    | .92  |
| 3           | -.173 | .030 | .29  | -.183 | .033 | .26  |
| 4           | .47   | .221 | <.01 | .079  | .006 | .62  |
| 5           | -.276 | .076 | .08  | -.323 | .104 | .04  |
| 6           | -.573 | .328 | <.01 | -.454 | .206 | <.01 |
| 7           | .185  | .034 | .25  | -.202 | .041 | .21  |
| 8           | -.222 | .049 | .17  | -.198 | .039 | .22  |
| 9           | -.349 | .122 | .02  | -.327 | .107 | .04  |
| 10          | -.05  | .003 | .75  | -.112 | .013 | .49  |

Note. ISA = Instantaneous Self-Assessment Workload Scale.
Table 8 shows the correlations of the average temperature inside point L with both the mean normalized ISA values and the individual ISA ratings; Participants 1, 2, 9, and 10 showed moderate negative correlations and Participant 7 showed weak negative correlations with the mean ISA normalized levels. Participants 4 and 7 showed a moderate to weak negative correlation with the individual ISA ratings.

Table 9 shows the correlations of the average temperature inside point M with both the mean normalized ISA values and the individual ISA ratings; Participants 1, 2, 4, 7, 9, and 10 showed a moderate negative correlation with the mean ISA normalized, whereas Participant 5 showed a weak negative correlation with the mean ISA normalized. Participants 2, 4, and 10 showed a moderate negative correlation with the individual ISA ratings, whereas Participant 7 showed a weak negative correlation with the individual ISA ratings. Figure 13 shows the temperature in points L and M for Participant 2 plotted against mean ISA normalized and individual ISA, representing an example of strong correlation.
whereas Figure 14 shows the same measures for Participant 6, representing the weakest correlation.

**Predictive Power of Combined Physiological Measures**

In this section we explore how the different physiological measures can be combined to produce the most accurate prediction of performance. Some of the measures presented earlier show promising correlations with the subjective ISA measure of mental workload.

A multiple linear regression was performed for each participant individually on more combinations of the predictor variables to test which one explains more of the variability in the response variable and how different physiological parameters can be combined for more reliable and valid capture of workload. Four combinations of the predictor variables were chosen:

1. Heart (R-R interval) and breathing rate data
2. The heart and breathing rate data and pupil diameter

*Figure 5. R-R versus mean Instantaneous Self-Assessment Workload Scale (ISA) normalized and individual ISA for Participant 10 (weakest correlation).*
The heart and breathing rate data, pupil diameter, and facial temperature inside points B, F, G, H, L, M, P, V.

Facial temperature inside points B, F, G, H, L, M, P, and V.

The reason behind the choice of the predictor variable combinations was to start with features from only one of the sensors and gradually add the others; Category 1 contains just the features produced by the Zephyr sensor, Category 2 adds pupil diameter to Category 1, Category 3 contains the combined features from the first two categories in addition to the facial thermography measures, and Category 4 contains just the facial thermography features.

Game performance, rather than ISA ratings, was selected as the response variable for this analysis. Game performance was used as it strongly correlates with the subjective ISA ratings and it is also a continuous variable; game performance is represented by the height at which the participants managed to maintain the yellow line on the screen.

Some of the predictor variables for some of the participants were highly correlated with each other. Intervariable correlation influences the ability of multiple linear regression to distinguish between the predictive ability of each individual variable. Our approach to this limitation was to systematically add and remove predictors based on the $F$ statistic; the tool used for this was stepwise regression in Matlab. The algorithm starts with a constant model and iteratively adds and removes predictors until the model can no longer be improved substantially.

Each of the sections in Table 10 shows the multiple linear regression results for each of the described groups of predictors for each participant. The Adjusted $R^2$ column contains the proportion of variability of the dependent variable accounted for by the regression model. Because the $R^2$ value increases by adding more predictor variables in the model, the adjusted $R^2$ value was reported in order to make the comparison between models more meaningful. The table also displays the $F$ statistic of the linear fit versus the constant model, testing the statistical significance of the model; the Predictor column contains the names of the predictors selected by the algorithm for each of the regressions. The Beta column contains the estimated standardized coefficients of the terms in the regression, indicating how many standard deviations the dependent variable will change with the change of one standard deviation in the predictor variable, allowing for a comparison of the relative contribution of each of the predictors. The $r$ statistic test for the significance of each term given the other terms in the model is used to test the null hypothesis that the term is equal to zero (versus the alternate hypothesis that the coeffi-
cients is different from zero). The associated $p$ values are also reported in the table.

The results presented in Table 10 show that for Combination 3, when using all the predictor variables, for seven out of 10 participants, the pupil diameter measure was demonstrated to be a good predictor of performance, followed by temperature in point P for six out of 10 participants. On average, facial thermography measures added 47.7% to the amount of variability explained by the regression model.

Figure 15 shows a boxplot summary of Table 10 in terms of adjusted $R^2$ and root mean square error (RMSE). It can be seen that for the predictors in Combination 3, the amount of variability explained is larger than all other combinations but close to Combination 4. At the same time, the RMSE is smallest for Combination 3, indicating a better fit compared with the other models. Based on the data collected in this study, for most of the participants, pupil diameter together with thermal

---

**Figure 6.** Pupil diameter versus mean Instantaneous Self-Assessment Workload Scale (ISA) normalized and individual ISA for Participant 9 (strong correlation for mean ISA normalized but weak correlation for individual ISA).
data measured around the nose area provided the best combination of predictors for inferring the level of performance.

**DISCUSSION**

This research presents novel insights into the relative value of physiological and subjective techniques for assessment of workload and human performance. The main novelty lies in the fact that multiple continuous physiological measures were recorded and synchronized with task performance and subjective ratings. The hypotheses explored in this study were as follows:

*Hypothesis 1*: There will be a measurable difference in subjective workload between the two levels of task difficulty.

This hypothesis was found to be true: The mental demand measured using NASA-TLX confirmed that there was a measurable difference

*Figure 7.* Pupil diameter versus mean Instantaneous Self-Assessment Workload Scale (ISA) normalized and individual ISA for Participant 7 (nonsignificant correlation for mean ISA normalized but weak positive correlation with individual ISA ratings).
Table 5: Breathing Rate Correlated With Subjective ISA Reports

| Participant | Mean ISA Normalized | Individual ISA |
|-------------|---------------------|----------------|
|             | R(37) | R² | p   | R(37) | R² | p   |
| 1           | .115  | .013 | .48 | .419  | .176 | <.01|
| 2           | .0005 | 0   | .99 | -.144 | .021 | .38 |
| 3           | -.095 | .009 | .56 | -.169 | .029 | .30 |
| 4           | .061  | .004 | .71 | .242  | .059 | .13 |
| 5           | .097  | .009 | .55 | .038  | .001 | .81 |
| 6           | -.258 | .067 | .11 | -.122 | .015 | .45 |
| 7           | .661  | .437 | <.01| .393  | .154 | .01 |
| 8           | .14   | .02  | .39 | .137  | .019 | .40 |
| 9           | .304  | .092 | .05 | .088  | .008 | .59 |
| 10          | .297  | .088 | .08 | .232  | .054 | .15 |

Note. ISA = Instantaneous Self-Assessment Workload Scale.

Figure 8. Breathing rate versus mean Instantaneous Self-Assessment Workload Scale (ISA) normalized and individual ISA for Participant 7 (strongest correlation with mean ISA normalized).
between the two levels of difficulty, and Stage 2 was perceived to be more mentally demanding than Stages 1 and 3.

Hypothesis 2: The subjective ratings of workload will be associated with changes in physiological measures.

Hypothesis 2 was partially proved: We explored which physiological measures showed a change in accordance to the change in mental workload as measured subjectively on the ISA. It was found that for some of the participants, the mean normalized ISA ratings showed a stronger correlation.
TABLE 6: Point P Temperature Correlated With Subjective ISA Reports

| Participant | Mean ISA Normalized | Individual ISA |
|-------------|---------------------|----------------|
|             | R(37) | R^2 | p   | R(37) | R^2 | p   |
| 1           | -.746 | .557 | <.01 | -.507 | .257 | <.01 |
| 2           | -.373 | .139 | .01  | -.07  | .005 | .66  |
| 3           | -.075 | .006 | .64  | -.137 | .019 | .40  |
| 4           | -.152 | .023 | .35  | -.429 | .184 | <.01 |
| 5           | -.167 | .028 | .30  | -.008 | .000 | .95  |
| 6           | .345  | .119 | .03  | .188  | .035 | .25  |
| 7           | -.401 | .161 | .01  | -.459 | .211 | <.01 |
| 8           | -.086 | .007 | .60  | -.042 | .002 | .79  |
| 9           | -.514 | .264 | <.01 | -.208 | .043 | .20  |
| 10          | -.329 | .108 | .04  | -.028 | .001 | .86  |

Note. ISA = Instantaneous Self-Assessment Workload Scale.

TABLE 7: Point V Temperature Correlated With Subjective ISA Reports

| Participant | Mean ISA Normalized | Individual ISA |
|-------------|---------------------|----------------|
|             | R(37) | R^2 | p   | R(37) | R^2 | p   |
| 1           | -.724 | .524 | <.01 | -.468 | .219 | <.01 |
| 2           | -.354 | .125 | .02  | -.05  | .003 | .76  |
| 3           | -.267 | .071 | .09  | -.1   | .010 | .54  |
| 4           | -.382 | .146 | .01  | -.381 | .145 | .01  |
| 5           | -.284 | .081 | .07  | -.132 | .017 | .41  |
| 6           | .146  | .021 | .37  | -.118 | .014 | .47  |
| 7           | -.296 | .088 | .06  | -.382 | .146 | .01  |
| 8           | -.107 | .011 | .51  | -.075 | .006 | .64  |
| 9           | -.035 | .001 | .83  | -.118 | .014 | .47  |
| 10          | -.307 | .094 | .05  | -.156 | .024 | .34  |

Note. ISA = Instantaneous Self-Assessment Workload Scale.

with some of their physiological measures than it did with the individual ISA rating.

Table 11 summarizes the results by displaying the number of participants who showed moderate to strong correlations with mean ISA normalized or individual ISA ratings for each of the physiological measures presented earlier. Overall, the correlations of the thermal data with the individual ISA ratings were weaker for all participants.

_Hypothesis 3:_ Multiple physiological measures can be used in combination to analyze workload.

Hypothesis 3 was tested by using a multiple linear regression on the data from each of the participants, showing that when using facial thermography data combined with other physiological data, the predictive model explains on average 47.7% more of the variability in performance compared with using solely a combination of R-R interbeat intervals, breathing rate, and pupil diameter. As mean performance across the participants was strongly correlated with the mean ISA normalized, it is an indication that these physiological measures could also provide good prediction results for
the level of subjectively experienced mental workload.

In their Discussion section, Ora and Duffy (2007) recommended that further examination under more controlled conditions and the test of additional psychophysiological measures, such as pupil dilation, should be performed in the hope of developing a more robust approach to the estimation of mental workload in a noninvasive way. In this study, the variation of demand was done in more controlled conditions, and additional physiological measures (such as heart rate, breathing rate, and pupil diameter) were collected and their relative group contribution was tested. In terms of facial thermography, the landmark tracking was done automatically and included more areas of the face. One of the limitations of the study was the small number of participants; for the limited number of participants (10), there was no physiological measure that proved to work best at predicting mental workload or performance levels across all participants. Although from a physiological point of view, people responded differently when being sub-

Figure 11. P, V points temperature versus mean Instantaneous Self-Assessment Workload Scale (ISA) normalized and individual ISA for Participant 1 (strong correlation example).
jected to the type of demand induced by the task, some of the physiological measures—especially pupil diameter and temperature in points G, M, and P—proved to be good and consistent indicators of the level of performance (and implicitly the level of demand) for more than half of the participants.

Further studies will concentrate on the collection of more data in environments closer to the real workplace setting and the use of machine learning algorithms to improve prediction accuracy, confirm feasibility of applying the physiological and analytical methods in situ, and ensure generalizability of results. Authors of future work should also consider how facial thermography measurements would vary over longer periods than have been examined in this study.

The results presented in this paper demonstrate that physiological measures, especially face temperature and pupil diameter, can be used for noninvasive real-time measurement of workload when combined with a facial-landmark-tracking algorithm, assuming mod-
Physiological Response to Mental Workload

Levels have been appropriately trained on previously recorded data from the user population. This proposition is feasible in a setting such as cockpits.

The demonstration of feasibility of physiological measures as a method as presented within this paper allows the identification of guidance for how this approach can be used in the future and requirements for further research. The methods presented in this article, with current technological capabilities, are better suited for workplaces where the subject is seated, but the methods can cope with a limited amount of head movement. Continuous real-time noninvasive workload measurement techniques are now a realistic proposition that will allow for improved design of human–machine systems, operating procedures, and operations scheduling in ways that will bring us closer to the goal of optimizing human well-being and overall system performance.

ACKNOWLEDGMENTS

The authors would like to thank the European Union for founding this research from the People Programme (Marie Curie Actions) of the European

---

TABLE 8: Point L Temperature Correlated With Subjective ISA Reports

| Participant | Mean ISA Normalized | Individual ISA |
|-------------|---------------------|----------------|
|             | R(37) | R² | p  | R(37) | R² | p  |
| 1           | -.533  | .284 | <.01 | -.249  | .062 | .12 |
| 2           | -.501  | .251 | <.01 | -.292  | .085 | .07 |
| 3           | -.196  | .038 | .23  | -.006  | 0   | .96 |
| 4           | -.08   | .006 | .59  | -.471  | .222 | <.01|
| 5           | -.289  | .084 | .07  | -.155  | .024 | .34 |
| 6           | -.025  | .001 | .87  | -.023  | .001 | .88 |
| 7           | -.373  | .139 | .01  | -.377  | .142 | .01 |
| 8           | -.16   | .026 | .33  | -.082  | .007 | .61 |
| 9           | -.594  | .353 | <.01 | -.157  | .025 | .33 |
| 10          | -.457  | .209 | <.01 | -.169  | .029 | .30 |

Note. ISA = Instantaneous Self-Assessment Workload Scale.

---

TABLE 9: Point M Temperature Correlated With Subjective ISA Reports

| Participant | Mean ISA Normalized | Individual ISA |
|-------------|---------------------|----------------|
|             | R(37) | R² | p  | R(37) | R² | p  |
| 1           | -.472  | .223 | <.01 | -.251  | .063 | .12 |
| 2           | -.674  | .454 | <.01 | -.511  | .261 | <.01|
| 3           | -.081  | .007 | .62  | .076   | .006 | .64 |
| 4           | -.419  | .176 | <.01 | -.509  | .259 | <.01|
| 5           | -.386  | .149 | .01  | -.248  | .062 | .12 |
| 6           | -.116  | .013 | .48  | .069   | .005 | .67 |
| 7           | -.542  | .294 | <.01 | -.358  | .128 | .02 |
| 8           | -.16   | .026 | .32  | -.071  | .005 | .66 |
| 9           | -.643  | .413 | <.01 | -.186  | .035 | .25 |
| 10          | -.543  | .295 | <.01 | -.584  | .341 | <.01|

Note. ISA = Instantaneous Self-Assessment Workload Scale.
Figure 13. L, M points temperature versus mean Instantaneous Self-Assessment Workload Scale (ISA) normalized and individual ISA for Participant 2 (strongest correlations).
Physiological Response to Mental Workload

Figure 14. L, M points temperature versus mean Instantaneous Self-Assessment Workload Scale (ISA) normalized and individual ISA for Participant 6 (nonsignificant correlations).
TABLE 10: Proportion of the Variability Accounted for by the Regression Model in the Response Variable

| Combination | Participant | Adjusted $R^2$ | RMSE | $F$ | $p$ | Predictor | Beta | $t$ | $p$ |
|-------------|-------------|----------------|-------|-----|-----|-----------|------|-----|-----|
| Combination 1: Mean RR, mean BR | 1 | .404 | 0.77 | 13.91 | <.01 | Mean RR | 0.56 | 4.34 | <.01 |
| | 2 | 0 | — | — | — | — | — | — | — |
| | 3 | .187 | 0.9 | 9.75 | <.01 | Mean BR | 0.45 | 3.12 | <.01 |
| | 4 | 0 | — | — | — | — | — | — | — |
| | 5 | 0 | — | — | — | — | — | — | — |
| | 6 | .434 | 0.75 | 30.16 | <.01 | Mean RR | 0.67 | 5.49 | <.01 |
| | 7 | .265 | 0.85 | 14.76 | <.01 | Mean BR | -0.53 | -3.84 | <.01 |
| | 8 | 0 | — | — | — | — | — | — | — |
| | 9 | .224 | 0.88 | 11.97 | <.01 | Mean RR | 0.49 | 3.46 | <.01 |
| | 10 | .116 | 0.93 | 6.01 | <.01 | Mean BR | -0.37 | -2.45 | <.01 |
| Combination 2: Mean RR, mean BR, pupil diameter | 1 | .404 | 0.77 | 13.91 | <.01 | Mean RR | 0.56 | 4.34 | <.01 |
| | 2 | 0 | — | — | — | — | — | — | — |
| | 3 | .187 | 0.9 | 9.75 | <.01 | Mean BR | 0.45 | 3.12 | <.01 |
| | 4 | .338 | 0.81 | 20.43 | <.01 | Pupil diameter | -0.59 | -4.52 | <.01 |
| | 5 | .092 | 0.95 | 4.85 | <.05 | Pupil diameter | -0.34 | -2.2 | <.05 |
| | 6 | .434 | 0.75 | 30.16 | <.01 | Mean RR | 0.67 | 5.49 | <.01 |
| | 7 | .265 | 0.85 | 14.76 | <.01 | Mean BR | -0.53 | -3.84 | <.01 |
| | 8 | .280 | 0.84 | 15.84 | <.01 | Pupil diameter | -0.54 | -3.98 | <.01 |
| | 9 | .696 | 0.55 | 88.39 | <.01 | Pupil diameter | -0.83 | -9.4 | <.01 |
| | 10 | .474 | 0.72 | 18.15 | <.01 | Mean RR | 0.45 | 3.82 | <.01 |
| | | | | | Pupil diameter | -0.62 | -5.24 | <.01 |
| Combination 3: Mean RR, mean BR, pupil diameter, temperature inside points B, F, G, H, L, M, P, V | 1 | .786 | 0.46 | 36.03 | <.01 | Pupil diameter | -0.33 | -3.72 | <.01 |
| | 2 | .888 | 0.33 | 51.42 | <.01 | Mean RR | -0.24 | -2.14 | <.05 |
| | | | | | Mean BR | 0.32 | 4.4 | <.01 |
| | | | | | G | -0.33 | -2.83 | <.01 |
| | | | | | M | 0.82 | 9.17 | <.01 |
| | | | | | P | -0.49 | -3.84 | <.01 |
| | | | | | V | 0.81 | 7.58 | <.01 |
| | 3 | .915 | 0.29 | 103.78 | <.01 | Pupil diameter | -0.13 | -2.39 | <.05 |

(continued)
| Combination | Participant | Adjusted $R^2$ | RMSE | $F$ | $p$ | Predictor | Beta | t  | $p$ |
|-------------|-------------|----------------|------|-----|-----|-----------|------|----|-----|
| 4           | 0.856       | 0.37           | 57.57| <.01|     | B         | -1.59| -16.33| <.01|
|             |             |                |      |     |     | G         | -0.66| -8.84| <.01|
|             |             |                |      |     |     | H         | 1.34 | 18.46| <.01|
|             | 0.26        | 2.61           | <.05 |     |     | M         | 0.58 | 7.03 | <.01|
|             | 0.21        | 2.87           | <.01 |     |     | V         |      |      |     |
| 5           | 0.692       | 0.55           | 29.51| <.01|     | Mean BR   | 0.28 | 2.54 | <.05|
|             | F           | -0.26          | -7.35| <.01|     | M         | 0.85 | 7.51 | <.01|
|             | G           | -0.63          | -2.8 |     |     | P         |      |      |     |
|             | P           | -0.32          | -2.35| <.05|     |           |      |      |     |
| 6           | 0.763       | 0.48           | 25.6 | <.01|     | Mean RR   | 0.62 | 5.47 | <.01|
|             | Pupil diameter | -0.27  | -2.7 | <.05|     | B         | 1.47 | 5.89 | <.01|
|             |             |                |      |     |     | F         | -1.38| -7.35| <.01|
|             |             |                |      |     |     | P         | -0.32| -2.35| <.05|
| 7           | 0.787       | 0.46           | 24.49| <.01|     | Mean RR   | -0.26| -2.46| <.05|
|             | Mean BR     | -0.3           | -3.03| <.01|     | Pupil diameter | -0.24| -2.34| <.05|
|             |             |                |      |     |     | L         | 0.37 | 2.43 | <.05|
|             |             |                |      |     |     | M         | 1.51 | 6.85 | <.01|
|             |             |                |      |     |     | P         | -1.46| -5.01| <.01|
| 8           | 0.712       | 0.53           | 24.5 | <.01|     | Pupil diameter | -0.67| -6.56| <.01|
|             | G           | -0.63          | -2.8 |     |     | M         | 1.65 | 4.83 | <.01|
|             | M           | -1.47          | -7.38| <.01|     | P         |      |      |     |
| 9           | 0.841       | 0.39           | 51.43| <.01|     | Mean BR   | -0.2 | -3.01| <.01|
|             | Mean BR     | -0.78          | -7.49| <.01|     | Pupil diameter | -0.45| -5.58| <.01|
|             |             |                |      |     |     | G         | 0.34 | 3.17 | <.01|
|             |             |                |      |     |     | P         |      |      |     |
| 10          | 0.7         | 0.54           | 30.55| <.01|     | Pupil diameter | -0.62| -6.54| <.01|
|             | G           | -0.45          | -5.58| <.01|     | V         | 0.49 | 5.01 | <.01|
|             | V           | 0.26           | 2.53 | <.05|     |           |      |      |     |
| Combination 4: | 0.708       | 0.54           | 31.71| <.01|     | B         | -0.94| -6.65| <.01|
| Temperature inside points B, F, G, H, L, M, P, V | | | | | | F | 0.58 | 3.92 | <.01|
|             | P           | 0.68           | 6.94 | <.01|     | M         | 1.01 | 10.1 | <.01|

(continued)
TABLE 10: (continued)

| Combination | Participant | Adjusted $R^2$ | RMSE | $F$ | $p$ | Predictor | Beta | $t$ | $p$ |
|-------------|-------------|----------------|------|-----|-----|-----------|------|-----|-----|
| 3           | P           | -0.76          | 5.22 | <.01|     |           |      |     |     |
|             | V           | 0.55           | 4.8  | <.01|     |           |      |     |     |
| 4           | B           | -1.55          | 15.18| <.01|     |           |      |     |     |
|             | G           | -0.64          | 8.09 | <.01|     |           |      |     |     |
|             | H           | 1.38           | 18.33| <.01|     |           |      |     |     |
| 5           | F           | 0.26           | 2.61 | <.05|     |           |      |     |     |
|             | M           | 0.58           | 7.03 | <.01|     |           |      |     |     |
|             | V           | 0.21           | 2.87 | <.01|     |           |      |     |     |
| 6           | B           | 0.79           | 2.61 | <.05|     |           |      |     |     |
|             | F           | -1.5           | -6.66| <.01|     |           |      |     |     |
|             | G           | -0.93          | -4.25| <.01|     |           |      |     |     |
|             | L           | 1.14           | 5.73 | <.01|     |           |      |     |     |
| 7           | B           | -0.49          | -3.27| <.01|     |           |      |     |     |
|             | G           | -0.97          | -3.88| <.01|     |           |      |     |     |
|             | L           | 1.08           | 4.32 | <.01|     |           |      |     |     |
|             | M           | 1.27           | 5.04 | <.01|     |           |      |     |     |
|             | P           | -0.83          | -3.07| <.01|     |           |      |     |     |
| 8           | 0           | —              | —    |     |     |           |      |     |     |
| 9           | G           | 0.89           | 6.91 | <.01|     |           |      |     |     |
|             | P           | -0.3           | -2.33| <.05|     |           |      |     |     |
| 10          | G           | 1.29           | 6.71 | <.01|     |           |      |     |     |
|             | L           | -1.21          | -3.9 | <.01|     |           |      |     |     |
|             | P           | 0.58           | 2.46 | <.05|     |           |      |     |     |

Note. RMSE = root mean square error; RR = heart rate; BR = breathing rate.

Union’s Seventh Framework Programme (FP7/2007-2013) under REA Grant Agreement No. 608322 and the Data Analysis and Interaction research team from Airbus Group Innovations UK for their support. We would also like to thank Robert Houghton for offering insightful comments that greatly improved the quality of the manuscript.

KEY POINTS

- One of the challenges posed by the future of air transportation, from a human factors perspective, is evaluating the level of mental workload to which the operators are subjected.
- Some methods of workload assessment have been difficult to implement in situ in a real work environment due to being intrusive (e.g., interrupting task or requiring uncomfortable equipment to be worn).
- We explore multiple physiological measures and their relative significance as indicators of performance and mental workload, demonstrating the feasibility of physiological measures as a method
of evaluating the level of mental workload in real time in a noninvasive manner.

**REFERENCES**

Ahstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics, 36*, 623–636. https://doi.org/10.1016/j.ergon.2006.04.002

Airbus. (2015). *Flying by numbers*. Leiden, Netherlands: Author.

Brennan, S. D. (1992). An experimental report on rating scale descriptor sets for the instantaneous self assessment (ISA) recorder (DRA Technical Memorandum (CAD5), 92017). Portsmouth: DRA Maritime Command and Control Division.

FLIR. (2012). *The ultimate infrared handbook for R&D professionals*. Retrieved from http://www1.flir.com/e/5392/research-development-guidebook/x7hz7/869543448

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, Netherlands: Elsevier Science.

Lehrer, P., Karavidas, M., Lu, S., Vaschillo, E., Vaschillo, B., & Cheng, A. (2010). Cardiac data increase association between self-report and both expert ratings of task load and task performance in flight simulator tasks: An exploratory study. *International Journal of Psychophysiology, 76*, 80–87. https://doi.org/10.1016/j.ijpsycho.2010.02.006

Murai, K., Hayashi, Y., Okazaki, T., & Stone, L. C. (2008). Evaluation of ship navigator’s mental workload using nasal...
temperature and heart rate variability. In 2008 IEEE International Conference on Systems, Man and Cybernetics (pp. 1528–1533). New York, NY: IEEE. https://doi.org/10.1109/ICSMC.2008.4811503

Ora, C. K. L., & Duffy, V. G. (2007). Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement. Occupational Ergonomics, 7, 83–94.

Parasuraman, R., & Mehta, R. (2015). Neuroergonomic methods for the evaluation of physical and cognitive work. In J. R. Wilson & S. Sharples (Eds.), Evaluation of human work (pp. 609–638). Boca Raton, FL: CRC Press.

Sensomotoric Instruments. (2011). iView X system manual. Teltow, Germany: Author.

Stemberger, J., Allison, R. S., & Schnell, T. (2010). Thermal imaging as a way to classify cognitive workload. In 2010 Canadian Conference on Computer and Robot Vision (pp. 231–238). https://doi.org/10.1109/CRV.2010.37

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2014). State of science: Mental workload in ergonomics. Ergonomics, 139, 1–17. https://doi.org/10.1080/00140139.2014.956151

Adrian Cornelius Marinescu is a PhD student at the University of Nottingham. He earned an MSc in alternative and renewable energy sources from the University of Bucharest, Faculty of Physics, in 2012.

Sarah Sharples is a professor of human factors at the University of Nottingham. She earned a PhD in human factors from the University of Nottingham in 1999.

Alastair Campbell Ritchie is an assistant professor at the University of Nottingham. He earned a PhD in bioengineering from the University of Strathclyde in 1997.

Tomas Sánchez Lopez is the research team leader for the information fusion team at Airbus Group Innovations. He earned a PhD in computer science from the Korean Advanced Institute of Science and Technology, South Korea, in 2008.

Michael McDowell is a researcher at Airbus Group Innovations. He earned an MEng in engineering mathematics from the University of Bristol in 2013.

Hervé P. Morvan is the Director of the Institute for Aerospace Technology at the University of Nottingham. He earned a PhD in computational fluid dynamics from Glasgow University in 2001.