ORIGINAL ARTICLE

Power law model for subjective mental workload and validation through air traffic control human-in-the-loop simulation

Norbert Fürstenau[1] · Thea Radüntz[2]

Abstract

We provide evidence for a power law relationship between the subjective one-dimensional Instantaneous Self Assessment workload measure (five-level ISA-WL scale) and the radio communication of air traffic controllers (ATCOs) as an objective task load variable. It corresponds to Stevens' classical psychophysics relationship between physical stimulus and subjective response, with characteristic power law exponent γ of the order of 1. The theoretical model was validated in a human-in-the loop air traffic control simulation experiment with traffic flow as environmental stimulus that correlates positively with ATCOs frequency and duration of radio calls (task load, RC-TL) and their reported ISA-WL. The theoretical predictions together with nonlinear regression-based model parameter estimates expand previously published results that quantified the formal logistic relationship between the subjective ISA measure and simulated air traffic flow (Fürstenau et al. in Theor Issues Ergon Sci 21(6): 684–708, 2020). The present analysis refers to a psychophysics approach to mental workload suggested by (Gopher and Braune in Hum Factors 26(5): 519–532, 1984) that was recently used by (Bachelder and Godfroy-Cooper in Pilot workload esimation: synthesis of spectral requirements analysis and Weber's law, SCL Tech, San Diego, 2019) for pilot workload estimation, with a corresponding power law exponent in the typical range of Stevens' exponents. Based on the hypothesis of cognitive resource limitation, we derived the power law by combination of the two logistic models for ISA-WL and communication TL characteristics, respectively. Despite large inter-individual variance, the theoretically predicted logistic and power law parameter values exhibit surprisingly close agreement with the regression-based estimates (for averages across participants). Significant differences between logistic ISA-WL and RC-TL scaling parameters and the corresponding Stevens exponents as ratio of these parameters quantify the TL/WL dissociation with regard to traffic flow. The sensitivity with regard to work conditions of the logistic WL-scaling parameter as well as the power law exponent was revealed by traffic scenarios with a non-nominal event: WL sensitivity increased significantly for traffic flow larger than a critical value. Initial analysis of a simultaneously measured new neurophysiological (EEG) load index (dual frequency head maps, DFHM, (Radüntz in Front Physiol 8: 1–15, 2017)) provided evidence for the power law to be applicable to the DFHM load measure as well.

Keywords Mental workload · Cognitive resource limitation · Psychophysics power law · Instantaneous self assessment · Air traffic control simulation

1 Introduction

The concept of mental workload (WL) addresses the demand a task (task load TL) imposes on the operators limited cognitive resources (e.g. processing, memory; Wickens and Hollands 2000; Wickens 2002)). According to these authors, WL research may be viewed in the context of prediction (e.g. multi-task performance), WL assessment imposed by equipment, and WL subjectively experienced by operators. A review on WL modeling and prediction in the complex air traffic control (ATC) work system was provided in (Loft

✉ Norbert Fürstenau
  norbert.fuerstenau@dlr.de

[1] Institute of Flight Guidance, German Aerospace Center (DLR), Lilienthalplatz 7, 38108 Braunschweig, Germany

[2] Unit Mental Health and Cognitive Capacity, Federal Institute for Occupational Safety and Health, Berlin, Germany

et al. 2007). They took into account changing task priorities for WL management and strategies of operators, and emphasized traffic density as indirect WL predictor due to task demands such as identifying, monitoring, and instructing aircraft (AC) via radio communication. Corver et al. showed (Corver et al. 2016) that traffic conflict, moderated by trajectory uncertainty, mediates the positive effect of traffic density on WL. The specific question of ATC complexity as WL driver was investigated, e.g. in (Djokic et al. 2010) who confirmed subjective WL to correlate strongly with traffic count and ATCO's communication load. In the present work, we used a simulated approach sector ATC work environment of a medium size German airport (with reduced risk of separation conflict), with variable traffic flow and online radio communication between ATCOs and (pseudo) pilots to provide evidence for the potential of the psychophysics approach (e.g. (Stevens 1975)) to workload for deriving quantitative WL-sensitivity parameters.

For this purpose, we validate the theoretically derived power law relationship between the quasi real-time one-dimensional subjective Instantaneous Self Assessment workload measure (ISA-WL) (Kirwan et al. 1997; Jordan 1992; Brennan 1992; Tattersall and Foord 1996) and the objective communication task load variable (frequency of ATCO's radio calls, RC-TL) as mediator between WL and environmental traffic load, by means of a human-in-the-loop (HitL) ATC-simulation experiment (Mühlhausen et al. 2018). Thereby, we formally combine the logistic RC-TL model with the recently published logistic WL model that was used for the analysis of the subjective ratings of operators (Fürstenau et al. 2020). In that previous work, we recorded during execution of the simulated ATC task, the periodic reporting of the subjectively experienced WL level as dependent on environmental load variable traffic flow n (aircraft per hour, AC/h), by means of the online five-level ISA questionnaire.

The ISA-WL and radio calls RC-TL data used for the present work represent only part of the complete set of subjective and objective WL measures (including expert ratings, NASA-TLX, cardiovascular (heart rate and HR variation), neurophysiological (EEG); for details see Sect. 2) that were registered online during the experiment and which required the least pre-processing effort for the analysis. The HitL ATC-simulation experiment was performed with a homogenous sample of experienced domain experts (ATCos) who also provided prior information on realistic traffic parameters for the selected airport approach sector.

The experiment within a realistic ATC approach radar and radio communication (between ATCos and (pseudo) pilots) work environment was primarily designed to validate the new robust neurophysiological real-time method of Dual Frequency Headmaps, (DFHM) for quantifying mental workload by means of the electroencephalogram (EEG)

(Radüntz 2017) (see Sect. 2.4). Initial ANOVA-based data analysis was published recently (Radüntz et al. 2019). The successful use in that work of the logistic ISA(n) model for validation of the objective EEG-based DFHM index (Radüntz et al. 2020a, b) provided the motivation for investigating in more detail nonlinear correlations between different WL and TL measures of our experimental data. Logistic dependencies of subjective workload on traffic count were reported before by (Lee 2005) who obtained significant fit parameters from ATC-simulation WL data with the seven-level ATWIT scale (Air Traffic WL Input (Stein 1985), see Sect. 2.2). A logistic model (comparable to our ISA(n) characteristic) was used also by (Averty et al. 2008) for the analysis of air traffic controllers decision-making in conflict risk detection.

One advantage of HitL simulations with highly trained domain experts is the online monitoring of different real-time data such as traffic flow and communication times and duration as environmental and TL variables, respectively, to be used as independent physical stimuli for subjective response within the psychophysics approach to WL. Moreover, a minimization of inter-individual variance is achieved through a homogenous sample of highly trained participants (Abich et al. 2013; Brookings et al. 1996).

Basic assumption for the derivation of our theoretical model was the cognitive resource or capacity limitation hypothesis (Kahnemann 1973; Wickens and Hollands 2000). All the above-mentioned subjective and objective-dependent WL and TL measures were correlated with traffic flow n (AC/h) as independent external load variable under nominal and non-nominal conditions (priority event $e = 0$, 1, two factor design). The derived power law in the present work corresponds to the classical stimulus–response relationship of Stevens (e.g., Stevens 1975; Link 1992). As proposed originally by Gopher et al. (1984, 1985), and as recently reported by Bachelder et al. (2019) the psychophysics approach suggests the power law application also to the relationship between objective task load as stimulus and subjective workload measures as response. In fact, Lehrer suggested in (Lehrer et al. 2010) the combined use of different measures due to well-known large inter-individual differences in sensitivities, because "it is known that some individuals respond more sensitively to task load changes in self-report measures, others in specific physiological measures". In the present context, the power law allows to predict theoretically and to estimate through (nonlinear) regression of experimental data the characteristic exponent that relates subjective WL (as response) to objective TL measures (as physical stimulus).

In the present work, the psychophysical power law is derived through combination of the logistic functions ISA(n) as dependent WL measure (subjective response) and RC(n) as communication TL load variable (physical stimulus; see

Sect. 4 and Appendix 2) which are interrelated through the independent environmental traffic flow variable n. The magnitude of the theoretically derived power law exponent γ is shown to correspond to the order of magnitude (≈1) of (Stevens') slope values of the generalized linear (log–log transformed) representation of the (subjective) response vs. (physical) stimulus. They were shown to be characteristic for a large number of sensory modalities (e.g., Stevens 1957). A theoretical basis for the psychophysical laws was derived by (Link 1992) (see Sect. 2.5). An information theoretic approach was provided by Norwich (e.g., Norwich and Wong 1997).

In what follows, we continue in Sect. 2 with a brief overview on aspects of different WL measures relevant for the present work. We introduce our study design in Sect. 3 and describe in Sect. 4 the theoretical background of our mental workload model with logistic and power law WL/TL characteristics for parameter prediction and regression-based parameter estimates. In Sect. 5, we present our experimental results which are discussed in Sect. 6 with regard to the theoretical predictions. Finally, in Sect. 7, we draw conclusions and outline further research. In Appendix A1, we provide tables with the detailed experimental (pre-processed) data for each participant, separated for experimental (traffic) scenarios, followed by A2 with mathematical details for the derivation of the theoretical model equations.

2 Mental workload and measures for (quasi) real-time applications

For the discussion of our results in Sect. 6, we will briefly address some aspects of mental workload, different subjective and objective WL measures which are relevant for the present work with focus on online (real-time) capabilities (ISA, ATWIT/WAK, SWAT, HR/HRV, EEG-DFHM), and the psychophysics (power law) approach.

2.1 Mental workload

Quantification of mental workload constitutes one of the main issues in cognitive ergonomics and human-factors research. Like many concepts in psychology, there is no singular agreed-upon definition or method for measuring mental workload. Much more, it is assumed that successful performance on a task or test requires cognitive resources, which can be seen as mental workload. In other words, mental workload is a theoretical construct referred to as "the cost incurred by the human operator to achieve a particular level of performance (Hart and Staveland 1988). Similar definitions were given by (Kahnemann 1973; Wickens and Hollands 2000), and (Xie and Salvendy 2000). Nevertheless, its quantification contributes to the evaluation of

human–machine systems, estimation of the appropriateness of automation levels, and enhancement of interface design. A good overview on different theoretical and practical aspects of workload with focus on transportation as our major field of interest is given in (Hancock and Desmond 2001).

As mentioned above, for measuring mental workload there are several methods available that can be categorized in two groups: objective and subjective methods. Objective methods rely upon quantification of performance or biophysiological data while the subjective methods consider the subjective rating given by the performer. Although all measurement methods aim to describe the relation between task demands and subject's ability to cope with them, several investigations reported dissociations among methods' results. A possible explanation might be that mental workload is a multidimensional concept that cannot be captured in all its facets by a single method. Apart from the task requirements, mental workload variations are caused by individual characteristics such as habituation, actual precondition, and coping styles (ISO-10075, 1991, 1996, 2004).

2.2 Subjective quasi real-time measures

Several researchers suggested that the subjectively experienced workload is of particular importance when evaluating subject's state (Yeh and Wickens 1984; Sheridan 1980). Johannsen et al. (1979) stated that "if an operator feels effortful and loaded, he is effortful and loaded". The most accepted subjective measure in ATC appears to be the multidimensional NASA task load index (TLX) based on questionnaires for capturing the different aspects constituting the experienced WL (Hart and Staveland 1988). NASA-TLX data together with expert ratings and ISA self reports (see below) were evaluated in a preliminary analysis of the present experiment to study the WL effect of a non-nominal event (Radüntz et al. 2019, see Sect. 3.2) during the HitL simulations. The main advantages of subjective methods are the relatively low data acquisition effort and the high user acceptance. Their main drawback is that they suffer from subjective distortion. They are influenced by memory lapses as the experienced workload took place at some time in the past (NASA-TLX) and they are subject to social desirability bias (Lehrer et al. 2010; Radüntz 2017). The questionnaire's items may not be readily understood or participants may lack the ability to introspect. What is more, they do not allow for fine-grained temporal sampling on the time scale of seconds and can alter the current workload state (Radüntz 2017).

In the present work, our interest was focused on the combination of the objective online communication TL measure with a subjective WL measure appropriate for (quasi) real-time data analysis of the simulator experiments. An early subjective quasi real-time WL-assessment technique was

introduced by (Stein 1985): the Air Traffic Workload Input Technique (ATWIT) using the seven-level WL Assessment Keypad (WAK). Lee et al. in (2005) reported on analysis of ATC-simulation ATWIT-WL data with nonlinear (sigmoid) dependency of WL on traffic count (see below and Sect. 6.1).

Two established subjective self-report measures suitable for near real-time application are SWAT (Subjective Workload Assessment Technique) and the above-mentioned one-dimensional ISA method. SWAT measures the three load dimensions, time, effort, and stress, each with three levels (Reid et al. 1989), while ISA monitors the experienced WL on a one-dimensional five-level scale via online self reports in fixed time intervals of a couple of minutes. In contrast to SWAT, it minimizes possible additional WL (due to the reporting) by not discriminating load dimensions (Brennan 1992; Jordan 1992; Kirwan et al. 1997; Tattersall and Foord 1996). The latter authors reported significant correlations of ISA ratings with cardiovascular HRV and task performance, although the primary task performance on a tracking task turned out poorer during periods when ISA responses were required. Of course this distortion certainly depends on the details of task and reporting method (verbal, keypad, touchscreen). Girard et al. adapted online ISA to a professional car driving simulator and reported significant correlation of ISA-WL with dynamic traffic density variation (Girard et al. 2005). The characterization of the reported subjective load levels is listed in Table 1:

Because the scale levels represent the subjective decision of participants on the experienced load during task execution the level differences may not be assumed to be equidistant. In the theoretical model of Sect. 4.1, we assume an equidistant ISA scale so that any deviation from linearity is included in the nonlinearities of the model equations. In a recent publication, we provided evidence for the logistic dependence of ISA-WL on the environmental traffic load variable n and derived a linearized ISA-WL-sensitivity index for subject clustering (Fürstenau et al. 2020). The subjective index was successfully applied for the validation of the neurophysiological DFHM WL index (Sect. 2.4; Radüntz et al. 2020a, b). Our logistic model-based data analysis agreed with results of Lee et al. (2005) and Lee (2005) based on ATC simulation with dynamic traffic variation. They reported on results of

logistic WL data fits based on a heuristic sigmoid function dependent on aircraft count within en-route sectors, with significant four-parameter estimates of ATWIT-based subjective WL measurements using the seven-level scale of the WL Assessment Keypad (WAK, see above (Stein 1985)).

2.3 Psycho-physiological measures heart rate and hr-variation

The analysis of bio-signals as objective measures (see also Sect. 2.4) offers the possibility to continuously determine mental workload. They do not interfere with participant's current workload state as they can be obtained on-the-fly during task execution. Their main issue is that user acceptance may be impaired because of the complexity of the registration system. However, recent developments in mobile sensor technology promise small, lightweight, and wireless systems (Radüntz 2017). Bio-physiological data include, among others, cardiovascular biomarkers which are easy to assess and were frequently used to analyse cardiovascular activity under a wide range of experimental conditions (Karavidas et al. 2006; Lehrer et al. 2010). The heart rate (HR) and the heart rate variability (HRV) are the most prominent biomarkers. Recently, Vanderhaegen et al (2020) reported on an experiment that showed synchronization between dynamic events with heart beats and its impact on non-conscious errors in control.

In most cases, HRV is characterized in the frequency domain by means of various spectral features. According to the definitions by (Mulder et al. 2004), the frequency range can be categorized in three bands: the low-frequency (LF: 0.02–0.06 Hz), mid-frequency (MF: 0.07–0.14 Hz), and high-frequency (HF: 0.15–0.4 Hz) bands. It was observed that under mental load the total spectral power decreased, whereby the spectral power between 0.02 and 0.20 Hz was particularly affected and contributed about 80% to the total spectral energy (Mulder and Mulder 1981).

Basic research on HRV as WL measure for adaptive automation was investigated by (Prinzel et al. 2003) with a tracking task, together with EEG (see Sect. 2.4) and event-related potentials. Lehrer et al. (Lehrer et al. 2010) reported an increase of association between self-report

Table 1 ISA workload categories after (Kirwan et al. 1997)

Level	WL Heading	Spare Mental Resources	Description
5	Excessive	None	Behind on task; loosing track of the full picture
4	High	Very Little	None essential tasks suffering. Could not work at this level very long
3	Comfortable busy pace	Some	All tasks well at hand. Busy but stimulating pace. Could keep going continuously at this level
2	Relaxed	Ample	More than enough time for all tasks. Active on tasks less than 50% of the time available
1	Underutilized	Very Much	Nothing to do. Rather boring

scale (using NASA-TLX) given immediately after each 5 min task and both expert ratings of task load and task performance in a flight simulator by means of cardiac data. We recently reported on analysis of HR and HRV measures within the present simulator experiment where we aimed at clarifying their inherent timescales (Radüntz et al. 2020a, b).

2.4 Neurophysiological (EEG-based) measures

The spectral power of EEG oscillations in different frequency bands (specifically α (4–7 Hz), β (8–13 Hz), θ (14–30 Hz) may be linked to different levels of workload by means of analysis of variance (ANOVA) (e.g. Lei and Roetting 2011; Aricó et al 2018). The potential of an EEG-based task engagement-index (based on the power ratio $\beta/(\alpha+\theta)$ recorded from four scalp sites, 40 s moving average, 2 s clock rate) within the context of adaptive automation was demonstrated by Prinzel et al. by means of a laboratory type multi-attribute cockpit-instrument tracking-task simulator experiment, using ANOVA for quantifying the significance of the engagement level (Prinzel et al. 2003). The important artifact rejection was based on a pre-set threshold voltage which for real-world applications of course would not be sufficient. Meanwhile, classifiers are increasingly used for the separation of workload levels. In previous publications, we have described the development and validation of the new DFHM WL index using a support vector machine classifier (based on frontal α-band and parietal θ-band powers), performed under laboratory conditions with standard task load batteries. Once calibrated for discriminating low, medium, and high WL levels, it was shown to require no retraining of the machine learning algorithm, neither for new subjects nor for new tasks (Radüntz 2016, 2017). For the present experiment, we used a commercial 25 active-electrode system (g.tec Ladybird) with 500 Hz sample rate and 0.5–50 Hz bandpass. The corresponding data from the present model-based data analysis showed the objective DFHM index to provide significant correlation with controller's subjectively experienced self rating ISA-WL measure under traffic load variation (Radüntz et al. 2020a, b). For testing the DFHM-WL index sensitivity, the participants in this analysis were separated into two groups (low and high WL sensitivity) according to their individual linearized WL-sensitivity parameters that were formally derived from the logistic ISA characteristic of the subjective self-report measures. Fürstenau et al. (2020). In Sects. 6.2, 6.3, we briefly address the potential of extending the resource limitation-based logistic and power law model approach to the new DFHM-WL index measure, by means of regression-based parameter estimates.

2.5 Psychophysics of Mental Workload

Despite the fact that subjective WL measures are widely accepted and used, there have been very few studies examining their methodological viewpoint. Based on laboratory experiments with standardized cognitive tasks (Gopher and Braune 1984; Gopher et al. 1985) proposed a scaling approach that can be traced back to the psychophysical measurement theory of Stevens (1975). Psychophysical research aims to describe the relationship between changes in the amplitude of a physical stimulus (e.g. brightness, loudness) and the subjective perception of these variations.

The classical Weber–Fechner law assumes a logarithmic relation between physical stimulus S and subjective perception $P = c \ln \left(S/S_t \right)$, with an experimentally determined constant c and a stimulus threshold S_t that denotes the intensity of the stimulus at a state with no perception (Buntain 2012). An improvement was introduced by Stevens (1975). In Stevens' law, the sensation magnitude is a power function of stimulus intensity and the corresponding generalized linear curve (double logarithmic scale) is described by the constant b and Steven's exponent γ (slope or sensitivity in log–log scale) that is characteristic for the type of stimulus.

$$\ln \left(P \right) = \ln \left(b \right) + \gamma \ln \left(S \right) \tag{1}$$

It is valid also for the stimulus–response transfer between sensor input (stimulus amplitude) and sensor neurons firing rate (action potential) (e.g. Birbaumer and Schmidt 2010). The power law exponent γ with a typically magnitude of the order of 1 was determined for a large number of different modalities (e.g. brightness, loudness, apparent length) to adjust the curve to the different psychophysical functions. Steven's law was derived from an information theoretic approach with P~perceived sensory (Shannon) information by Norwich et al. (1987), Norwich and Wong (1997). Within this context, it represents an approximation for lower amplitude stimuli with prolonged sampling time, while the Fechner law represents an approximation for the large amplitude brief stimulus duration. With regard to workload, Gopher et al. in (1984) argued that, … "if the human information processing system can be assumed to invest … hypothetical processing facilities to enable the performance of tasks then subjective measures can be thought to represent the perceived magnitude of this investment, in much the same way that the perception of …" a physical stimulus is changed with variation of its magnitude. Gopher et al. based their formal power law relationship on the measured average values across the sample of 55 participants of perceived load for each of 21 single and dual-task conditions of a task load battery, with tasks guided by Wickens' multiple resources paradigm (e.g. Wickens and Hollands 2000). In contrast to a standard psychophysical (stimulus–response) experiment, in their WL experiment, there existed no a priori physical quantity (e.g. brightness or sound

pressure) that induced the subjective judgement, and that would allow to derive the two parameters (γ, b) of the power law function (1) through regression analysis. Instead, they derived a physical stimulus scale by means of the amount of (Shannon) information attributed to the task load battery.

Recently, Bachelder and Godfroy-Cooper (2019) reported on the application of the psychophysics power law to the analysis of a pilot workload estimation simulator experiment. They used a flight compensatory tracking task with Bedford hierarchical unidimensional WL scale (a modified Cooper-Harper rating scale) designed to identify operators spare mental capacity, while completing the task. The physical stimulus S determining WL in Eq. (1) was derived from the measured standard deviations of control error rates. Theoretically predicted Stevens exponents of different tasks were in the range $0.24 \leq \gamma \leq 0.41$ and compared favorably with those obtained from regressions of the data using the power law (1): $0.21 \leq \gamma \leq 0.37$, i.e. the order of magnitude was comparable with those of the classical psychophysics experiments.

A basic theoretical foundation for the power law was provided by Link with the stochastic brain wave discrimination theory (Link 1992) that allows for formal derivation of psychophysical laws. Starting point was the probability for reaching a decision threshold through random sampling of the difference between stimulus and referent waves that defined a logistic response function with exponential dependence on wave amplitude difference and threshold. Stevens' power law was derived from sensation matching by combining the two corresponding logistic functions. The ratio of two normalized subjective response thresholds A_S/A_P relate two simultaneously measured sensations with logistic response probability functions. The product of this ratio with the log(normalized sensation of physical stimulus S/S_0) equals the log(normalized subjective response P/P_0) in the generalized linear form of Stevens law (Eq. (1); $S_0 =$ standard stimulus). Based on the cognitive resource limitation hypothesis as our theoretical starting point (Sect. 4), we use a comparable formal procedure for the derivation of the ISA(RC) power law, however, in the present approach through combination of the discrete logistic ISA(n)-WL response and objective RC(n) task load stimulus characteristics, with the variables assumed as statistical means from averages across a sufficiently large random sample of participants.

3 Experiment

Details of our experimental setup and procedures together with initial results were provided in previous publications on the validation of the new neurophysiological DFHM WL index, with different subjective and objective WL measures

as reference (see Sect. 2) (Mühlhausen et al. 2018; Radüntz et al. 2019). Here, we give a brief overview with details relevant for the validation of the power law WL index only, based on the combination of ISA-workload data and the ATCo's frequency of radio calls with pilots (RC, calls / h). The experiment was designed within a collaboration between the Federal Institute for Occupational Safety and Health (BAuA) in Berlin and the Institute of Flight Guidance of the German Aerospace Center (DLR) in Braunschweig. Simulation experiments with data acquisition were performed at the Air Traffic Management and Operations Simulator (ATMOS) of the DLR. The investigation was approved by the local review board of the BAuA and all procedures were carried out with the adequate understanding and written consent of the participants.

3.1 Procedure and subjects

Every subject completed eight simulation scenarios in randomized order within two consecutive half days and communicated online with pseudo-pilots who simulated the cockpit crews, each one responsible for several aircraft (AC).

Our sample consisted of 13 approach controllers, 3 tower controllers, and 5 employees of the DLR that exhibited adequate expertise to handle the arrival management simulation and interact with the pseudo-pilots. In total, we had $N = 21$ subjects between the ages of 22 and 64 years (2 female, 19 male, mean age 38 ± 11) with different work experience who came from different airports and were familiar with different work positions.

3.2 Experimental design and workload assessment

The experiment was conducted for investigation of workload effects under different task-load levels j = 1,…,8 in a standard approach sector radar work environment. The load levels were realized through four different traffic flow conditions n_j (25, 35, 45, and 55 aircraft AC/h) and a dichotomous priority-flight request event $e = 0, 1$. The combination of both independent variables led to eight simulation scenarios (8 scenarios: j = 1, …, 4 without event e = 0 and j = 5, …, 8 with priority event e = 1).

Radio communication between ATCo's and pilots represents a major contribution to the total task load, besides monitoring the traffic on the radar display (traffic count n) for anticipating possible separation conflicts (Manning et al. 2001; Averty et al. 2004; Djokic et al. 2010; Corver et al. 2016). Because communication (task) load under nominal conditions increases with traffic count, it seemed appropriate to use a one-dimensional WL measure for the experiment. Generation of traffic was realized by means of well-trained pseudo-pilots in a separate room with computer systems for controlling the simulated pre-defined air

traffic according to the clearances via the simulated radio connection to the ATCo at the approach radar work place. Registration of the start and stop times of ATCo's radio calls provided time series that allowed to derive for the eight scenarios the average radio call duration (RD/seconds) and moving averages as well as the average across the whole scenario of the frequency of radio calls (RC / calls per hour).

Participants periodically judged their subjectively experienced WL in fixed time intervals of 5 min by means of the Instantaneous Self Assessment (ISA) self-report method (Brennan 1992; Jordan 1992; Kirwan et al. 1997). Their judgement based on an one-dimensional five-level integer scale with values corresponding to (1) under-utilized, (2) relaxed, (3) comfortable, (4) high, and (5) excessive (for details see Sect. 2.2). The realization by means of a touchscreen for selecting the experienced scale level allowed for minimum distortion (Tattersall and Foord 1996, Sect. 2.2). According to prior information from experts familiar with the selected approach sector, for $n \leq n_1 = 25$ AC/h subjects were expected to experience low load, while $n_2 < n \leq n_3 = 45$ AC/h was the standard operating range with $n_3 = n_c =$ sector capacity) with high load. n_c as prior knowledge was also derived theoretically from the average separation minimum of given traffic mix (3.1 nm/AC) and average approach speed of 140 kts. The highest traffic flow ($n_4 = 55$ AC/h) exceeded the realistic maximum traffic n_c and served for driving the load over the acceptable limit according to experts comments.

Scenarios without priority event ($e = 0$) had a duration of 20 min with four ISA reports, whereas scenarios including the priority event "sick passenger on board" ($e = 1$) at simulation time $t_S = 10$ min took 25 min and contained five ISA reports. For the theoretical modeling and data analysis, we used as dependent variables the scenario means $<ISA>(\underline{n}_j)$ and $<RC>(n_j)$ calculated over the whole time series as WL-rating and TL-value estimate, respectively, for each participant in the eight scenarios. Tables with pre-processed raw data and results for individual participants are provided in our previous publication (Fürstenau et al. 2020) and in Appendix 1 for completeness. In what follows, we restrict the theoretical predictions, regression analysis and discussion to the means across the 21 subjects.

4 Theory

In this section, we derive a theoretical psychophysical power law ISA(RC) with exponent γ from the parametric representation of communication load RC(n), and WL self-report ISA(n), with asymptotic upper limits ISA$_u$, RC$_u$ as prior information. With suitable normalization and transformations (S(RC), P(ISA)) into a generalized linear relationship

$y_P \sim \gamma \, y_S$ we obtain a formal equivalence to Stevens' law (Eq. (1)), with $y_S \sim \ln(S)$, $y_P \sim \ln(P)$.

Starting point for our theoretical model was the assumption of cognitive resource limitation (Kahnemann 1973; Wickens 2002). The dynamics of growth of a population or magnitude of a corresponding continuous variable that increases with time t through consumption of a limited resource may be formalized through the Verhulst differential equation with the logistic (sigmoid) function as solution (see Appendix A2). By replacing the usual time variable by the independent environmental traffic load variable n, we used this function as theoretical model for the characteristics of the measured averages $<<>>$ across the participant sample of the scenario means of subjective $<ISA>(n)$ and rate of radio calls $<RC>(n)$ [calls/h]. The reported subjective value ISA(n) WL level is assumed to measure the fraction of limited overall cognitive resources (attention, processing, memory) required for the specific task RC(n). In what follows, we will use I(n), R(n) where appropriate.

4.1 Logistic ISA(n) model

The logistic resource limitation approach for prediction and regression-based estimates of ISA(n) model parameters was used for deriving a linearized WL-sensitivity index in Fürstenau et al. (2020). It allowed for subject clustering within the neurophysiological DFHM index validation (Sect. 2.4, Radüntz et al. 2020a, b). A comparable logistic model approach was used also by Lee et al. for analysis of ATC-simulation WL data using the 7-level ATWIT method (Lee 2005) and by (Averty et al. 2008) for formalizing ATCo's decision analysis in the context of collision risk judgement. Main feature is the asymptotic approach to an upper WL boundary.

For analyzing the measured ISA data, we used prior information for the detailed design of the logistic workload characteristic to be fitted to the experimental data (see Sects. 2.2, 3.2). On one hand, prior knowledge concerns the selected traffic flow range $25 \leq n \leq 55$ (AC/h) to be handled by the controllers and on the other hand, the ISA scale. The latter by definition is limited to the range between ISA: $= I_d = 1$ and $I_u = 5$ with five integer values $1 \leq I(n) \leq 5$. In the most simple approach, this leads to the assumption of constant minimum and maximum ISA levels of ISA(n): $= I(n) = I_d = 1$ for $0 \leq n \leq 25 =$ underload, and $I_u = 5$ for $n \geq 55 =$ excessive load. If a linear increase is assumed in between, with slope $a \leq (5-1)/(55-25) = 0.13$ (AC/h)$^{-1}$, this yields as intersection $I(n=0) = 1-25\,a = -2.33$. In reality, an idealized linear $I(n \mid a, b)$ characteristic would be different for different individuals because of inter-individual variation of task load sensitivity and transition to underload and overload (see Fürstenau et al. 2020). Consequently, a random sample of participants would generate distributions with density functions for slope

a (sensitivity > 0) and intersection b (both negative and positive values possible). By assuming a variable $n_0 \leq n_1 = 25$ for the underload transition, we get $b \geq -2.33$, and with $a > 0$, $n_0 > 0$: $-2.33 \leq b < 1$.

For a more realistic model of the average ISA ratings, we refer to the above-mentioned standard formalism for resource limited growth and assume asymptotic convergence of $\lim I(n >> n_4 = 55) = I_u$ to be modeled by a logistic (sigmoid) function:

$$I(n) = \frac{I_u}{1 + exp\left\{-\frac{n-\mu}{\nu}\right\}} = \frac{5}{1 + kexp\left\{-\frac{n}{\nu}\right\}} \quad (2)$$

With shift parameter $\mu = \nu \ln(k)$, $k = I_u/I_d - 1$ and scaling coefficient ν for the convergence towards the upper and lower asymptote. ν also characterizes as sensitivity index the maximum slope $I' = dI/dn = I_u/4\nu$ at inversion point $n = \mu$ with $I(\mu) = I_u/2$). For the nominal traffic ($e = 0$), we have $k = 4$ and $\mu = \ln(4)\nu$ (for mathematical details see Appendix A2). As initial guess, we select for $e = 0$ the shift parameter value $\mu := \mu_t := 35$ AC/h, because according to domain experts a priori information, it corresponds to the center between underload $n_1 = 25$ and sector capacity limit $n_3 = n_c = 45$, representing the optimum (nearly linear) operational range for the given conditions, sufficiently far away from the nonlinear sections (see Fig. 1). A reasonable uncertainty value may be selected as $|\delta\mu_t| = 5$, i.e. half the distance to the boundaries. As shown in Fig. 1, the characteristic features for the nominal case ($e = 0$, solid curve) are the predicted effective ISA range between approximately 2 and 3.5 and the only weak nonlinearity for the given load variable range $25 \leq n \leq 55$ AC/h, with slope value $I'(n = \mu := 35) \approx I_u/4\nu = 0.0495$ (AC/h)$^{-1}$ (i.e. significantly smaller than the initial rough estimate) For comparison with other WL measures and derivation of the power law, we define the normalized ISA metric $p_I = I(n)/I_u$ through division by the upper asymptotic value I_u. Via definition of the transformed ISA variable $P = p/(1-p) = I(n) / (I_u - I(n))$, we arrive at the exponential dependence $P(n) = 1/k \exp(n/\nu)$. Taking the logarithm transforms this exponential characteristic into the generalized linear model $y(n) = a_g n + b_g$ with parameters $a_g = 1/\nu$, $b_g = -\ln(k)$:

$$y_p(I(n)) = ln[P] = \frac{n-\mu}{\nu} = \frac{1}{\nu}n - ln(k) \quad (3)$$

For the nominal case ($e = 0$) with $\mu_t = 35$ AC/h ($= n_2$, operational traffic), the theoretically predicted slope value (WL sensitivity) is obtained as $a_{gt} = 1/\nu_t = \ln(4) / \mu_t = 0.0396$ (AC/h)$^{-1}$ or $\mu_t = 25.25$ AC/h, and intersection $b_{gt} = -\mu_t/\nu_t = -\ln(4) = -1.3863$.
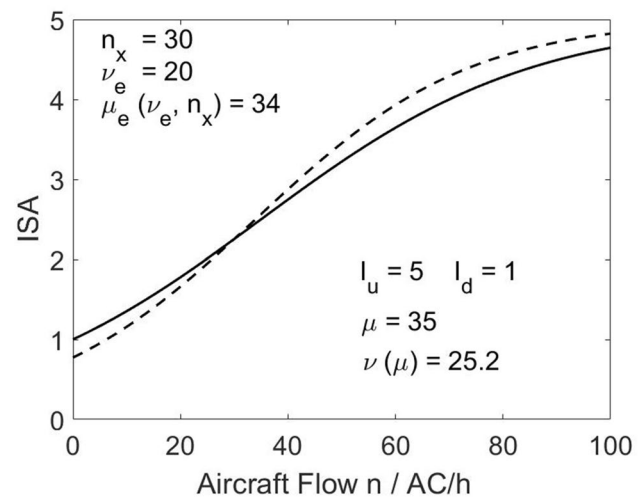

Fig. 1 Theoretical ISA(n) characteristics (Eq. (2)) for nominal (solid curve, $e = 0$: $\mu = 35$, $\nu(\mu) = 25.2$) and non-nominal scenarios (dashed line, priority event, e = 1: $\nu_e = 20$, $n_x := 30$, $\mu_e (\nu_e, n_x) = 33.9$). Intersection point $(n_x, I_x) = (30, 2.3)$. Abscissa: independent traffic load variable $0 \leq n \leq 100$ / AC/h. Ordinate: ISA-WL with ISA$(n = 0) := I_d = 1$ for e = 0, asymptotic limit $I(n) = I_u = 5$ for $n >> n_c = 45$; for details see text and Appendix 2

We expect any effect of the priority request in simulation runs $j = 5$–8 to generate an increase of slope of $y_p(n, e = 1)$ from the nominal value ($1/\nu_e > 1/\nu$ or $\nu_e < \nu$), however, only for traffic load larger than a threshold value n_x, i.e. $n \geq n_x >$ underload traffic n_1. This generates an intersection between the $e = 0$ and $e = 1$ sigmoids at n_x defining a critical threshold for onset of the priority effect (bifurcation of $e = 0$ into separate $e = 0$, $e = 1$ characteristics for $n_x > n_1$, with $I_x > I_1$ and $I(e = 1) > I(e = 0)$ for $n > n_x$).

Basically, for the non-nominal ($e = 1$) simulations, parameter estimates (μ_e, ν_e) have to be determined by two-parameter (k_e, ν_e) regression of the experimental data using model Eqs. (2) or (3) due to lack of prior knowledge on the magnitude of the WL effect of the priority event (in contrast to $e = 0$). However, a one-parameter model (like for $e = 0$) may be derived by means of a plausibility argument (prior knowledge) for the intersection coordinate (n_x, I_x) between $e = 0$ and 1 characteristics that in turn allows for deriving a relation between μ_e (or k_e) and ν_e: $\mu_e(\nu_e, n_x)$ or $k_e(\nu_e, n_x)$. For the non-nominal scenarios ($e = 1$), the shift parameter μ_e is derived as (for details see Appendix A2)

$$\mu_e = n_x\left(1 - \frac{\nu_e}{\nu}\right) + \nu_e ln(4) \quad (4)$$

A prior estimate of n_x may be obtained with reference to the multiple resources theory (Wickens 2002). The nominal traffic management task and the major part of additional decision-making due to priority request are both

perception–cognition–communication tasks. The additional task due to the priority request consists in checking for the possibility of a direct route to final approach depending on the traffic situation. Both nominal and priority flights require traffic monitoring on the approach radar display and communication with pilots, that use overlapping mental resources. However, we may argue that the task of excluding a potential separation conflict for the changed routing option generates additional WL only under higher traffic load ($n > n_x$), with n_x between underload and operational traffic ($n_1 = 25 < n_x < n_2 = 35$ AC/h). Only small additional mental resources and corresponding neglectable WL change is expected for $n < n_x$. Consequently, a plausible prior value is $n_x :\approx 30$ AC/h with (plausible) maximum n_x-uncertainty given by the n_2–n_1 interval: $\delta n_x := (n_2 - n_1)/2 = \pm 5$, yielding $\delta I_x = \pm 0.5$ (error propagation including the independent shift parameter ($\mu := 35$) uncertainty $\delta\mu := \pm 5$). These estimates based on domain experts prior knowledge allows for deriving a plausible a-priori estimate for the bifurcation point $\{n_x, I_x\} = \{30, 2.25\}$. Figure 1 depicts two theoretically predicted ISA(n) characteristics according to Eq. (2) for the extended traffic flow interval $0 \le n \le 100$ AC/h.

The solid line represents the nominal scenarios (e = 0) with $\mu := \mu_t = 35$ ($\nu(\mu_t) = 25.2$) ($\mu_t =$ inversion point, center of the nearly linear range between underload n_1 and $n_c = n_4$). The dashed sigmoid shows an example with increased WL sensitivity $1/\nu_e$ (e = 1: $\nu_e = 20$) with intersection at $n_x := 30$ and $\mu_e(\nu_e, n_x) = 33.9 < \mu = 35$, according to Eq. (4). For $n > n_x$, the sigmoid exhibits the predicted subjective WL increase for e = 1, whereas for $n < n_x$ (underload range), the priority scenarios are expected to follow the e = 0 curve (i.e. dashed continuation to be ignored). The simulated traffic range $25 \le n \le 55$ covers the nearly linear section of the sigmoid curves. This predicted quasi linearity was used in our previous ISA data analysis (Fürstenau et al. 2020) for derivation of a linearized WL-sensitivity index (see Appendix 2) that was successfully applied to the analysis of the simultaneously monitored neurophysiological DFHM index (see Sects. 2.4, 6.2, 6.3, Radüntz et al. 2020a, b) with regard to participant clustering.

4.2 Logistic RC(n) model

Assuming a nearly linear increase of radio communication between ATCo and pilots with traffic flow n for small RC (calls/h) (i.e. for small $n < n_1$ $R(n) \sim n$ with asymptotic approach to the maximum R_u), the logistic $R(n)$ characteristic is given by

$$R(n) = R_u \left[\frac{2}{1 + \exp\{-n/\rho\}} - 1 \right] \tag{5}$$

With $n/2\rho := x$, the normalized rate of radio calls $R(n)$ / $R_u := s(n)$ may be written in short as tanh(x) (for mathematical details see Appendix A2). It is easily verified that for $n >> n_c$ the dimensionless variable $R(n)/R_u := s(n) = 1$.

If we introduce as prior knowledge an estimate of average radio call duration of $T_D \approx 4$ s (see Sect. 5.2), an estimate for the asymptotic maximum number of calls per hour may be obtained by $R_u := 3600 / (T_D(\text{ATCo}) + T_D(\text{Pilot}) + T_D(\text{Pause})) \approx 3600 / (4 + 4 + 1) = 400$ calls / h. Taking $R_u := 400$ as prior knowledge, Eq. (5) turns into a one-parametric model. A rough theoretical estimate for the scaling parameter ρ may be obtained from a linear extrapolation of the maximum slope at $n = 0$ as $\Delta s /\Delta n = 1 / 2\rho_t$ yielding $\rho_t :\approx n_c/2 = 22.5$ (see Appendix 2, Eq. A2.11; with $\Delta s = 1$, and $\Delta n :=$ capacity limit $n_c = n_3$). The slope at the inversion point (linearized sensitivity) is predicted as $1 / 2\rho = 0.02 > 1 / 2\nu = 0.01$, i.e. larger than the WL sensitivity.

Through normalization and logarithmic transformation, the nonlinear characteristic (5) may be transformed into a generalized linear model, comparable to $y_p(n)$ (Eq. 3). With the normalized and transformed radio calls variable $S = (1 + s) / (1 - s) = (R_u + R(n)) / (R_u - R(n))$, we arrive at the exponential dependence $S(n) = \exp(n/\rho)$. Taking the logarithm transforms this exponential characteristic into the generalized linear form of the radio calls sigmoid characteristic $y_s(n) = \ln(S) = a_{sg} n + b_{sg}$ or

$$y_s(n) = ln(S) = \frac{1}{\rho}n \tag{6}$$

with slope $1/\rho := a_{sg}$ as RC task load sensitivity parameter and $b_{sg} = 0$ (see Appendix A2 for details). The choice of variable name S and index s indicates the usage of the transformed RC variable as physical stimulus for the (transformed) subjective ISA-WL variable P (for report of subjective perception of the physical stimulus) according to Eq. (1) (see following Sect. 4.3).

In contrast to the ISA(n, e) curves with prediction (for $n > n_x$) of the non-nominal scenarios ISA(n, e = 1) > ISA(n, e = 0), we may expect for RC(n) the inverse behavior: $R(n, e = 1) < R(n, e = 0)$. According to (Sperandio 1978) approach, controllers under (suddenly) increased traffic load (in our case the occurrence of a priority request as non-nominal event with increased task load) prefer switching of control strategy to standard procedures with global routing for most AC, i.e. global approach sequence with pilots responsible for controlling the standard separation distance. Consequently for ATCos, control of the first AC in the AC sequence will be sufficient, resulting in decreased RC(n) with corresponding decrease of ISA-WL, and attention resources free for focus on the priority event (see discussion in Sect. 6). Because our initial model assumption, RC($n = 0$)) = 0 should be true

for both $e = 0$ and 1 the intersection of both characteristics is predicted at $\{n_{sx}, R_x\} = \{0, 0\}$. Fig. 2 depicts predicted theoretical radio call rate (calls/h) characteristics for nominal traffic ($e = 0$) and scenarios with non-nominal event ($e = 1$ with somewhat decreased TL sensitivity, i.e. increased ρ_e, value selected as example):

4.3 Power law model for ISA(RC)

The power law for the ISA(RC) characteristic may be derived from the parametric representation [ISA(n), $R(n)$] by introducing $n(R)$ as obtained from Eq. (5), into Eq. (2) (for details see Appendix A2). Using prior information on upper (asymptotic) limits $I_u = 5$, $RC_u = 400$ calls/h and lower limits $I(n = 0) = I_d = 1$, $RC(n = 0) = 0$ the normalized nonlinear ISA(RC) characteristic $p(s)$ with $p = I/I_u$, $s = RC/R_u$ is obtained as a two-parametric model (γ, k) with $\gamma = \rho/\nu$ and $\mu/\nu = \ln(k) = -b_g$ (see EQ. (3))

$$p(s) = \frac{1}{1 + k \left[\frac{1-s}{1+s} \right]^\gamma} \tag{7}$$

With $k = 4$ for the nominal case ($e = 0$), this is reduced to a model with power γ as the single free parameter and a theoretical estimate obtained from the stimulus–response ratio $\gamma_t := \rho_t/\nu_t = 22.5/25.25 = 0.89$, i.e γ is predicted to be of the order 1 as usually observed for psychophysics power law exponents measured in classical stimulus–response experiments (e.g. (Stevens 1957; Link 1992; Bachelder and Godfroy-Cooper 2019) and references therein). As expected and shown in the following Fig. 3 for three examples ($\gamma = 0.8$,

1.0 and 1.2), all characteristics converge independently from the single parameter γ to $p = 0.2$ (ISA = 1) for $s = 0$ (RC = 0), and to $p = 1$ (ISA = 5) for s = 1 (RC = RC_u = 400).

Like for ISA(n) the nonlinear power law Eq. (7) may be transferred into a generalized linear relationship that is obtained after transformation of p and s into the dimensionless variables $P(p) = p / (1 - p)$ and $S(s) = (1 + s) / (1 - s)$, respectively, (for details see Appendix 2):

$$y_p = \ln(P) = \gamma \ln(S) - ln(k) = \gamma y_s + b_s \tag{8}$$

with $b_s = b_{gt} = -\ln(4)$ and γ defining the slope of the generalized linear (log–log) form $y_p(y_s)$ of Stevens law (Eq. 1) corresponding to $P = bS^\gamma$ in linear coordinates with $b := 1/k = \exp(-\mu/\nu)$.

For our experimental scenarios $j = 5 \dots 8$ with additional task load due to the non-nominal event $e = 1$, the intercept $b_s = b_{gt} = -\ln(4)$ of the $e = 0$ scenarios is replaced by the second free model parameter b_{se}. The unknown shift parameter μ_e ($< \mu$) of the generalized linear characteristic together with $\gamma_e > \gamma$ defines a two-parameter power law model with offset change $b_{se} < b_{gt} = -1.386$ and an intersection with the nominal characteristic at R_x or s_x, respectively (for details see e.g. Eq. (A2.16) in Appendix 2). Again, like for ISA(n), it appears plausible that for $e = 1$ additional task load leads to ISA(RC)-WL increase only for radio call frequencies RC > $R_x = R(n_x)$ (corresponding to $p > p_x$ and $y_p > y_{px}$ for $n > n_x$). R_x characterizes the communication underload threshold. Based on our prior numerical prediction of parameters n_x: = 30, $\rho_t \approx 22.5$, $\nu_t = 25.3$, $b_{gt} = -\ln(4)$, we may derive a rough theoretical prediction
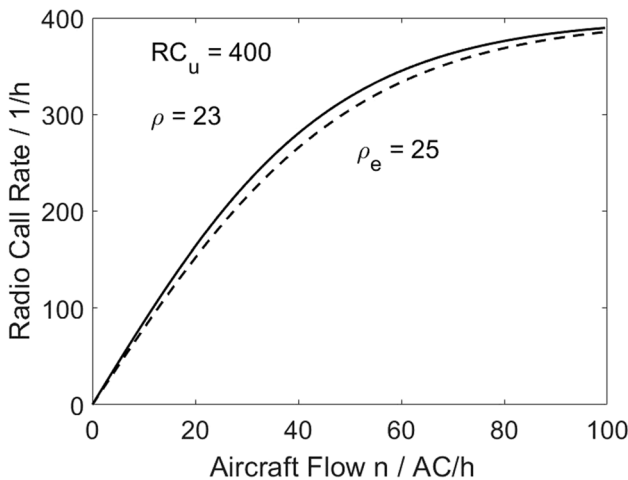

Fig. 2 Plot of theoretical radio calls rate $R(n)$ (Eq. 5) with sensitivity parameter $\rho = 23$ (solid line: nominal traffic $e = 0$) and $\rho_e = 25$ (dashed line: priority event, $e = 1$,). Maximum slope with linear increase at origin $R(n = n_x = 0) = 0$. Asymptotic limit of calls per hour for n > > n_c, R_u: = 400 as prior information (for details, see text and Appendix 2)
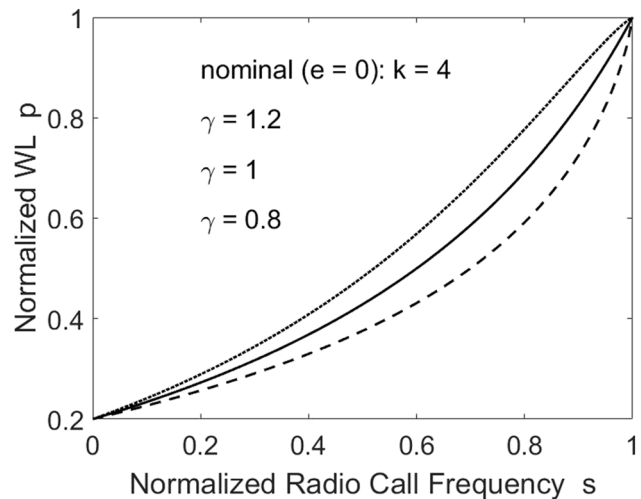

Fig. 3 Theoretical power law characteristics $p(s)$ for nominal case ($e = 0$: $k = 4$) with normalized variables using Eq. (7), with $\gamma = 0.8$, 1.0, 1.2, from top to down. Abscissa: normalized radio calls rate RC/R_u; ordinate: normalized WL ISA/I_u. For details see text and Appendix 2

for the hypothesized intersection coordinates of the generalized linear power law (8) of nominal ($e = 0$) and priority event ($e = 1$) characteristics, to be compared with the experimental results in Sect. 5.4 (for details see Appendix 2): $\{R_x, I_x\} = \{R_u \tanh(n_x/2\rho), I_u/(1 + 4\exp[-n_x/\nu])\}$ or $\{y_{sx}, y_{px}\} \approx \{n_x/\rho_t, n_x/\nu_t - \ln(4)\} = \{1.3, -0.2\}$.

Like for non-nominal ($e = 1$) scenarios of ISA(n) in Sect. 4.1, we can also derive for the ISA(RC) power law characteristic a generalized linear regression model with only one free parameter by utilizing the prior estimate of the $e = 0$, $e = 1$ intersection coordinate $n_x := 30$. Introducing into Eq. (8), an expression for the offset $b_s(e = 1) := b_{se}(\gamma_e) = y_{px} - \gamma_e y_{sx}$ yields for the one-parametric non-nominal model

$$y_{pe}(y_{se}) = \gamma_e(y_{se} - y_{sx}) + y_{px} \tag{9}$$

leaving γ_e as single free parameter of the non-nominal model equation that is valid for $y_s \geq y_{sx}$ (for details see Appendix 2). This means that corresponding to Fig. 1 for the logistic ISA(n) characteristics in linear coordinates, also the power law characteristic exhibits a bifurcation of ISA(RC) at intersection coordinate R_x into separate branches for the nominal and non-nominal scenarios (i.e. for $RC > R_x$: ISA(RC | $e = 1$) > ISA(RC | $e = 0$)). It should be kept in mind that all the above theoretical predictions are valid only for the means of a sufficiently large statistical sample of participants.

5 Experimental results

In what follows, we use the above theoretical characteristics and numerical predictions for (nonlinear) regression analysis of the experimental subjective ISA-WL and objective radio calls communication (RC-TL) data with logistic and power law models. This analysis is based on the set of scenario means averaged across the 21 participants ($<<\text{ISA}(n_j)>>$, $<<\text{RC}(n_j)>> (j = 1, ..., 8$; see Appendix 1 for complete preprocessed dataset). In contrast to the traffic flow n (AC/h) as independent environmental load variable the measured time series of radio calls between controller and pilots represents a resource limited controller activity with upper limit R_u, well defined by simple considerations of available and required communication time (see Sect. 4.2).

After presenting the experimental ISA(n) and RC(n) results with regression analysis for scaling parameter estimates ν and ρ in Sect. 5.1 and 5.3, respectively, we focus in Sect. 5.4 on the correlation between ISA-WL and RC-TL data. In what follows (where not mentioned otherwise), we include as uncertainties for parameter regression estimates (ν, ρ, γ) standard errors ε of means (= standard deviation / $\sqrt{N}$), with 95% confidence intervals CI = ε t, and Student's t(95%) ≈ 2.1 for $N - 1 = 20$ degrees of freedom. Linear and nonlinear (iterative) regressions were performed with the Matlab® statistics toolbox using "fitlm" and "nlinfit".

5.1 Logistic < <ISA> >(n) characteristic

For the present purpose, we analyze the means across participants with the generalized linear version of the logistic model (Eq. (3)). We quantify the scaling (sensitivity) parameter ν for the nominal ($e = 0$) scenarios through application of the one-parameter model, using the theoretical intercept $b_{gt} = -\ln(4) = -1.3863$. The lower ISA scale limit $I_d = 1$ allowed for deriving the dependency between slope and shift parameter $\mu = \nu \ln(4)$. The non-nominal ($e = 1$) case with increased slope $1/\nu_e$ (and consequently k_e, μ_e) requires a two-parameter estimate (ν_e, k_e) due to the a-priori unknown intersection $I_{de} < I_d(n = 0 \mid e = 0) = 1$ of the non-nominal sigmoid. Both regressions provide an experimental estimate for the predicted intersection at (n_x, I_x) between the $e = 0$ and $e = 1$ curves. The logistic fit model for $e = 1$ neglects the small deviation originating from the (expected) merging of the $e = 0$ and $e = 1$ characteristics for $n < n_x$. Figure 4 depicts in semi-log coordinates the result of fitting transformed ISA variable $y_p(I) = \ln(p(n)/(1 - p(n)))$, $p = I(n)/I_u$, with Eq. (3).

The slope parameter ($\pm$ stderr) for $e = 0$ is estimated as $a_g = 1/\nu = \ln(4)/\mu = 0.0380$ (± 0.0004) with T test p(T = 110) = 1.7 10^{-6}. It corresponds to $\nu = 26.32$ (± 0.3) and $\mu = 36.49$. This result provides evidence that the theoretical offset $b_{gt} = -\ln(4)$ derived for the generalized linear logistic model is in fact a good approximation for the $e = 0$ scenarios.

As expected, the two-parameter regression of the $e = 1$ group of simulations (with priority event) yields less precise parameter estimates (stderr): $a_{ge} = 0.0471$ (0.0013) or $\nu_e = 1/a_{ge} = 21.231$, with $p(|T| = 37) = 0.0007$; $b_{ge} = -\mu_e/\nu_e = -1.670$ (0.05), with $p(|T| = 32) = 0.001$. Nevertheless, the CI(95%) in Fig. 4 clearly separate the transformed logistic $<<y_I>>(n)$ fits for the two factor-2 groups.

Through the inclusion of the theoretical intercept $b_g := b_{gt} = -\ln(4)$ as prior knowledge for $e = 0$, and two-parameter regression (a_{ge}, b_{ge}) for $e = 1$ the crossing coordinates of the generalized linear fits confirm (for the participant sample means) the minimum traffic flow $n = n_x$ as underload threshold: $\{n_x, y_{px}\} = \left\{\frac{(b_{ge} - b_g)}{(a_g - a_{ge})}, a_g n_x + b_g\right\} = \{31.2, -0.201\}$, and through back-transformation $I_x = I_u/(1 + \exp(-y_{px})) = 2.25$, in agreement with the theoretical predictions within the given uncertainty (for details see Appendix 2). Estimates of uncertainty (sterr.) may be derived from those of the above parameters through error propagation yielding: $\{\delta n_x, \delta y_{px}\} = \{0.6, 0.03\}$ and $\delta I_x = 0.013$. i.e. the experimental uncertainty $\delta n_x/n_x \approx 2\%$ is an order of magnitude smaller than the prior estimate (5/30 ≈ 0.2). So for the average across participants, the experimental results confirm the theoretical prediction that below threshold n_x (see Sect. 4.1) the priority event induced additional task load does
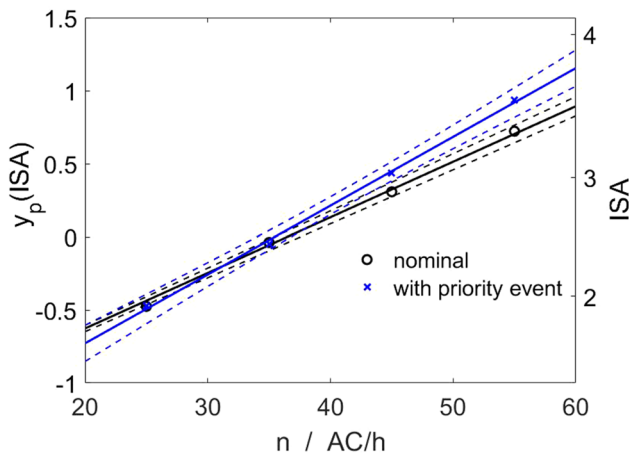
Fig. 4 Transformed ISA measurements (participant sample means of the four scenario averages for $e=0$ ($j=1$–4: circles) and for $e=1$ (with priority event, $j=5$–8, crosses). Abscissa: traffic load n (aircraft / hour); ordinate left: log(natural) of transformed ISA, right: ISA scale. Solid lines: linear regressions with 95% confidence intervals (dashed) using generalized linear logistic model with one-parameter regression (ν, Eq. 3) for $e=0$ scenarios, and with two-parameter regression (μ, ν) for $e=1$. Intersection of $e=0$, 1 lines observed at $(n_x, y_{Ix}) \approx (31, -0.2)$

Fig. 5 Radio call duration RD(n) (ordinate) as dependent on traffic flow n (Abscissa). Measured scenario mean values $j=1$–8, each averaged over the 21 subjects sample, separated for factor 2 ($e=0$: crosses $j=1$–4, nominal traffic; $e=1$: squares $j=5$–8, with priority event). Least squares fits: solid/dashed lines for $e=0/1$

not generate reporting of any additional workload, of course with large inter-individual variation (as detailed in (Fürstenau et al. 2020)).

5.2 Radio Call Duration < <RD> >(n)

Figure 5 depicts the observed linear decrease of radio call duration (RD, as mean over the 21 subject sample) with increasing traffic flow. This result is in agreement with findings of Djokic et al. (2010). According to these authors, the radio frequency occupation time as determined by radio call rate RC (frequency of radio calls, not to be confused with physical radio transmission frequency) and radio call duration RD represents the communication load as significant factor determining the workload. They report an increase of perceived WL with increasing overall frequency occupation time and with decreasing RD. We will show below that this agrees with our results with regard to RC(n) and ISA(RC). In terms of controller strategy, reduction of call duration may be understood as a method to reduce or stabilize workload in case of task load increase, e.g. through increase of traffic (Sperandio, 1978).

From underload (25 AC/h) to overload (55 AC/h) RD reduces from ca. 4 to 3.6 s/call, i.e. a decrease of 10%, independent of factor 2 ($e=0$ or 1). This is consistent with (Manning et al. 2001) who measured for en-route sector radar control an average ($\pm$ sterr) of 3 ($\pm$ 1) s. Assuming the same duration for the pilot response, the duration of communication events (e.g. for pilots clearing request) is 7–8 s. From this number, we may derive an asymptotic upper limit
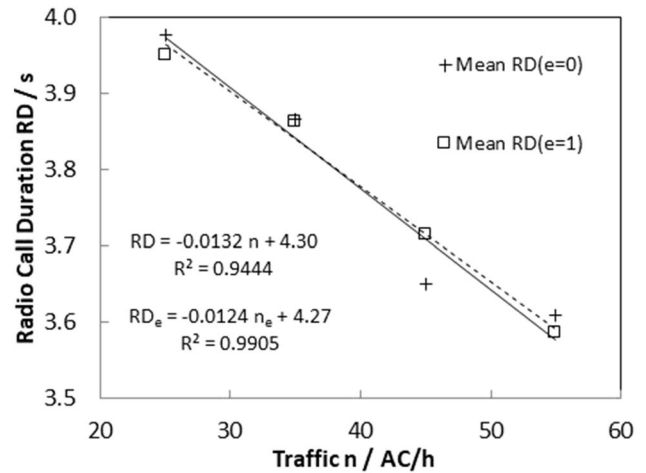
of radio call frequency as a rough estimate when we add a minimum average interruption between ATCos calls of 1 s. With $2 \times 4 + 1 = 9$ s, we obtain as maximum $RC_u = 3600 / 9 \approx 400$ calls/h.

5.3 Logistic radio call-frequency characteristic < <RC> >(n)

The iterative logistic two-parameter fit (RC_u, ρ) with Eq. (5) of ATCO's frequency of radio call (RC / calls/h) for both factor-2 cases $e=0$, 1 is presented in Fig. 6.

The regressions exhibit a common quasi exponential convergence of ($e=0$, 1) towards $RC_u \approx 400$ h^{-1}, precisely ($\pm$ sterr): $RC_u(e=0) = 388$ (± 10) and $RC_u(e=1) = 401$ (± 12), that agrees with the theoretical prediction in the previous section. Within standard errors, parameter estimates R_u ($\pm 3\%$) are the same. Also, scaling parameter estimates $\rho = 19.6$ (± 0.98) $\rho_e = 21.9$ (± 1.1) are reasonably close to the linearized theoretical prediction ($\rho_t \approx 22.5$) in Sect. 4.2. Only weak evidence is observed for a difference of scaling parameters ρ, ρ_e between nominal and non-nominal scenarios ($e=0$, 1, respectively) with measured relative sterr. of $\pm 5\%$.

The evidence for a common asymptotic limit (400 h^{-1}) is tested with the generalized linear one-parameter (ρ) model (6) using normalized variables < <RC> >$/R_u := s$, and transformation $S(s)$ (see Appendix 2, Eq. A2.12) for a linear regression as depicted in Fig. 7.

The slope estimates (with sterr) with linear regression are $1/\rho := a_s = 0.0479$ (0.0006) with $p = 5$ 10^{-6} (|T|= 77), $a_{se} = 0.0459$ (0.0004) with $p = 2$ 10^{-6} (|T|= 105). Standard

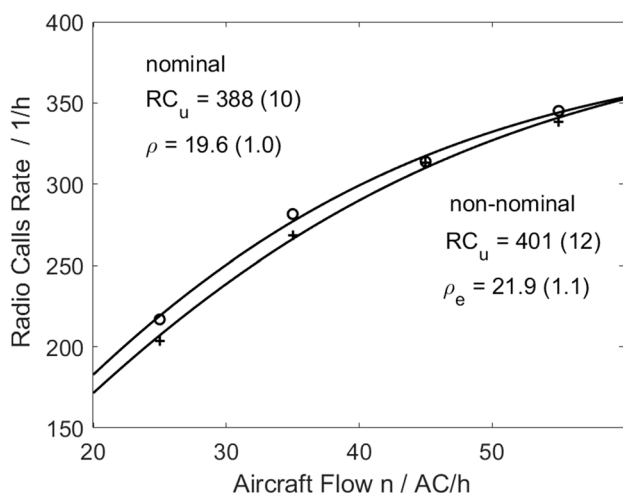**Fig. 6** Frequency of ATCos' radio calls for the eight scenarios $RC(n_j)$, $j = 1$–8, separated for factor 2 ($e = 0$: circles, $j = 1$–4; $e = 1$: crosses, $j = 5$–8) represented by scenario means averaged across the 21 subjects sample. Nonlinear regressions based on two-parametric (RC_u, ρ) logistic Eq. (5) with $RC(n = 0) = 0$

Fig. 7 Frequency of ATCos' radio contacts for the eight scenarios n_j, $j = 1$–8, separated for factor 2 ($e = 0$, $j = 1$–4, circles); $e = 1$, $j = 5$–8, crosses)). Ordinate $y_s(n)$: transformed RC variables of normalized radio call rates $R(n_j)/R_u$ (scenario means averaged across participants). Linear regressions (solid lines) based on one-parametric (scaling ρ) generalized linear (log-lin) form of logistic model (Eq. 6) with $R_u := 400$ as prior knowledge. Dashed/dotted lines: 95% CI for $e = 0/1$ condition

errors $\leq 1\%$ and T test suggest to reject the 0-hypothesis, i.e. within the 95% CI a significant effect of the priority event leading to reduced radio call rate is observed.

5.4 Power Law Characteristic of $<<ISA>>(<<RC>>)$

The ISA(RC) power law (Eqs. 7, 8) was derived from the parametric representation $\{R(n \mid \rho), I(n \mid \nu)\}$ (for details see Appendix 2). By eliminating n, it provides the direct dependence of the subjective ISA-WL response on objective communication task load (RC-TL) through the exponent γ. With ν and ρ estimates in Sects. 5.1 and 5.3, we can predict the power law exponent via the theoretically derived ratio $\gamma = \rho/\nu$ of the measured logistic parameter values for comparison with the direct estimate of γ (independent of n) using Eqs. (7, 8), and with the theoretical estimate in Sect. 4.3. The numerical estimates are collected in Table 2, separated for nominal/non-nominal scenarios, γ and γ_e, respectively. A significant increase is observed for the priority event scenarios γ_e ($e = 1$) as compared to $\gamma(e = 0)$. This means that the sensitivity (= slope) $a_g = 1/\nu$ of the transformed subjective ISA(n)-WL characteristics increases significantly more for the priority scenarios than slope $a_{sg} = 1/\rho$ of the transformed objective communication variable $R(n)$.

The nonlinear (iterative) LSQ fit $<<ISA>>(<<RC>>)$ with model Eq. (7) of the eight measured scenario means, each averaged over the participants and separated for factor 2 ($e = 0, 1$) is depicted in Fig. 8.
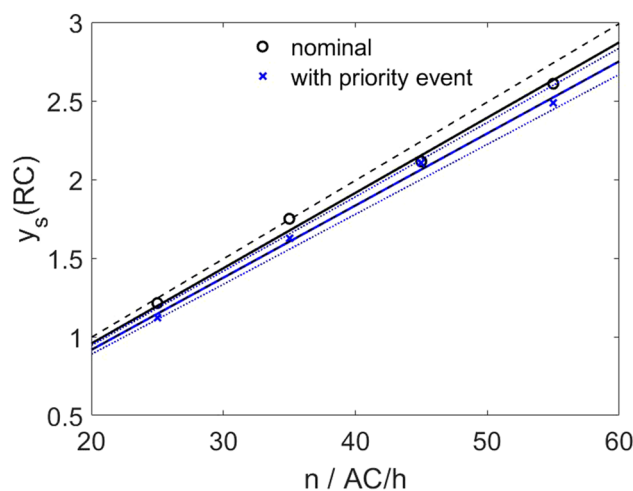
In normalized coordinates, the $s(n) = R(n)/R_u$ measurement range covers $0.5 \leq s \leq 0.9$ with $p(n) = \text{ISA}/I_u$ range $0.38 \leq p \leq 0.68$ which may be compared with the theoretical prediction in Fig. 3. Like ISA(n) also ISA(RC) exhibits a significant increase for priority event scenarios ($e = 1$, dashed line) as compared to the nominal case (solid line), partly due to the inverse behavior of the calls vs. traffic flow RC(n) in Fig. 6 and 7. With the (k, γ) parameter estimates ($e = 0$: {4, 0.792}; $e = 1$: {5.28, 1.025}), the measured crossing coordinates are calculated as $\{R_x, I_x\} = \{214, 1.96\}$ with standard errors of the order 2%. These values are close to the low traffic underload region, in reasonable agreement with the theoretical prediction $\{R_x, I_x\} \approx \{233, 2.3\}$ (see Sect. 5.1 and Appendix 2, Eq. A2.18). For comparison, the linear regression of the transformed $y_P(p)$ vs. $y_S(s)$ data (with the 95% confidence intervals) with the generalized linear model Eq. (8) is depicted in Fig. 9.

The one-parameter fit estimate of slope γ ($\pm$ stderr) for the nominal scenarios $e = 0$ (with theoretical value $k := k_t = 4$, intercept $b_s := -\ln(k) = -1.386$) is obtained as $\gamma = 0.7933$ (0.011; $t = 69.9$, $p = 6.4 \cdot 10^{-6}$). For the non-nominal scenarios ($e = 1$), the two-parameter fit yields: $\gamma_e = 1.025$ (0.069); $t = 14.7$, $p = 0.046$; $b_{se} = -\ln(k_e) = -1.668$ (0.132); $t = -12.6$, $p = 0.006$, i.e. $k_e = 5.302$, with somewhat reduced confidence as depicted in Fig. 9. The values confirm the above results of the iterative NL regression with sufficient significance according to t tests within 95% CI. Introducing

Table 2 Comparison of theoretical predictions with experimental parameter estimates

Parameter estimate (sterr)	ν log-lin model Equ. (3)	ρ log-lin model Equ. (6)	γ power lawmodel Equ. (7) nonlin. regression	γ log–log power law model Equ. (8), lin. regression	γ T-testp(T)
Theory ($e=0$)	$\approx \mu_t/\ln(4)$	$\approx n_c/2$	ρ_t/ν_t	$(y_{px}+\ln(4))/y_{sx}$	
$\mu_t:=35, n_c=n_3$	$=25.25$	$=22.5$	$=0.89$	$=0.89$	
Nominal Scenarios ($e=0$) 1-param. fit	26.3 (0.3)	20.9 (0.3)	0.79 (0.01)	0.79 (0.01)	$6.4\ 10^{-6}$ (70)
Priority Scenarios ($e=1$)					
2-parameter fit	21.2 (0.6)		1.03 (0.03)	1.03 (0.07)	0.046 (15)
1-parameter fit		21.8 (0.2)		1.06 (0.02)	$7.8\ 10^{-6}$ (66)
$e=0/1$-Intersection	$\{n_x, I_x\} =$	$\{n_x, R_x\}$	$\{R_x, I_x\} =$	$\{y_{sx}, y_{px}\} =$	
Theory	$\{30, 2.3\}$	$:= \{0,0\}$	$\{233.1, 2.3\}$	$\{1.33, -0.2\}$	
Experiment (sterr)	$\{31.2\ (0.6), 2.3\ (0.3)\}$		$\{214\ (4), 1.96\ (0.04)\}$	$\{1.2, -0.42\}$	

Column 2, 3: Logistic parameter estimates for y_p and y_s vs. traffic flow n (from separate one- and two-parameter regressions with generalized linear (log-lin) model for e=0, 1, respectively). Column 4, 5: Comparison of power law exponent $\gamma = \rho/\nu$ (from logistic parameters) with independently estimated γ using nonlinear and generalized linear model $y_p(y_s)$ (rel.sterr. < 2% for $1/\nu$ and $\approx 1\%$ for ρ). Last row: intersection coordinates of e=0, 1 characteristics. For details, see text
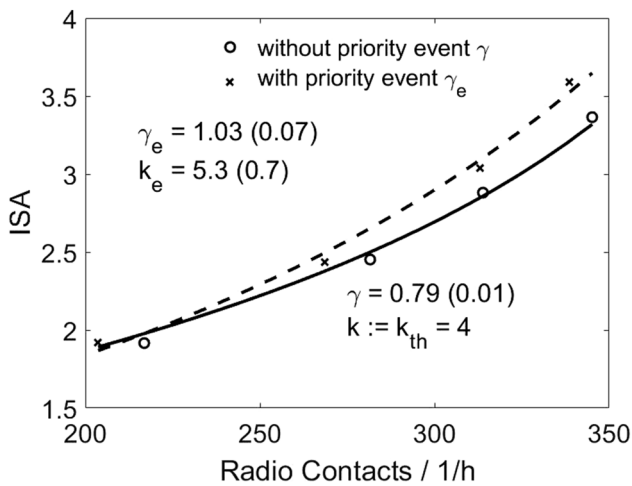

Fig. 8 $<<ISA>>$ vs. $<<RC>>$ scenario means ($j=1$–8) averaged across participant sample together with model-based nonlinear fit (Matlab NLINFIT, using Eq. 7). Separated for $e=0$ (circles, $j=1$–4, solid regression line) and $e=1$ (crosses, $j=5$–8, dashed regression line). Power law fit parameters (k, γ) with standard errors. Intersection coordinate of $e=0$, 1 curves observed at $\{R_x, I_x\} \approx \{215, 2\}$. For details, see text and Appendix 2
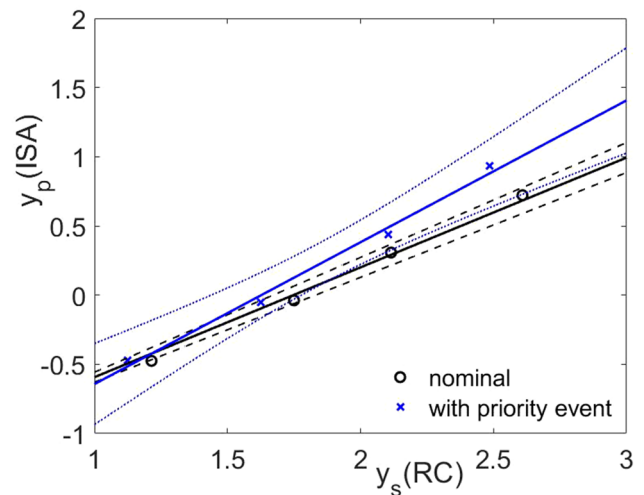

Fig. 9 Normalized, transformed ($y_P(p)$ vs. $y_S(s)$) scenario means ($j=1$–8) $<<ISA>>$ ($<<RC>>$) averaged across the 21 participant sample together with model-based generalized linear (log–log) fit (using Eq. (8), solid lines), separated for factor 2. ($e=0$: $j=1$–4, circles, 1-parameter fit (γ), dashed lines: 95% CI; $e=1$: $j=5$–8, crosses, two-parameter fit (γ_e, b_{se}) dotted lines: 95% CI). For details see text and Appendix 2

these parameter estimates into the intersection equations for the fit parameters (Appendix 2, Eqs. A2.18, A2.19) provides the bifurcation coordinates in transformed double-log coordinates $\{y_{sx}, y_{px}\} = \{1.216, -0.422\}$ corresponding to $\{R_x, I_x\} = \{217, 2.0\}$, in agreement (within given uncertainties of 2%) with the above results of the nonlinear fit. It also agrees reasonably well with the rough theoretical estimates obtained with the a priori assumptions in Sect. 4.3

$(n_x := 30, \rho := \rho_t = 22.5)$, yielding $\{y_{sx}, y_{px}\}_t = \{1.33, -0.20\}$. By means of the experimental estimate $n_x = 31.2 (\pm 0.6)$ in Sect. 5.1 (i.e. uncertainty reduced by factor 10 as compared to the prior guess) and a relationship for the intercept $b_{se}(\nu, \nu_e, \rho_e, n_x)$ (Eq. A2.20, Appendix 2), we obtain a one-parameter fit with Eq. (9) also for the priority scenarios ($e=1$), yielding $\gamma_e = 1.058 (0.016)$; $t=65.6$, $p=7.8\ 10^{-6}$, i.e.

a significantly improved confidence. Table 2 summarizes the results of the model-parameter estimates based on the nonlinear and generalized linear characteristics fitted to the scenario means averaged across the 21 subject samples.

6 Discussion

Based on the cognitive resource limitation hypothesis and prior knowledge (domain experts, WL, and TL scales), logistic two-parameter models (shift, scaling) were designed for theoretical parameter prediction and regression-based parameter estimates of the experimental subjective ISA-WL and objective communication (RC-TL) data. Traffic flow n (aircraft per hour entering the approach sector) served as (environmental) independent load parameter (four levels of traffic load) that defined the scenarios in the human-in-the-loop simulation experiments of approach traffic control. A non-nominal event (priority request, $e = 1$) was hypothesized to increase WL measures at higher load levels ($n > n_x >$ under-load n_1). For nominal traffic load scenarios ($e = 0$) prior knowledge and simple communication time considerations allowed to derive asymptotic upper scale limits for normalization of dependent variables (ISA, RC) and simplification of model equations to one free (scaling) parameter (ν_t and ρ_t) for ISA-WL and RC-TL, respectively. As a consequence the transformation (from linear scales) of logistic ISA and RC characteristics into generalized linear characteristics (log–log scales) y_p(ISA), y_s(RC) allowed to derive a $y_p(y_s))$ power law relationship, independent of environmental load n, with theoretical exponent $\gamma_t = \rho_t / \nu_t$ as formal equivalent to the classical stimulus–response (Stevens) law of psychophysics.

6.1 Subjective Workload Measure ISA(n)

A covariance and nonlinear regression analysis of the complete set of individual ISA(n) data (see Table A1.1 in Appendix 1) in a previous article (Fürstenau et al. 2020) provided initial evidence for the validity of the logistic model. In fact a normalized ISA-WL-sensitivity index derived from a linear approximation to Eq. (2) was successfully applied to clustering of the participants with regard to a new neurophysiological WL index (DFHM, see Sect. 2.4) that was measured simultaneously with the ISA and communication data within the present experiment (Radüntz et al. 2020a, b). With a plausibility argument for a numerical estimate of μ ($:= (n_1 + n_3)/2 = 35 =$ operational traffic n_2) also the theoretical prediction for the logistic scaling parameter ν (theory: $\nu = \mu/\ln(4)$) exhibited surprising agreement with the corresponding regression parameter estimate (see Table 2, Sect. 5.4). The experimental results confirmed the theoretically predicted subjective ISA range (for means across participant sample) for

the given range of traffic flow levels $n_1 \leq n \leq n_4$ which were selected by domain experts as realistic a-priori values for the specific approach area ($\leq n_3$), with n_4 as excessive load.

For the non-nominal scenarios (including a priority event $e = 1$), a theoretical estimate of the hypothesized sensitivity increase ($1/\nu_e > 1/\nu$, for $n > n_x$) did not exist so that also the second model parameter had to be estimated through regression. The numerical parameter estimates are included in Table 2 (Sect. 5.4). For the predicted intersection between nominal and priority event scenarios, an experimental value $n_x = 31.2$ (± 0.6) was obtained with Eq. A2.9 (Appendix 2) using parameter estimates (a_g, a_{ge}, b_{ge}). This is reasonably close to the initial theoretical guess $n_x = 30$ (± 5), however, with an uncertainty reduction by an order of magnitude.

It should be pointed out that the significance of the ISA-WL increase of the non-nominal scenarios was based on a conservative analysis of data, because we averaged the ISA values across the full scenario times, whereas the (pseudo) pilots' priority request during the $e = 1$ scenarios was introduced after simulation time $t_s = 10$ min. It means that any influence on workload could be effective only during $t_s > 10$ min so that the restriction on the ISA averaging for $t_s > 10$ min should increase the scaling parameter differences between $e = 0, 1$, however, at the cost of increased uncertainty. This successful validation of the logistic ISA(n) model, including the bifurcation coordinates (n_x, I_x) provided the initial evidence for the basic hypothesis of cognitive resource limitation.

Our logistic modeling approach may be compared with results of a simulated ATC HITL-simulation experiment reported by Lee et al. (2005). They used a heuristic sigmoid function based nonlinear 4-parameter regression for ATWIT-WL data analysis (see Sect. 2.2) which exhibited significant parameter estimates.

6.2 Objective task load measure radio calls rate RC(n)

The theoretical predictions in Sect. 4.1 and 4.2 showed that within the traffic load range $n_1 \leq n \leq n_4$, in contrast to the nearly linear increase of the subjective workload ISA(n) the radio calls rate RC(n) as task load (TL) exhibits a clear nonlinear resource limitation behavior (see Fig. 2). This correlates (nonlinearly) with the decreasing mean call duration from 4 to 3.6 s and it agrees with the experimental results (Fig. 6). It is interesting (and counter-intuitive) that the RC(n)-TL appears systematically higher for the nominal traffic case ($e = 0$) as compared to the case of a non-nominal event ($e = 1$), in contrast to the ISA behavior where just the opposite is observed. Although the decrease of slope $a_{se} = 1/\rho_e$ amounts to only $\approx 4\%$ ($\approx$ decrease of call rate at $n = n_c = 45$ the difference is significantly larger than the standard error of 1% for $1/\rho$. Therefore, at the same

traffic level n, the average of controllers basic communication activity ($<<RC>>$ calls/h) is estimated significantly lower with the additional task ($e = 1$ scenarios, $j = 5-8$), with regard to sterr. and predicted by theory.

As explanation, the radio calls rate is thought to be decreased during a priority event (reducing the respective part of task load) as a strategy to be able to spend cognitive resources for solving the additional priority task, because under high traffic volume both the nominal traffic and the priority event contribute to communication load i.e. the same processing modalities (Wickens 2002). Sperandio in (Sperandio 1978) categorized the strategy change under increasing load as workload homeostasis. For approach controllers, he found that under low traffic they preferred a direct approach strategy with extensive use of aircraft (AC) performance data for individual control (direction, altitude, speed,…), whereas under high traffic they switched to standard procedures with global routing for most AC, i.e. global approach sequence with standard separation distance so that control of the first AC in the sequence would be sufficient (see also Sect. 4.2). Nevertheless, under higher traffic volume despite reduced communication load with $e = 1$, the complete subjectively experienced workload due to auditory and visual processing increases due to the additional task and obviously is included into the ISA($e = 1$) WL rating.

The approach to the asymptotic limit R_u of RC communication load as objective TL measure (see Fig. 6) is reflected by the comparable asymptotic behavior of the objective neurophysiological DFHM(n) workload measure (dual frequency head map index DF(n), see Sect. 2.4) which was measured simultaneously (Radüntz et al. 2020a, b). The following Fig. 10 depicts the measured data of all eight scenarios for the DF(n) scenario means averaged across participants together with a nonlinear regression analysis using model Eq. (5) (originally designed for communication task load RC(n) with RC($n = 0$) $= 0$). To highlight the similarity to the RC(n) characteristic (Fig. 6, Sect. 5.3), the fit was not separated for the two conditions $e = 0$, $e = 1$. The default mode EEG activity during the rest measurements was reflected in a corresponding offset (average across participants (DF$_0 = 25 \pm 5$) that was measured separately for each participant before and after each simulation run and (as average of bias runs) subtracted from DF(n) data before regression.

Analyzed data for $j = 1-8$ were limited to the simulation time interval following the priority event ($e = 1$ scenarios) at $t_s = 10$ min, within 0.5–3 min after t_s. Because no prior information was available for deriving an estimate of the effective asymptote a two-parameter fit had to be used for estimating DF$_u$ ($= 62.8$ (± 0.2)) and scaling parameter δ ($= 11.8 (\pm 0.2)$, $t = 37$, $p = 3 \cdot 10^{-9}$). This result indicates that the neurophysiological DFHM index derived from the EEG as response to the ATCO's activity appears to reflect the task
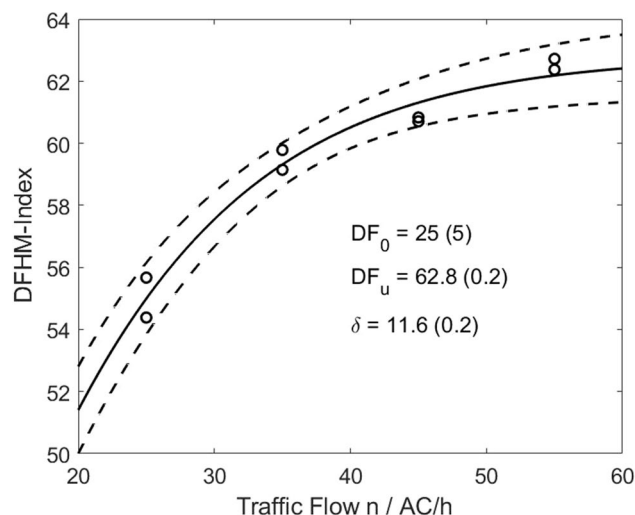

Fig. 10 Neurophysiological DF(n) WL index (analyzed simulation time interval $t_s = 10.5-13$ min, immediately following priority request of e $= 1$ scenarios at $t_s = 10$ min, as dependent on traffic load n_j, $j = 1-8$, for both e $= 0$ and e $= 1$ scenarios (circles: average across participants for scenario means). Solid curve: 2-parameter nonlinear iterative lsq. regression using logistic Eq. (5) after offset subtraction. Inset: Parameter estimates, offset DF$_0$, asymptote DF$_u$, scaling parameter δ ($\pm$ sterr). Dashed curves: 95% confidence interval

load, measured through the communication load as a kind of physical stimulus (DF(RC)), corresponding to the ISA(RC) power law. The load sensitivity of the new DFHM WL index relative to the RC(n)-TL sensitivity may be estimated from the ratio of the logistic function slopes at $n = 0$ (inversion point), i.e. the logistic scaling parameter ratio $\rho/\delta = 20.9 /11.6 = 1.80$ (± 0.04). This ratio is exactly the prediction for a power law (psychophysical) stimulus (RC)–response (DF) exponent γ_d according to Eq. A2.16 (with $m_s = m_p = 0$, Appendix 2). It appears of interest that in contrast to subjective ISA(n)-WL sensitivity ($1/\nu$), the DFHM index with $1/\delta > 1/\rho$ exhibits a higher sensitivity than RC-TL with regard to dependence on traffic flow (for low n).

6.3 Power law stimulus–response relationship ISA(RC)

The subjective ISA report on a one-dimensional online WL measure reflects the load due to the different limited resources attributed to perceptual input (visual, auditory), response demand (vocal), and cognitive (judgement, decisions, strategy change) modalities with corresponding processing stages (Wickens 2002). That is why ATCO's ISA report was not expected to correlate linearly with the radio communication activity RC (calls/h) derived from the logged simulation data (time series of traffic flow, pilots clearance request, ATCO's communication times and duration). Another reason is given by the communication time

restrictions (required time / available time for radio calls < 1) with upper limit R_u of calls frequency (see Sect. 4.2). The predicted theoretical RC(n) characteristic contrasts with the nearly linear ISA correlation with n (for the given simulated traffic range; compare Fig. 1 and Fig. 2). This is theoretically reflected by the ISA(RC) power law relationship formalized by Eqs. (7) and (8) (see also Appendix 2).

The average frequencies of radio calls between controller and pilot ($<<RC>>_j$ (calls/h) as mean per scenario $j = 1 \ldots 8$, averaged across the participants) represent as a response demand a more direct measure of task load than the traffic flow n_j as environmental load variable, visualized in the radar display. The ISA(RC) and transformed $y_p(y_s)$ model functions used for the regression-based model parameter estimates for $e = 0, 1$ (γ, b_s), and the derived intersection coordinates (R_x, I_x) provided experimental evidence (Figs. 8 and 9) for the theoretically predicted power law characteristic (Eqs. 1, 7, 8, 9). The numerical predictions were achieved with a plausible assumption on the logistic shift parameter $\mu_t := 35$ AC/h = operational traffic n_2 (together with theoretically predicted relationship $\nu_t = \mu_t / \ln(4)$), and with a RC(n)-slope linearization, respectively (yielding $\rho_t \approx n_c / 2 = 22.5$ AC/h). The value of the theoretically predicted dimensionless power law exponent $\gamma_t = \rho_t / \nu_t = 0.89$ corresponds to the order of magnitude of typical psychophysics (Stevens) exponents (e.g. Link 1992). Considering the uncertainty of the theoretical prediction with $\delta\mu_t / \mu_t \approx 15\%$, γ_t compares well with the experimental regression-based estimates in Table 2 (Sect. 5.4).

ISA(RC) dependence on frequency of radio calls (Figs. 8, 9) like ISA(n) in Fig. 4, exhibits a significant sensitivity increase with priority event scenarios ($e = 1$) as compared to the nominal case ($e = 0$), i.e. $1/\nu_e > 1/\nu$, despite the inverse behavior of RC(n) (Figs. 6, 7) as compared to ISA(n). This factor-2 effect again exhibits a clear onset at the intersection point $\{R_x, I_x\}$ and $\{y_{sx}, y_{px}\}$ of the $e = 0$, $e = 1$ curves in Figs. 8 and 9, respectively. The (k, γ) parameter estimates obtained with an iterative regression procedure (Fig. 8, Sect. 5.4) are consistent with those obtained with the generalized linear model (Fig. 9, Table 2). The crossing coordinate I_x of the $e = 0, 1$ bifurcation obtained with the power law analysis, is in reasonable agreement with that one obtained from the logistic ISA(n) characteristics ($I_x = 2.3$) in Sect. 5.1 (Fig. 4). Our results for the RC(n) characteristics in Sect. 5.3 indicated the strategy change of ATCOs radio communication in $e = 1$ scenarios for $n > n_x$ (see Sects. 4.2, 6.2) to stabilize the communication TL. The result exhibited an inverse behavior of the TL sensitivity ($1/\rho_e < 1/\rho$) as compared to the ISA-WL sensitivity ($1/\nu_e > 1/\nu$). That is why the theoretically derived relationship $\gamma = \rho/\nu$ could explain the (slightly) improved significance of the observed bifurcation of the

power law characteristic at R_x with slope increase (log–log coordinates) $\gamma_e/\gamma = 1.03 / 0.79 = 1.30$ (± 0.04) as compared to the logistic slope increase of the ISA-WL-sensitivity ratio $1/\nu_e / 1/\nu = 26.3/21.3 = 1.23$ (± 0.04) (see Table 2 in Sect. 5.4). The measured exponent $\gamma < 1$ for the nominal traffic condition is consistent with the result of Bachelder et al. $0.2 < \gamma < 0.4 < 1$ (see Sect. 2.5) (Bachelder and Godfroy-Cooper 2019).

For the non-nominal ($e = 1$) scenarios, a theoretical estimate of the offset parameter $b_{se} = -\ln(k_e)$ due to the predicted sensitivity increase ($\gamma_e > \gamma$ for RC $> R_x$) did not exist a priori (in contrast to ISA(n) with $n_x := 30$ by plausibility) so that for e = 1 also the second model parameter k_e or b_{se} had to be estimated through regression that increased the uncertainty of the $e = 1$ parameter estimates as compared to the nominal case, $e = 0$. However, using the experimental estimates of $n_x, \nu, \rho, \nu_e, \rho_e$ derived from the logistic ISA(n) and RC(n) models as prior information (Sects. 5.1, 5.3), an estimate for b_{se} was obtained via Eq. (A2.20) in Appendix 2. This allowed for a one-parameter (γ_e) model also for the $e = 1$ condition (Eq. 8), valid for RC $\geq R_x$, with a reduction of uncertainty $\delta\gamma_e$ by a factor 4 (see Table 2).

To summarize, our predicted and measured power law exponents γ are of the order of 1, in agreement with the classical Stevens exponents (e.g. Link 1992; Stevens 1957). Moreover, for the nominal condition ($e = 0$), we obtained $\gamma < 1$ (i.e. subjective response sensitivity $1/\nu <$ objective TL (stimulus) sensitivity $1/\rho$), in agreement with the recent result of (Bachelder and Godfroy-Cooper 2019) (see Sect. 2.5).

The preliminary regression analysis of the new neurophysiological EEG-DFHM index in the previous Sect. 6.2 (Fig. 10) with regard to the logistic dependence on traffic load n suggested the generalizability of the power law hypothesis also to the DF(RC) correlation. Figure 11 depicts the linear one-parameter regression of the normalized transformed DF(RC) data (after subtraction of offset $DF_0 = 25$) for the estimate of exponent γ_d. Asymptote $DF_u = 63.6$ as prior information required for normalization (with slight increase due to algorithmic requirements) was derived from the nonlinear logistic DF(n) regression.

The estimate of the exponent $\gamma_d = 1.66 \pm 0.05$ was obtained independently of the environmental load parameter n. It confirms the theoretical prediction in the previous section where the exponent was calculated based on the theoretically derived ratio of the logistic scaling parameters $\gamma_\delta = \rho/\delta = 1.80 \pm 0.04$. Consequently, the power law fit supports the initial result that the neurophysiological DFHM index exhibits a higher sensitivity ($\sim 1/\delta$) to environmental load (for low n) than RC-TL ($1/\rho$) and than ISA-WL as well.
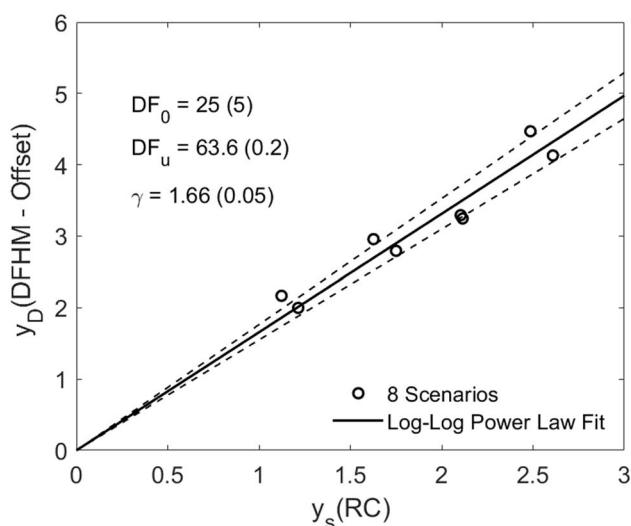
Fig. 11 Transformed normalized DFHM workload index (after offset DF_0 subtraction) vs. transformed radio calls rate $s = RC/R_u$ in log–log scale (circles: average across participants for both $e = 0$, 1 scenario means $j = 1$–8; analyzed simulation time interval $t_s = 10.5$–13 min, following priority event at $t_s = 10$ min for $j = 5$–8). Solid line: 1-parameter (γ) lin. regression power law fit (gen. lin. model Eq. (A2.16) with $b_s = 0$, $\Delta_s = \Delta_d = 1$ and 95% confidence interval (dashed). For details, see text and Appendix 2

7 Conclusion

The goal of the present work was to provide evidence by means of selected WL/TL measures of an ATC-simulation experiment for a psychophysics (stimulus – response) power law model with (Stevens) exponent as combined WL/TL index (e.g. Stevens1957, 1975; Link 1992). It combines two logistic model-based (WL, TL)-measures via a common independent stress generating environmental load variable (simulated traffic flow n (aircraft/h)). This goal means an expansion of a previous publication (Fürstenau et al. 2020) where we provided evidence for the logistic dependency of air traffic controllers' subjective quasi real-time ISA-WL reports on the environmental traffic load n as a consequence of the cognitive resource limitation hypothesis (e.g. Kahnemann 1973; Wickens and Hollands 2000). This result agrees with previous reports of a comparable sigmoid dependency of ATWIT-WL self-report on traffic count in (Lee 2005) (see Sects. 2.2, 6.1). Moreover, our results support the early psychophysics approach to workload by Gopher and Braune (1984), and they are consistent with a recent application of Stevens law to workload data analysis by (Bachelder and Godfroy-Cooper 2019) (see Sects. 2.2, 2.5).

The present stimulus–response form of the power law (Eq. 1) was obtained by combination of logistic two-parameter models after transformation of (ISA, RC) measures into normalized variables (P(ISA), S(RC)). The formal derivation relied on the prior knowledge on asymptotes (R_u, I_u) of the two-parametric {ISA($n \mid I_u$, ν), RC($n \mid R_u$, ρ)} characteristics, with logistic sensitivity parameters (ν, ρ). Based on the generalized linear (log–log) version of the power–law model $y_p(P(\text{ISA}/I_u)) \sim \gamma y_s(S(\text{RC}/R_u))$, the experimental exponent estimates γ and γ_e under nominal and non-nominal traffic load conditions, respectively, could be derived directly as slope parameters from the transformed ISA(RC) data, independently of the environmental load variable n. Within sterr. they agreed with the theoretically predicted ratio of the logistic WL/TL sensitivities $\gamma_t = (1/\nu_t)/(1/\rho_t) = 0.8$ ($\gamma_{te} = 1.0$)

With regard to formal aspects of model development, this procedure exhibits some analogy to the derivation of Stevens law within the wave discrimination theory of psychophysics (Link 1992) (see Sect. 2.5). Specifically, our theoretical prediction of the power law exponent $\gamma = \rho/\nu$ may be compared with the ratio of two normalized subjective response thresholds A_S/A_P that relate two simultaneously measured stimulus and perception sensations to logistic response probability functions. This analogy of our deterministic resource limitation-based workload model with Link's stochastic wave discrimination theory of psychophysics suggests both approaches to explain complimentary (stochastic vs. deterministic) aspects of (one-dimensional) subjective real-time workload measures in terms of stimulus–response relationships (see also Appendix 2: "Resource Limitation Model").

The subjective ISA-WL and objective communication RC-TL data analyzed in the present work served as a reference for a new neurophysiological workload measure (DFHM index) based on real-time EEG-data (Radüntz 2017; Radüntz et al. 2020a, b). Initial model-based analysis of the simultaneously measured DFHM-WL index provided evidence (see Fig. 10 in Sect. 6.2) that also this new neurophysiological measure follows the same logistic dependence on the environmental traffic load n (with scaling δ) as the communication task load variable RC(n) (scaling ρ). Consequently, the data provided evidence also for a power law relationship for DF(RC) with exponent $\gamma_d = \rho/\delta$ (Fig. 11 in Sect. 6.3), like ISA(RC) with $\gamma = \rho/\nu$. This suggests the hypothesis that the neurophysiological DFHM WL index may be treated as objective (bio-) physical stimulus inducing the subjective ISA-WL response in a power law relationship with exponent $\gamma_D = \delta/\nu$, formally equivalent to the RC-TL stimulus generating the ISA-response according to ρ/ν. A causal sequence may thus be hypothesized, according to operator activity driven by traffic count or flow: n (AC/h) communication (RC(ρ)) neural activity DFHM(δ) conscious response ISA(ν), quantified by logistic sensitivities and power law (stimulus – response) exponents γ, γ_d (Sect.2.5) and γ_D. Based on the cognitive resource limitation

hypothesis, these parameters quantify the dissociation between subjective and objective WL and TL measures, respectively.

The psychophysical scaling offers a theoretically founded approach for a WL parameter γ that combines subjective and objective measures. It was shown to predict the (near real-time) ISA-WL means across a sufficiently large sample of well-trained domain experts in a work environment reasonably well described by one-dimensional TL/WL variables. The ISA(RC) characteristic after transformation into normalized variables P(S) represents a response or judgement measure on the perceived workload, with normalized task load S corresponding to the physical stimulus of classical psychophysics experiments formalized by Stevens law.

Ongoing research addresses the more detailed (logistic and power law) model-based analysis of the DFHM index and HR/HRV WL data of the present experiment to clarify the generalizability of the psychophysics approach to workload. One potential major field of application of the discussed predictive mental WL models for online WL measures is adaptive automation for WL stabilization in safety critical (aeronautic) human system interfaces (Parasuraman and Hancock 2001; Prinzel et al. 2003).

Appendix 1: Individual ISA and RC averages across scenarios

The following two tables list the 168 scenario means $<ISA>(n_{jk})$ and $<RC>(n_{jk})$ of the eight scenarios $j = 1…8$ for each of the participants $k = 1–21$, with $j = 1…4$ for nominal traffic ($e = 0$) and $j = 5…8$ for traffic with non-nominal (priority) event (e = 1) after ca. 10 min of simulation time. The $<ISA>_{jk}$ (short $<I>_{jk}$) provide the data base for the regression analysis and discussion in Sectis. 5 and 6.

See Tables 3, 4.

Table 3 ISA means averaged over scenario time series ($<ISA>(n_{jk})$, $j = 1–4$ for nominal ($e = 0$), $j = 5–8$ for non-nominal (priority event) scenarios. Participants $k = 1 … 21$; subject code 14–34) as used for the model-based data analysis. Last three rows: column means, standard deviations, standard errors

Subject	Scenario means $<ISA>$ $j = 1–4$, $e = 0$				Scenario means $<ISA>$ $j = 5–8$, $e = 1$			
Code | n_j	25.00	35.00	45.00	55.00	25.00	35.00	45.00	55.00
14	1.50	3.25	3.25	3.25	1.60	2.40	2.80	3.60
15	2.00	2.75	3.50	3.50	2.40	2.80	3.60	4.40
16	2.00	3.00	3.50	4.25	2.60	2.60	3.40	4.20
17	2.00	2.00	4.00	3.50	1.80	2.60	3.20	3.80
18	3.00	3.25	3.25	3.50	3.00	3.00	3.20	4.20
19	3.00	3.00	3.00	3.50	2.40	2.80	3.20	3.40
20	2.00	2.75	3.25	3.75	1.40	2.60	3.00	3.20
21	1.25	2.25	2.00	3.00	2.75	1.80	2.60	3.40
22	2.00	3.00	3.25	2.75	1.40	2.80	3.00	2.80
23	1.75	2.00	3.00	4.00	1.80	2.00	3.00	3.40
24	2.00	2.25	2.00	3.75	1.60	3.00	3.40	3.40
25	1.00	1.50	2.25	2.75	1.60	1.40	2.80	3.40
26	2.25	2.00	2.75	3.25	2.00	2.60	3.00	3.60
27	2.00	2.00	2.25	2.75	2.00	2.20	2.00	2.60
28	1.25	2.50	2.75	3.50	1.60	2.80	2.80	4.00
29	2.25	2.75	3.25	4.67	2.20	3.20	4.00	4.60
30	1.50	2.00	2.00	2.50	1.20	2.00	2.80	2.80
31	1.25	2.00	2.00	2.75	1.40	2.00	2.40	2.80
32	2.00	2.50	3.75	4.25	1.80	2.20	4.20	4.60
33	1.75	1.75	2.50	2.25	1.80	1.80	1.60	3.20
34	2.50	3.00	3.00	3.25	2.00	2.60	3.80	4.00
Mean μ_p	1.92	2.45	2.88	3.37	1.92	2.44	3.04	3.59
stdev σ	0.53	0.52	0.62	0.62	0.48	0.47	0.61	0.60
sterr $\sigma/\sqrt{N}$	0.12	0.11	0.14	0.14	0.10	0.10	0.13	0.13

Table 4 RC means averaged over scenario time series ($<RC>(n_{jk})$, $j=1$–4 for nominal ($e=0$), $j=5$–8 for non-nominal (priority event) scenarios. Participants $k=1$–21; subject code 14–34) as used for the model-based data analysis. Last three rows: column means, standard deviations, standard errors

Subject	Scen. means $<RC(n)>$ / AC/h, $j=1$–4, $e=0$				Scen. means $<RC(n)>$ / AC/h, $j=5$–8, $e=1$			
Code | n_j	25	35	45	55	25	35	45	55
14	180	249	294	309	172.8	187.2	276	319.2
15	246	309	336	336	235.2	300	340.8	333.6
16	198	303	297	339	225.6	259.2	288	343.2
17	210	243	297	318	182.4	264	273.6	321.6
18	234	270	321	330	204	247.2	290.4	300
19	219	231	255	288	187.2	228	261.6	264
20	240	330	327	423	196.8	304.8	388.8	360
21	222	273	276	306	300	280.8	319.2	328.8
22	234	357	297	384	225.6	314.4	333.6	364.8
23	198	240	291	273	187.2	259.2	283.2	331.2
24	237	303	315	381	196.8	295.2	307.2	350.4
25	270	285	378	408	256.8	300	384	403.2
26	222	267	327	366	204	268.8	290.4	357.6
27	204	237	321	363	206.4	264	331.2	314.4
28	174	270	282	336	180	247.2	292.8	340.8
29	195	237	219	294	160.8	211.2	290.4	297.6
30	243	387	414	405	220.8	292.8	381.6	386.4
31	252	330	384	399	213.6	314.4	360	393.6
32	201	276	339	354	184.8	278.4	316.8	376.8
33	195	297	360	378	189.6	297.6	333.6	381.6
34	180	219	261	258	146.4	225.6	230.4	240
$\mu_s(<RC>)$	216.9	281.6	313.9	345.1	203.7	268.6	313.0	338.5
stdev σ_s	26.33	44.00	45.95	46.85	33.71	34.94	42.18	41.44
$\sigma_s / \sqrt{N}$	5.75	9.60	10.03	10.22	7.36	7.62	9.20	9.04

Appendix 2: Derivation of theoretical model equations

Resource limitation model

The most simple formal approach for growth dynamics under limited resources (e.g. processing and memory capacity) of a characteristic system variable (energy consuming cognitive activity) is given by the nonlinear (1st order, nonlinear 2nd degree) logistic or Verhulst differential equation:

$$df/dt = \dot{f}(t) = \kappa f(t)(1 - f(t)) \qquad (A2.1)$$

With normalized function $f(t) = F(t)/\Delta F$, growth interval $\Delta F = F_u - F_d$ between upper F_u and lower F_d limit and rate constant κ. The well-known general solution is the logistic or sigmoid function (2) and (A2.2) with $\Delta = 0$) as solution of (A2.1) that exponentially approaches the lower and upper asymptotic limits with time constant $\tau = 1/\kappa$. Equation (A2.1) is the most simple version of a Bernoulli equation with time dependent coefficients $\kappa_1(t)$, $\kappa_2(t)$ for the linear and quadratic terms that allows for arbitrary time varying growth $\kappa > 0$ and decay rates $\kappa < 0$. An application for resource limited growth

under external disturbance was demonstrated in (Fürstenau et al. 2016). A stochastic Bernoulli-Langevin equation is obtained by adding a random noise term to (A2.1). By transformation into the equivalent Fokker–Planck stochastic partial differential equation (Risken 1988) with deterministic (Bernoulli) drift and stochastic (e.g. Gaussian) diffusion term it allows for modeling the time dependent dynamics of the probability density with mean $f(t)$. This procedure possibly could provide an approach to formally connect the stochastic wave discrimination theory of psychophysical laws (Link 1992) with our resource limitation-based power law that connects subjective ISA-WL with objective load measures.

Logistic ISA(n)-WL model

In what follows, we replace the time parameter of Eq. (A2.1) by the environmental (traffic) load variable n. In a more general form the logistic ISA(n, μ, ν)-WL Eq. (2) besides shift parameter m:= μ and scaling parameter τ:= ν includes an additional offset parameter Δ that allows for offset adaptation to different subjective and objective WL measures which are assumed to follow the mental resource limitation

hypothesis. In normalized form (for ISA-WL: division by the asymptotic resource limit $p(n) := I(n)/I_u$), the logistic model function is written as:

$$p(n) = \frac{1 + \Delta}{1 + \exp\left\{-\frac{n-m}{\tau}\right\}} - \Delta = \frac{1 + \Delta}{1 + k \exp\left\{-\frac{n}{\tau}\right\}} - \Delta$$
(A2.2)

with $k = \exp\{m/\tau\} = I_u (1+\Delta)/I_d - 1$. It is easily verified that the hypothesized ISA(n) characteristic in Sect. 4.1 with $m := \mu$ is expressed by $\Delta = 0$. RC(n) in Sect. 4.2 is defined by $m = 0$ and $\Delta = 1$ so that (A2.2) equals $\tanh(x)$ with $x = n/(2\rho)$ (see below). For ISA(n), the constant k can be expressed by the scale limits ($I_d = 1$, $I_u = 5$):

$$p(n = 0) = p_d = \frac{I_d}{I_u} = \frac{1}{1 + k}$$
(A2.3)

i.e. $k = I_u/I_d - 1 = \exp\{\mu/\nu\}$ yielding $\mu = \nu \ln(4)$. This reduces Eq. (A2.2) and Eq. (2)) to a one-parametric model for the nominal scenarios ($e = 0$). The asymptotic limits are given as $\lim I(n \to \infty) = I_u$ and $\lim I(n \to -\infty) = 0$ (with $n < 0$ as mathematical design aspect only, with no real meaning).

$$I(n) = \frac{5}{1 + 4 \exp\left\{-\frac{n}{\nu}\right\}}$$
(A2.4)

with maximum slope at $n = \mu$: $dI/dn = I'(n = \mu) = I_u/4\nu$. For the $e = 1$ scenarios, we set (with index e for $e = 1$) $\mu := \mu_e$, i.e. $k := k_e$ and $\nu := \nu_e < \nu$ (see Sect. 4.1):

$$I_e(n) = \frac{5}{1 + k_e \exp\left\{-\frac{n}{\nu_e}\right\}}$$
(A2.5)

The different shift and scaling parameters (μ_e, ν_e) of Eqs. (A2.4), (A2.5) for the scenarios with priority event generate the intersection at coordinates (n_x, $I_x(n_x)$) between $e = 0$ and $e = 1$ characteristics. Via a plausibility argument in Sect. 4.1, a theoretical estimate of n_x was obtained from the prior information on the traffic scenarios $n_1 \ldots n_4$: $n_x := 30$ (± 5), with $I_x = I(n_x) = 2.25$ (± 0.5). Setting equal Eqs. (A2.4), (A2.5) at $n := n_x$ yields for $k(e = 1) := k_e$, $\mu(e = 1) := \mu_e$

$$k_e = (I_u/I_{de} - 1) = \exp\{m_e/n_e\} = 4 \exp\{n_x(1/n_e - 1/n)\}$$
$$\mu_e = n_x + \nu_e(\ln(4) - n_x/\nu)$$
(A2.6)

with $k_e > k(e = 0) = 4$ and $I_{de} < I_d(e = 0) = 1$. Introducing k_e into Eq. (A2.5) (or replacing m in Eq. (A2.2) by μ_e and $\rho = 0$) yields

$$p_e(n) = \frac{1}{1 + \exp\left\{-\frac{n - \mu_e}{\nu_e}\right\}} = \frac{1}{1 + k_e \exp\left\{-\frac{n}{\nu_e}\right\}}$$
(A2.7)

Numerical estimates for k_e and μ_e based on prior estimates on intersection n_x and nominal scaling parameter ν provide a one-parameter equation also for the non-nominal case ($e = 1$). It is depicted in Fig. 1 (Sect. 4.1) as dotted curve with $\nu(\mu := 35) = 25.3$, $\nu_e := 20$, $\mu_e = 34$ and slope $I'_e (\nu_e) = 0.0625$ together with the solid curve for the nominal ($e = 0$) characteristic.

The generalized linear (semi-logarithmic) form of the logistic characteristic $y_p(n) = 1/\nu\, n - \ln(4) = a_g n + b_g$ (Eq. (3)) utilizes the normalized ISA-WL $p(n) := I(n)/I_u$ (with prior knowledge $I_u = 5$) together with the semi-logarithmic scale of the nonlinear transformation $P(\text{ISA}) := p(\text{ISA}) / (1 - p(\text{ISA}))$. The latter is obtained from Eq. (A2.2) by some basic algebraic operations. For the non-nominal scenarios ($e = 1$), the generalized linear form is obtained from Eq. (3) by replacing $\nu := \nu_e$ and $k := k_e$ yielding ($y = \ln(P(\text{ISA}))$):

$$y_{pe}(n) = \ln(P_e) = \frac{1}{\nu_e}(n - n_x) + \frac{n_x}{\nu} - \ln(4) = a_{ge}n + b_{ge}$$
(A2.8)

With slope $a_{ge} = 1/\nu_e$ and intercept $y_{pe}(n = 0) = b_{ge} = n_x\left(\frac{1}{\nu} - \frac{1}{\nu_e}\right) - \ln(4)$, yielding $y_{pe}(n = n_x) = y_p(n = n_x) = y_{px} = n_x/\nu_t - \ln(4) = -0.20$.

Using the plausibility argument for n_x as prior knowledge (see Sect. 4.1), the intersection coordinates between nominal ($e = 0$) and priority ($e = 1$) characteristics for the generalized linear model are $\{n_x, y_{px}\} = \{n_x, a_g n_x - \ln(4)\} = \{30, -0.20\}$, with uncertainty $\delta\mu := \pm 5$ yielding $\{\delta n_x, \delta y_{Ix}\} = \{\pm 5, +0.43/ -0.32\}$. If for the regression analysis of the priority event scenarios ($e = 1$) no a-priori knowledge on n_x is assumed and both slope a_{ge} and intersection parameter b_{ge} have to be estimated from the data, the intersection coordinates are obtained by combination of Eqs. (3), (A2.8) as:

$$\{n_x, y_{px}\} = \left\{ \frac{(b_{ge} - b_g)}{(a_g - a_{ge})}, a_g n_x - \ln(4) \right\}$$
$$= \left\{ \frac{(-\mu_e/\nu_e + \ln(4))}{(1/\nu - 1/\nu_e)}, a_g n_x - \ln(4) \right\}$$
(A2.9)

with $b_g := b_{gt} = -\ln(4) = -1.3863$. With the regression-based parameter estimates ($a_g = 0.0380$ (4 10^{-4}), $a_{ge} = 0.0471$ (1.3 10^{-3}), $b_{ge} = -1.670$ (0.05)) we get $n_x = 31.2$ (0.56) that confirms the initial estimate (30 ± 5), however, with uncertainty reduced by factor 10.

Linearized logistic ISA model

The results in Sects. 4.1, 5.1 show that the nearly linear part of the logistic characteristic may be sufficiently

represented by a linearized model for the selected traffic flow range $25 \leq n \leq 55$. In the vicinity of the logistic shift parameter μ with $n = \mu + \Delta n$, $|\Delta n| = |n - \mu| << \mu$, $I(n = \mu) = I_u/2$ the logistic characteristic is approximated by

$$I(n) \approx b_{lt} + a_{lt}n \approx \frac{I_u}{2}\left(1 - \frac{\mu}{2\nu}\right) + \frac{I_u}{4\nu}n$$
$$= \frac{I_u}{2}\left(1 - \frac{\ln(4)}{2}\right) + \frac{I_u \ln(4)}{4\mu}n$$

where the slope in the vicinity of μ, $dI/dn \approx a_{lt} = I_u/(4\nu) = I_u \ln(4)/(4\mu)$ is obtained through neglection of quadratic and higher powers of the exponents $(n - \mu)/\nu$.

Within the linear approximation a dimensionless index $s_{ak}(a_k, b_k)$ can be defined for individual participants $k = 1 \dots 21$ via slope and intercept (a_k, b_k) by normalization of the independent and dependent variables $(n, I(n))_k$ through division by the mean value of the traffic variable n $(n_1 + n_4)/2$ and the individual ISA intervals $(I_{max} + I_{min})/2$, respectively. In what follows, we skip the participant index k. Normalized variables are written with small letters or index n, yielding $n_n = \frac{n}{(n_1 + n_4)/2} = \frac{n}{40}$, and for the individual ISA scenario means $< ISA >_{jk}$:

$$\langle isa \rangle = \frac{\langle ISA \rangle}{\frac{\langle ISA \rangle_{max} + \langle ISA \rangle_{min}}{2}} = \frac{ISA}{40a + b} = s_b + s_a n_n$$

where the individual maximum and minimum ISA values are taken from the linear regression predictions at the maximum n_4 and minimum n_1 of the predictor variable. This yields the dimensionless linear sensitivity index

$$s_a = \frac{isa_{max} - isa_{min}}{n_{n\,max} - n_{n\,min}} = \frac{1}{1 + \frac{b}{40a}}$$

with anticorrelation $s_b = 1 - s_a$. s_a allows for clustering of participants k with individual sensitivities a_k as used in (Radüntz et al. 2020a, b) for discriminating between subjects of low and high WL sensitivity.

Logistic radio calls frequency model RC(*n*)

The hypothesized logistic radio calls characteristic $R(n)$ is obtained from Eq. (A2.2) by setting the offset parameter $\Delta := 1$ yielding Eq. (5). With normalized call rate $s := R(n)/R_u$ and $x := n/(2\rho)$ Eq. (5) can be written in the form $s(x) = (1 - \exp(-2x))/(1 + \exp(-2x)) = (\exp(x) - \exp(-x))/(\exp(x) + \exp(-x))$ which is the definition of the hyperbolic function tanh(x)

$$s(n) = \left[\frac{2}{1 + \exp\{-n/\rho\}} - 1\right] = \tanh\left(\frac{n}{2\rho}\right) \quad (A2.10)$$

With turning point at the origin $(n, s) = (0, 0)$ it starts with slope $s'(n) = ds/dn$ given by

$$s' = \frac{1}{2\rho cosh^2(n/2\rho)} \quad (A2.11)$$

Maximum slope (at $n = 0$) is $s'(n = 0) \approx \Delta s/\Delta n = 1/2\rho$. With a linear extrapolation of the maximum slope at $n = 0$, $\Delta s = 1$ and $\Delta n := $ capacity limit $n_c = n_3$ yield the theoretical estimate for the radio calls scaling parameter $\rho_t : \approx n_c/2 = 22.5$.

Like for ISA-WL, the generalized linear form for $R(n)$ is obtained via nonlinear transformation S(s) of the normalized TL-variable s(n) that is obtained from Eq. (5) or (A2.10) by some basic algebraic operations: $S(s) = (1 + s(n)) / (1 - s(n))$, followed by taking the logarithm:

$$y_s(n) = ln(S) = \frac{1}{\rho}n \quad (A2.12)$$

With slope $a_{gs} = 1/\rho$ and intersection between nominal and non-nominal characteristics at the origin $n_{xs} = 0$.

Psychophysics power law model ISA(RC)

In the simulation experiment with external load variable n [AC/h] defining simulation scenarios, the function pair $(R(n), I(n))$ defines the parametric dependence of the subjectively perceived and reported ISA(n) WL on objectively measured frequency of radio calls (R(n)). To derive the hypothesized psychophysical power law relationship between subjective response P(ISA(*n*)) (representing fraction of used cognitive processing resources) and objective stimulus measure (S(RC(*n*))), the ISA(*n*)-WL scale $(1 \leq I(n) \leq 5)$ and the metric TL RC(*n*) measure [number of radio calls / h] have to be normalized and nonlinearly transformed into (S, R) (see Eqs. (A2.8), (A2.12)). The suitable transformations for the ISA-WL and RC-TL variables have been derived before for the generalized linear forms of the logistic characteristics. The power law may be derived by combining these linear versions. A more general expression is obtained by transformation of Eq. (A2.2) into the generalized linear form that includes a general bias term Δ and the general scaling and shift parameters τ and m, respectively:

$$y(n) = ln\left(\frac{\Delta + p(n)}{1 - p(n)}\right) = ln(P(n)) = \frac{1}{\tau}n - \frac{m}{\tau} \quad (A2.13)$$

Based on this general expression, we define separate equations for stimulus ($y_s(n)$) and perception (of mental load) ($y_p(n)$)

$$y_s(n) = ln\left(\frac{\Delta_s + s(n)}{1 - s(n)}\right) = ln(S(n)) = \frac{1}{\varrho}n - \frac{m_s}{\varrho} \quad \text{(A2.14)}$$

$$y_p(n) = ln\left(\frac{\Delta_p + p(n)}{1 - p(n)}\right) = ln(P(n)) = \frac{1}{\nu}n - \frac{m_p}{\nu} \quad \text{(A2.15)}$$

By solving Eq. (A2.14) for n and introducing the expression into (A2.15), we obtain the generalized linear form of the power law:

$$y_p(y_s) = ln(P(n)) = \gamma y_s + \frac{m_s - m_p}{\nu} = \gamma y_s + b_s \quad \text{(A2.16)}$$

For the nominal scenarios ($e = 0$) with $m_s = 0$ and $m_p = \mu$ we have $b_s = b_{gt} = -ln(4)$ so that (A2.16) = Eq. (8) as depicted in Fig. 3 of Sect. 4.3, i.e. a one-parameter model for regression-based estimate of γ. For the non-nominal scenarios due to the lack of prior information on $b_{se} = b_s(e = 1)$, a two-parameter fit is required. With exponent γ of the power law for ISA(RC) defined by the ratio of the scaling coefficients of stimulus ρ and perception ν: $\rho/\nu := \gamma$ and $m_s = 0$, $m_p = \mu$ the linear equation in log–log coordinates (A2.16) may be written in in linear coordinates

$$P(S) = exp\left\{-\frac{\mu}{\nu}\right\}S^\gamma = \frac{1}{k}S^\gamma \quad \text{(A2.17)}$$

In correspondence to the intersection $\{n_x, I_x = \text{ISA}(n_x)\}$ of the nominal and non-nominal ISA(n) characteristics the intersection $\{s_x, p_x\}$ for the (normalized) $e = 0$ and $e = 1$ power law equations is derived from Eqs. (3, 5) as

$$\{s_x, p_x\} = \left\{ \tanh\left(\frac{n_x}{2\rho}\right), \left(1 + 4\exp\left(-n_x/\nu\right)\right)^{-1} \right\} \quad \text{(A2.18)}$$

For the generalized linear form in (log–log)-coordinates of the power law (Eq. (A2.16)), the intersection coordinates are obtained from Eqs. (3) and (A2.12) with $n = n_x$:

$$\{y_{sx}, y_{px}\} = \left\{ \frac{n_x}{\rho}, \frac{n_x}{\nu} - ln(4) \right\} \quad \text{(A2.19)}$$

Based on the combination of Eq. (A2.16) for $e = 0$ and $e = 1$ at crossing coordinate $n = n_x$ (using $y_p(n_x) = y_{pe}(n_x)$), a theoretical relationship between the $e = 1$ intersection parameter b_{se} and the unknown power law exponent γ_e in this case is derived as:

$$y_{pe}(y_{se} = 0) = b_{se} = \left(\frac{1}{\nu} - \frac{\gamma_e}{\rho}\right)n_x - ln(4) \quad \text{(A2.20)}$$

With prior estimate $n_x = 31.2$ (± 0.6) obtained from the logistic ISA(n) regressions for $e = 0, 1$, we can use this relationship for a one-parameter regression estimate of γ_e ($e = 1$)

with the generalized linear form (A2.16) of the power law model:

$$y_{pe} = \gamma_e\left(y_{se} - y_{sx}\right) + y_{px} = \gamma_e y_{se} + n_x\left(\frac{1}{\nu} - \frac{\rho_e}{\varrho\nu_e}\right) - ln(4) \quad \text{(A2.21)}$$

with $y_{sx} = \gamma_e n_x/\rho = \rho_e n_x/(\nu_e \rho)$ and $y_{px} = n_x/\nu - ln(4)$, where we use the theoretical relationship $\gamma_e = \rho_e/\nu_e$ and logistic regression parameter estimates as prior information.

Vice versa, without prior estimate of n_x, we may use the parameter estimates (γ, γ_e, b_{se}, and $b_s = -ln(4)$) from the one-parameter fit for $e = 0$ and the two-parameter fit for $e = 1$ to quantify the intersection of the linear $e = 0, 1$ characteristics from the two linear Eq. (A2.16):

$$\{y_{sx}, y_{px}\} = \left\{ \frac{b_{se} + ln(4)}{\gamma - \gamma_e}, \frac{b_{se}\gamma + ln(4)\gamma_e}{\gamma - \gamma_e} \right\} \quad \text{(A2.22)}$$

References

Abich J IV, Reinerman-Jones L, Taylor G (2013) Establishing workload manipulation utilizing a simulated environmen. In: Shumaker R (ed) Lectur Notes Computer Science (LNCS. Springer, Berlin, pp 211–220

Aricó P, Borghini G, Flumeri GD, Sciaraffa N, Babiloni F (2018) Passive BCI beyond the lab:current trends and future directions. Physiol Meas 39:08TR02. https://doi.org/10.1088/1361-6579/aad57e

Averty P, Collet C, Dittmar A, Athenes S, Vernet-Maury E (2004) Mental workload in air traffic control: an index constructed fom field tests. Aviat Space Environ Med 75:333–341

Averty P, Guittet K, Lezaud P (2008) An ordered logit model of air traffic controllers conflict risk judgement. Air Traffic Control Q 16(2):101–125

Bachelder E, Godfroy-Cooper M (2019) Pilot workload esimation: synthesis of spectral requirements analysis and Weber's law. SCL Tech, San Diego

Birbaumer N, Schmidt RF (2010) Biologische Psychologie, 7th edn. Springer Medizin Verlag, Heidelberg

Brennan S (1992) An experimental report on rating-scale descriptor set for the instanteneous self assessment (ISA) recorder. Technical report, DRA Maritime Command and Control Devision, Portsmouth

Brookings JB, Wilson GF, Swain CR (1996) Psychophysiological responses to changes in workload during simulated air traffic control. Biol Psychol 42:361–377

Buntain C (2012) Psychophysics and just noticable difference. (D. o. University of Maryland, Ed.) Report 4, CMSC828D.

Corver SC, Unger D, Grote G (2016) Predicting air traffic controller workload: Trajectory uncertainty as the moderator of the indirect effect of traffic density on controller workload through traffic conflict. Hum Factors 58(4):560–573. https://doi.org/10.1177/0018720816639418

Djokic J, Lorenz B, Fricke H (2010) Air traffic control complexity as workload driver. Transp Res Part C 18:930–936

Fürstenau N, Mittendorf M (2016) Bernoulli-Langevin wind speed model for the simulation of storm events. Z. Naturforsch. 71(12) a:1167–1173. https://doi.org/10.1515/zna-2016-0238

Fürstenau N, Radüntz T, Mühlhausen T (2020) Model based development of a mental workload sensitivity index for subject clustering. Theor Issues Ergon Sci 21(6):684–708. https://doi.org/10.1080/1463922X.2020.1711990

Girard J, Wilczyk M, Barloy Y, Simon P, Popieul J (2005) Towards an on-line assessment of subjective driver workload. In: I. C. University of Iowa (Ed), Proceedings of Driving Simulation Conference, North America 2005, pp 382–391. Orlando, Fl. Retrieved 2020, from http://www.nads-sc.uiowa.edu/dscna/2005/papers/Towards_an_on-line_assessment_subjective_driver_workload.pdf

Gopher D, Braune R (1984) On the psychophysics of workload: why bother with subjective measures. Hum Factors 26(5):519–532

Gopher D, Chillag N, Arzi N (1985) The psychophysics of workload - a second look at the relationship between subjective measures and performance. In: Proceedings of the Human Factors Society (29th Annual Meeting), pp 640–644

Hancock P, Desmond PA (eds) (2001) Stress, workload, and fatigue. Lawrence Erlbaum Publishers, Mahwah

Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): result of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) Human mental workload. Springer, Berlin, pp 139–183

ISO-10075 (1991) Ergonomic principles related to mental workload-general terms and definitions.

Johannsen G, Morey N, Pew R, Rasmussen J, Sanders A, Wickens C (1979) Final report of experimental psychology group. In: Morey N (ed) Mental workload its theory and measurement. Springer, Berlin, pp 101–114

Jordan C (1992) Experimental study of the effect of an instantaneous self assessment workload recorder. Technical report, DRA Maritime Command and Control Division, Portsmouth

Kahnemann D (1973) Attention and effort. Prentice Hall, Englewood Cliffs

Kirwan B, Evans A, Donohoe L, Kilner A, Atkinson LT, MacKendrick H (1997) Human factors in the atm system design life cycle. In: Chit M (ed) FAA/Eurocontrol ATM R&D seminar. Eurocontrol, Paris

Lee PU (2005) A non-linear relationship between controller workload and traffic count. Proc Hum Factors Ergon Soc. https://doi.org/10.1177/154193120504901206

Lee PU, Mercer J, Smith N, Palmer E (2005) A non-linear relationship between controller workload, task load, and traffic density: the straw that broke the camel's back. In: Proceedings of International Symposium Aviation Psychology, pp 438–444

Lehrer P, Karavidas M, Lu S-E, 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. Int J Psychophysiol 76:80–87. https://doi.org/10.1016/j.ijpsycho.2010.02.006

Lei S, Roetting M (2011) Influence of task combination on EEG spectrum modulation for driver workload estimation. Hum Fact 53:168–179. https://doi.org/10.1177/0018720811400601

Link SW (1992) The wave theory of difference and similarity. Lawrence Erlbaum Associates and Routledge, New York

Loft S, Sanderson P, Neal A, Mooij M (2007) Modeling and predicting mental workload in en route air traffic control: Critical rewiew and broader implication. Hum Factors 49:376–399

Manning C, Mills S, Fox C, Pfleiderer E, Mogilka H (2001) The relationship between air traffic control communication events and measures of controller taskload and workload. In: Proceedings of 4th USA/Europe Air Traffic Management R&D Seminar. Santa Fe, NM: FAA / Eurocontrol. Retrieved 3 10, 2021, from https://atmseminar.eurocontrol.fr/past-seminars/4th-seminar-santa-fe-nm-usa-december-2001/papers/paper-161

Mühlhausen T, Raduentz T, Tews A, Guerluek H, Fürstenau N (2018) Research design to access the mental workload of air traffic controllers. Proceedings of 1st International Conference of Human Systems Engineering and Design: Springer AISC advances in intelligent systems and computing. CHU Université de Reims Springer, Berlin, pp 415–421

Mulder G, Mulder LJ (1981) Information processing and cardiovascular control. Psychophysiology 18:392–402. https://doi.org/10.1111/j.1469-8986.1981.tb02470.x

Mulder L, Waard D, Brookhuis K (2004) Estimating mental effort using heart rate variability. In: Stanton N, Hedge A, Brookhuis K, Salas E, Hendrick H (eds) Handbook of human factors and ergonomics methods. CRC Press, London, pp 201–208

Norwich KH (1987) On the theory of Weber fractions. Percept Psychophys 42(3):286–298

Norwich KH, Wong W (1997) Unification of psychophysical phenomena: the complete form of Fechner's law. Percept Psychophys 59(6):929–940

Parasuraman R, Hancock PA (2001) Adaptive control of mental workload. In: Hancock PA, Desmond PA (eds) Stress, workload, and fatigue. Lawrence Erlbaum Associates, Mahwah, pp 305–320

Prinzel LJ III, Parasuraman R, Freeman FG, Scerbo MW, Mikulka PJ (2003) Three experiments examining the use of electroencephalogram, event-related potentials, and heart-rate variability for real-time human-centered adaptive automation. NASA/TP-2003-212442, NASA Langley Research Center, Hampton

Radüntz T (2016) Kontinuierliche Bewertung psychischer Beanspruchung an informationsintensiven Arbeitsplätzen auf Basis des Elektroenzephalogramms. Dissertation. (Mathematisch-Naturwissenschaftliche Fakultät, Hrsg). Berlin, Humboldt-Universität zu Berlin, Berlin

Radüntz T (2017) Dual frequency head maps: a new method for indexing mental workload continuously during execution of cognitive tasks. Front Physiol 8:1–15. https://doi.org/10.3389/fphys.2017.01019

Radüntz T, Fürstenau N, Tews A, Rabe L, Meffert B (2019) The effect of an exceptional event on the subjectively experienced workload of air-traffic controllers. In: Longo L, Leva MC (eds)

Human mental workload: models and applications. Springer International Publishing, Amsterdam

Radüntz T, Fürstenau N, Mühlhausen T, Meffert B (2020a) Indexing mental workload during simulated air traffic control tasks by means of dual frequency head maps. Front Physiol 11:300. https://doi.org/10.3389/fphys.2020.00300

Radüntz T, Mühlhausen T, Freyer M, Fürstenau N, Meffert B (2020b) Cardiovascular biomarkers' inherent timescales in mental workload assessment during simulated air traffic control. Appl Psychophysiol Biofeedback. https://doi.org/10.1007/s10484-020-09490-z

Reid G, Potter S, Bressler J (1989) Subjective workload assessment technique (SWAT): a user's guide. American National Standards Institute, New York

Risken H (1988) The Fokker-Planck equation, 2nd edn. Springer, Berlin

Sheridan T (1980) Mental workload-what is it? Why bother with it? Hum Factors Soc Bull 23:1–2

Sperandio JC (1978) The regulation of working methods as function of workload among air traffic controllers. Ergonomics 21(3):195–202

Stein E (1985) Air traffic controller workload: An examination of workload probe. DOT/FAA/CT-TN84/24, DOT/FAA, Atlantic City

Stevens SS (1957) On the psychophysical law. Psychol Rev 64(3):153–181

Stevens SS (1975) Psychophysics: introduction to its perceptual, neural and social prospects. Wiley, New York

Tattersall AJ, Foord PS (1996) An experimental evaluation of instantaneous self-assessment as ameasure of workload. Ergonomics 39(5):740–748

Vanderhaegen F, Wolff M, Mollard R (2020) Non-conscious errors in the control of dynamic events synchronized with heartbeats: anew challenge for human reliability study. Saf Sci. https://doi.org/10.1016/j.ssci.2020.104814

Wickens C (2002) Multiple resources and performance prediction. Theor Issues Ergon Sci 3(2):159–177. https://doi.org/10.1080/14639220210123806

Wickens CD, Hollands JG (2000) Attention, time sharing, and workload. Engineering psychology and human performance, 3rd edn. Prentice-Hall, Upper Saddle River, pp 439–479

Xie B, Salvendy G (2000) Review and reappraisal of modelling and predicting mental workload in single and multitask environments. Work Stress 14(1):74–99

Yeh YY, Wickens C (1984) The dissociation of subjective measures of mental workload and performance. techreport, Engineering-psychology Research Laboratory, University of Illinois at Urbana-Champaign.