Dynamic Workload Assessment in Complex Situations: DYWA

DLR-IB-FL-BS-2026-89

Document properties

Title	Dynamic Workload Assessment in Complex Situations: DYWA
Subject	Development of a workload instrument for complex and dynamic situations
Institute	Flight Guidance
Authors	Gampe, Anja; Hamann, Anneke; Carstengerdes, Nils
Classification	A/I (open)
IB Number	DLR-IB-FL-BS-2026-89
DOI Report	www.doi.org/10.60575/mws3-6f54
DOI Materials	www.doi.org/10.5281/zenodo.22890862
Date	19/10/2026
Version	1.0

Contents

Abstract ____ 4

1. Introduction ____ 5

2. Construction of the Dynamic Workload Assessment (DYWA) ____ 7

2.1. Design Requirements 7

2.2. Instrument Development 8

2.3. The Dynamic Workload Assessment 9

2.3.1. Instrument Structure 9

2.3.2. Qualitative Event Annotation 10

3. Administration Procedure ____ 11

4. Data Processing and Analysis ____ 12

5. Evaluation of Psychometric Properties ____ 15

6. Potential Applications ____ 16

7. Conclusion ____ 17

References ____ 18

Appendix ____ 20

Abstract

In this document we present the new DYnamic Workload Assessment: DYWA. Currently available instruments for self-reported workload assessment are either applied retrospectively or at static intervals during a situation without capturing the temporal dynamics of workload and the stressors that operators react to. The DYWA is an easy-to-administer and non-intrusive instrument. We recommend using this instrument in studies to gain more insight into I) the temporal dynamics of workload, II) specific stressors of situations, III) their association with the corresponding workload, and IV) concurrent measures like performance, communication or physiological responses.

1. Introduction

Workload is an important concept in human factors and ergonomics because it directly influences operator performance, safety, and well-being (Hart, 1986; Young et al., 2015). It describes the demands imposed by the work environment and the extent to which these demands require the allocation of cognitive, physical, and perceptual resources. Contemporary conceptualizations define workload as the stress created by task and environmental demands that elicits a corresponding strain response in the individual (Babaei et al., 2025; Bartolomei et al., 2026; Longo et al., 2022). Within this framework, workload is not an inherent property of a task but rather emerges from the interaction between external demands and the operator's available resources, skills, and adaptive capacity (Gopher & Donchin, 1986; Hancock, 1989). Consequently, identical task demands may produce different levels of experienced workload depending on the individual, the operational context, and the temporal evolution of the task.

Because workload is a latent and dynamic construct, its assessment has been the focus of human factors research for several decades. A wide variety of measurement approaches has been developed, each emphasizing different aspects of workload. The multidimensionality of the construct complicates the assessment: Rather than considering workload as a single outcome variable, Xie and Salvendy (2000) argued that comprehensive assessment requires multiple complementary indicators describing both its temporal characteristics and its functional consequences. Specifically, they proposed that workload can be characterized through instantaneous workload, representing the workload at a given moment; average workload over a task period; accumulated workload resulting from prolonged exposure; peak workload during periods of highest demand; overall workload summarizing the complete task experience; effective workload associated with productive resource allocation; and ineffective workload reflecting resource expenditure that does not contribute to task performance. This framework illustrates that workload is inherently conceptualized dynamically and that different indicators provide distinct yet complementary information about operator state.

Despite the broad consensus that workload is multidimensional and evolves over time, many existing assessment methods remain associated with practical limitations (Boumann et al., 2023). Performance-based measures are a well-known and often utilized way of assessing workload. Changes in workload are inferred from changes in task performance, such as response times, error rates, or secondary-task performance (Wickens, 1981). Although these approaches provide valuable insights into operator functioning and may be used to construct temporal dynamics of workload, they may also reflect compensatory strategies rather than workload itself and are often difficult to interpret in isolation (Martins, 2016; Wickens, 2008). Moreover, in the context of higher levels of automation, participants or operators, whose task is to monitor the automation, may not display enough measurable behavior. Changes in the cognitive processes underlying workload may go unnoticed.

Physiological measures, including heart rate variability, electroencephalography, electrodermal activity, and pupillometry, can provide objective indicators of mental and physical effort as well as a continuous assessment with high temporal resolution (Charles & Nixon, 2019). Yet, they often require specialized equipment, complex data processing and can be confounded with task characteristics, other mental states or arousal-related processes (Fairclough, 2009; Hamann & Carstengerdes, 2026; Mansikka et al., 2016; Tattersall & Foord, 1996; Zhao et al., 2025).

Subjective assessment techniques remain among the most widely used approaches because they are inexpensive, easy to administer, and capture the operator's own perception of task demands. There are instruments available that aim to capture the multidimensionality of workload, such as the NASA Task Load Index (NASA-TLX, Hart, 1986), which evaluates workload across six dimensions: mental demand, physical demand, temporal demand, effort, performance, and frustration. Another example of multidimensional

workload assessment is the DLR Workload Assessment Tool (DLR-WAT, Brandenburger et al., 2023; Grippenkoven et al., 2018). Adapted from the NASA-TLX, it divides mental workload further into information acquisition, knowledge retrieval, decision-making and revises the scale between underload and overload. The Subjective Workload Assessment Technique (SWAT, Tsang & Velazquez, 1996) was developed to assess workload using different dimensions like perceptual, verbal, spatial, visual, or manual task demands, thereby recognizing that different task characteristics contribute differently to the overall workload experienced by an operator. While capturing different aspects of the workload resulting from a task, none of these measures is able to capture its temporal dynamics since the ratings are always done post-hoc and globally for the situation or simulation run. The Instantaneous Self-Assessment Scale (ISA: Tattersall & Foord, 1996) is a well-known instrument in aviation research. It is specifically designed to provide information on workload changes during a dynamic situation, consists of only one item that is rated on a 1-5 Likert scale, and is thus quick to administer during ongoing experiments, running simulations or realistic tasks. The ISA rating is usually administered every x minutes and can thus provide an estimation of temporal dynamics of workload, but not of its multidimensional composition or relation to certain task characteristics. It thus risks oversimplifying a complex dynamic signal by forcing a discrete, linear interpretation of what is often a nonlinear, fluid process. Moreover, interruptions can fragment the operator's attention, introduce artificial workload spikes, or miss critical transitions in workload state. A combination of multiple workload rating scales during and after the task may combine the strengths and alleviate the weaknesses of each individual measure, but add to the burden on the participant.

In sum, performance measures are informative, but not always available and possibly influenced by additional factors like strategy use. Physiological data provide a continuous assessment, but are cumbersome and prone to confounding with other cognitive and physiological states. Self-report is easy to administer, but either unidimensional or global, lacking temporal dynamics altogether. Consequently, there remains a need for assessment approaches that preserve the multidimensional nature of workload while reducing measurement complexity. Developing simple measures capable of capturing key temporal and qualitative characteristics of workload may therefore facilitate both experimental research and real-world human-system evaluation.

2. Construction of the Dynamic Workload Assessment (DYWA)

2.1. Design Requirements

To address the limitations of existing workload assessment approaches, a set of design requirements was derived from the human factors literature. These requirements aim to ensure that the assessment method captures the complexity of workload while remaining applicable in realistic operational settings.

First, the method should account for the **multidimensional nature of workload** by integrating both quantitative and qualitative assessment approaches. Although quantitative measures provide information about workload intensity (Brandenburger et al., 2023; Hart, 1986; Tattersall & Foord, 1996; Zak et al., 2020), they offer limited insight into the underlying causes of workload (Babaei et al., 2025; Bartolomei et al., 2026; Longo et al., 2022). Complementary qualitative data are therefore required to identify the specific stressors, task characteristics and demands experienced by individual participants, supporting a more comprehensive interpretation of workload.

Second, the method should enable **continuous quantitative assessment** throughout the course of a task or situation. In contrast to retrospective assessments, continuous measurement allows the analysis of temporal workload dynamics (Guastello, 2017; Thelen, 2005), including instantaneous, average, peak, and cumulative workload (Xie & Salvendy, 2000). Capturing these dynamics is essential because workload fluctuates over time and critical workload peaks may not be reflected in a single overall rating.

Third, workload assessment should be **non-intrusive**, allowing data to be collected without interrupting task execution or altering the natural progression of the situation. Minimizing interference is particularly important in dynamic and safety-critical environments, where interruptions may influence both operator performance and the workload being measured.

Furthermore, the method should be **practical and easy to administer**, requiring minimal participant training and imposing only a small time burden on both participants and researchers. Low implementation effort facilitates the application of the method in experimental as well as operational contexts.

Finally, the method should demonstrate **good to excellent psychometric properties**, including reliability, validity, and sufficient sensitivity to detect meaningful differences and changes in workload. Robust psychometric performance is a prerequisite for producing valid and interpretable workload measurements and for enabling meaningful comparisons across participants, tasks, and experimental conditions.

Based on these considerations, the following design requirements were defined: (1) capture the multidimensional nature of workload through the integration of quantitative and qualitative measures; (2) provide continuous quantitative assessment enabling the extraction of metrics like instantaneous, average, peak, and cumulative workload; (3) identify participant-specific workload sources and stressors; (4) support non-intrusive data collection without interrupting task execution; (5) require minimal training and administration effort; and (6) exhibit good to excellent psychometric properties.

2.2. Instrument Development

The DYWA was developed to capture both the temporal development of perceived mental workload and the contextual factors underlying workload changes. The DYWA is based on a two-dimensional visual scale and enables participants to retrospectively reconstruct the temporal development of their perceived mental workload over the course of a completed task or situation without relying on external prompts. Unlike traditional post-task questionnaires, the instrument provides a graphical representation of workload dynamics while also allowing qualitative annotation of workload-related events.

The DYWA was developed by systematically translating the identified design requirements into specific design features. Each requirement informed one or more aspects of the instrument's structure, content, and administration procedure.

Taking the ISA (Tattersall & Foord, 1996) as a starting point because of its easy administration, high face validity and proven utility in applied research, the newly designed instrument was based on the same data range (1-5) including the same verbal anchors. This has the added advantage of comparability between studies using ISA and DYWA.

To capture the multidimensional nature of workload that ISA lacks, DYWA combines a continuous quantitative workload rating with a qualitative component. The quantitative rating provides a standardized measure of perceived workload over time, whereas the qualitative component allows participants to document the specific factors that contributed to elevated workload. This combination enables both the quantification of workload intensity and the interpretation of its underlying causes.

To enable the assessment of workload dynamics, the quantitative component was designed to support measurements throughout the whole task execution time span. Consequently, workload can be analyzed as a time series, allowing the derivation of instantaneous, average, peak, and cumulative workload measures rather than relying solely on a single retrospective rating.

The newly developed qualitative component was designed to collect participant-specific stressors whenever participants perceived an increase or decrease in workload. The responses can be restricted to a list of predefined categories or can be used without restriction allowing participants to report stressors in their own words, thereby capturing context-specific demands that may not be represented in standardized questionnaires.

To avoid interrupting the ongoing task and minimize interference with task performance, the assessment procedure was designed to require only a brief retrospective interaction. The quantitative rating can be completed within a few seconds, while qualitative information is collected by adding a few words, thereby reducing the burden on participants.

Ease of administration was considered throughout the development process. The instrument uses an intuitive response format that requires little explanation, allowing participants to use it without extensive training. In addition, the assessment procedure was designed to be easily integrated into experimental protocols and operational settings.

Finally, the instrument was designed to enable robust psychometric evaluation. Standardized quantitative ratings allow the assessment of reliability, validity, and sensitivity, while the qualitative component complements the numerical data by supporting the interpretation of workload changes without replacing quantitative measurement.

2.3. The Dynamic Workload Assessment

The DYWA constitutes a novel way of capturing and analyzing temporal dynamics of workload. Its continuous workload profile and the event annotations enable the simultaneous assessment of temporal workload dynamics and participant-specific workload sources within a single instrument. Figure 1 provides an aviation-specific example.

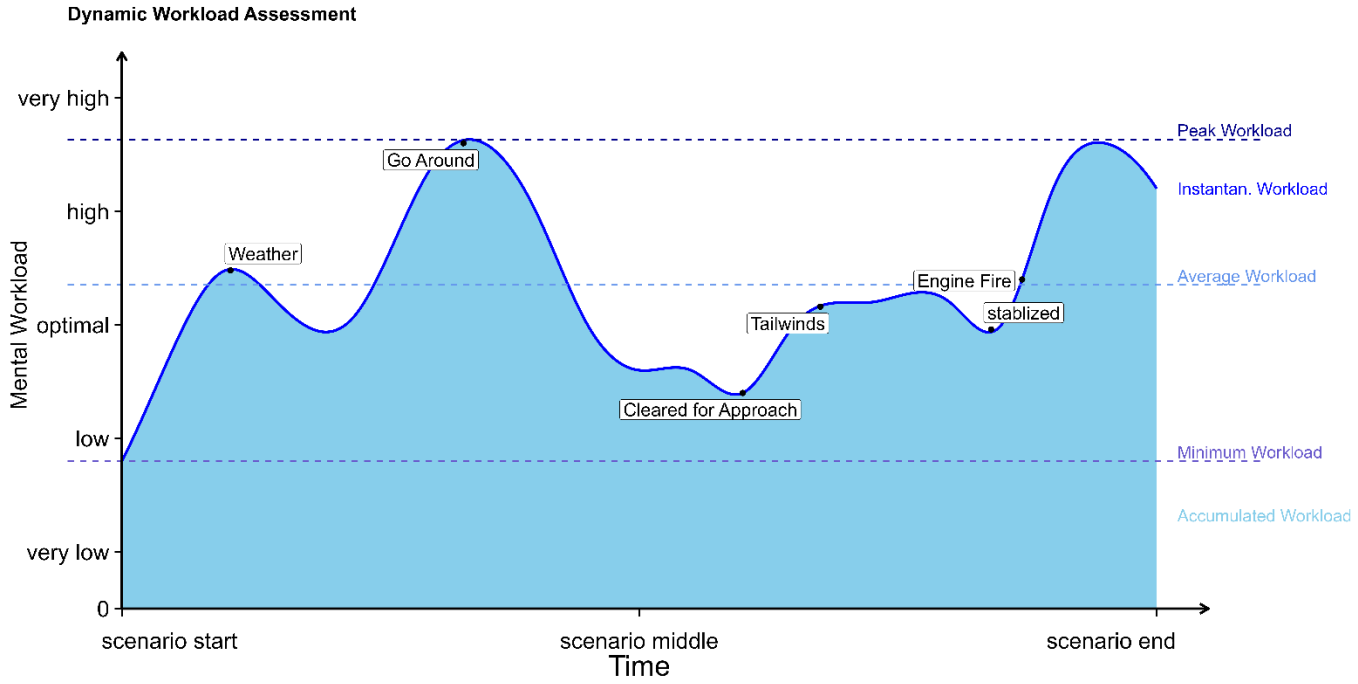


Figure 1: Dynamic Workload Assessment filled with example annotations for a pilot flying a complex approach.

2.3.1. Instrument Structure

The instrument consists of a two-dimensional coordinate system that can be presented on paper or implemented digitally. The **horizontal axis (x-axis)** represents the temporal progression of the task or situation, extending from the beginning on the left to the end on the right. This axis allows participants to express changes in perceived workload over time rather than providing a single overall rating. The **vertical axis (y-axis)** represents the participant's subjective mental workload. The scale is aligned with the ISA and ranges from *Very Low* at the lower end, through *Optimal* in the middle, to *Very High* at the upper end

(Tattersall & Foord, 1996). Participants indicate their perceived workload by drawing a continuous workload profile across the duration of the task, thereby visualizing fluctuations in workload over time.

2.3.2. Qualitative Event Annotation

To complement the quantitative workload profile, the instrument includes an event annotation component. Participants can indicate workload-relevant events by placing an arrow at the corresponding point on the workload profile. The arrow identifies both the approximate time of occurrence and the associated workload level. Each annotated event is accompanied by a short free-text description of the perceived workload source (e.g., *unforeseen disruption, time pressure, or complex decision*), and may be restricted to a list of predefined categories by the experimenter. These annotations provide qualitative information about the factors contributing to changes in workload and facilitate the interpretation of the quantitative workload trajectory.

3. Administration Procedure

The dynamic workload assessment can be administered in either paper-based or digital form (e.g., on a tablet, touchscreen, or digital whiteboard). Before administration, the template is adapted to the respective task or operational scenario by providing a descriptive title and, where appropriate, contextual labels that facilitate temporal orientation while maintaining a consistent scale structure across applications. See the Appendix for generic examples and templates in German and English.

The printable templates, a digital html-based version of the DYWA as GUI and an R script to generate own examples and templates are available at www.doi.org/10.5281/zenodo.22890862.

Prior to the assessment, participants receive standardized instructions explaining that the workload curve should represent their subjective perception of mental workload throughout the completed task or situation. Participants are informed that there are no correct or incorrect responses and that the curve should reflect their individual workload experience rather than an objectively expected pattern.

To avoid interference with task performance, the workload profile is drawn and annotated immediately after task completion / scenario end. Participants retrospectively reconstruct the temporal development of their workload by drawing a continuous curve across the two-dimensional scale. Conducting the assessment as closely after the task as possible minimizes recall bias while preserving the uninterrupted execution of the task. A brief transition period between task completion and the assessment may be provided to allow participants to consolidate their recollection of the workload progression.

When using the paper-based version, different colours may be used for the workload curve and event annotations to improve readability during subsequent analysis. In digital implementations, these elements can be visually distinguished using different graphical representations, e.g. lines and text boxes.

For tasks with a clearly defined temporal structure, deviating from the example provided in this report, the x-axis may be divided into predefined task phases (e.g., *Preparation, Execution, Debriefing*). These temporal reference points support participants in reconstructing the progression of their workload while maintaining the continuous nature of the workload profile.

4. Data Processing and Analysis

The workload profile generated by the instrument can be processed either directly in digital form or by digitizing paper-based recordings. For digital implementations, the workload curve is stored as a sequence of two-dimensional coordinates ((x, y)), representing workload as a function of task progression. Event annotations are stored together with their corresponding temporal positions and free-text descriptions. For paper-based assessments, the completed workload profile is digitized by scanning or photographing the chart and extracting the curve coordinates using image-processing software (e.g., MATLAB, Python-based tools, or WebPlotDigitizer[1]). Event annotations are subsequently transcribed into a digital format. All workload values below *very low* must be set to *very low*, and all values above *very high* must be set to *very high*, respectively.

When comparisons with external time-resolved data sources (e.g., physiological measurements, behavioral observations, or system logs) are required, the reconstructed workload profile can be aligned with the corresponding real-time data. Depending on the study design, annotation software (e.g., ELAN[2]) may be used to synchronize event annotations with external recordings.

For more detailed analyses, the workload profile may be segmented either according to predefined experimental phases or based on characteristics of the workload trajectory itself, such as periods of increasing workload, stable workload, or workload decline. Such segmentation enables phase-specific analyses and facilitates comparisons across different task stages.

From the continuous workload profile alone, several quantitative workload metrics can be derived such as average, peak or variability. The event annotations provide complementary information on workload sources, for example the number of annotated events, their temporal distribution, or their occurrence at different workload levels. Qualitative analyses may categorize reported stressors according to domains or demand types, depending on the research objectives.

The quantitative workload profile and qualitative event annotations can also be analyzed jointly. This triangulation enables the identification of relationships between changes in perceived workload and the participant-reported factors contributing to those changes, thereby supporting a more comprehensive interpretation of workload dynamics.

Finally, analyses for individual participants, for teams or groups can be undertaken. In the latter cases, participants' workload profiles may be compared regarding, for example, synchronicity of workload variations, the occurrence of commonly reported events or critical phases.

See Table 1 for example metrics that could be extracted and analysed from a DYWA assessment. Researchers are encouraged to choose from them, and are welcome to add own metrics, based on the respective research question and study design.

[1] https://automeris.io/
[2] https://archive.mpi.nl/tla/elan

Level	Metric	Description	Example Calculation
Individual	Instantaneous workload	Workload at a specific time point	$W(t_i)$
	Average workload	Mean workload across the task	Mean(W)
	Peak workload	Maximum workload	Max(W)
	Minimum workload	Minimum workload	Min(W)
	Cumulative workload	Area under the workload curve	Trapezoidal integration
	Variability of workload	Variance of workload	SD, SE, Entropy, RMSSD, Slope
	Time to peak / Time to recovery	Time from beginning of experiment to maximum workload / Time from maximum workload to next medium workload	$t_{Max(W)} - t_0$ $t_{W(3)} - t_{Max(W)}$ for $t_{W(3)} > t_{Max(W)}$
	Baseline	Individual baseline at start or carry-over effect from previous task	W(0)
	Signed deviation from optimal	Underload vs overload	$W(t_i) - 3$
	Event frequency	Number of annotated workload events	Count
	Event categories	Classification of reported stressors	Thematic coding
	Valence of events	Events that decrease workload vs. events that increase workload	Thematic coding + workload at the events with corresponding distances to previous event $W(E_i) - W(E_{i-1})$
	Triangulation	Linking workload trajectory with annotated events	Mixed-method analysis
Team	Synchronisation	Shared / divergent curves of team workload during task	Cross-correlation of quantitative individual measures, duration of divergent workloads
	Workload distribution	Distribution of workload in team	Comparison of chosen metrics from the "Individual" category between team members, task share proportions
	(un-)Shared events	Events annotated by all team members	Events that share annotations at (dis-) similar points in time and/or valence
Group	(un)Common events	Events used across all / singular participants	Events used in less than 20% or more than 80% of the participants

	Manipulation check operationalisation	Task load differences across conditions; occurrence of induced events	Mean(W) under condition1 ⇔ Mean(W) condition 2; binary decision on occurrence (yes, no)
	Critical phases	Definition of phases with high … medium … low workload time-based or event-related	Percentiles of moving windows across all participants, selection of time phases or event-related phases
	(Critical) stressor category	Definition of stressors for domains like decision making, navigation, monitoring, alarms, etc. that impacted workload	Thematic coding grouping of domains (for workload increase or decrease)

Table 1: Examples of metrics that can be derived from DYWA.

5. Evaluation of Psychometric Properties

The following table give an overview of common psychometric criteria and their degree of fulfilment regarding the DYWA, Table 2. The need for empirical evidence is highlighted where deemed indispensable.

Criterion	Degree of Fulfillment	Explanation
Objectivity	High	Data collection: Low risk of experimenter bias because of minimal interaction between participant and experimenter (depending on the implementation) Analysis and interpretation: defined scale and verbal anchors reduce risk of biased interpretation, digitalization of time-series is objective
Reliability	Tbd.	Need for empirical evidence regarding: Test-retest reliability Split-half reliability
Validity	Given / Tbd.	Content validity is given due to high face validity Need for empirical evidence regarding: Construct validity (convergent/discriminant) Criterion validity (concurrent/predictive)
Acceptance	High	Short, low intrusiveness, easy to understand
Fairness	High	Item is uncorrelated with personal characteristics
Efficiency	High	One item, comparatively fast, easy to administer
Utility	High	Concept of workload has high relevance in research and practice; no other suitable instrument available for dynamic workload assessment fulfilling the requirements
Transparency	High	Instructions are short and easy to understand, an example can be shown
Non-fakeability	Medium	Subjective character of the item includes possibility to fake the absolute level of workload, but the combination with annotated events decreases the likelihood, number of annotated events is subject to participant's intention
Reasonableness	High	Short, low intrusiveness, easy to understand

Table 2: Evaluation of Psychometric Properties of DYWA.

6. Potential Applications

Human-factors studies: The proposed instrument is particularly suited for experimental studies investigating the temporal dynamics of mental workload. By reconstructing workload trajectories rather than providing a single overall rating, the method enables the identification of workload peaks and transitions between task phases.

Training: Repeated use of the instrument after each training run and comparisons over time may provide insight into individual learning curves and further training needs. This could facilitate tailored trainings for both novices and experienced operators.

Evaluation of HMI and operational concepts: The instrument may support comparative evaluations of interface designs, automation concepts, or decision-support systems by revealing differences in workload development throughout task execution.

Combination with other continuous measures: The continuous workload profile can be synchronized with physiological measurements, behavioral observations, eye-tracking data, or system logs to facilitate multimodal analyses of workload.

Explorative analyses: The qualitative event annotations enable researchers to identify workload sources that may not be captured by predefined questionnaire items and can therefore support exploratory analyses in novel operational domains or with novel interfaces or automation procedures.

7. Conclusion

This paper introduces a dynamic workload assessment method that combines a continuous visual workload profile with qualitative event annotations to capture both the temporal development and the perceived sources of mental workload. The method was developed based on design requirements derived from the human factors literature as well as actual needs in the aviation domain to understand the development of mental workload holistically, aiming to address limitations of existing workload assessment approaches related to temporal resolution, multidimensionality, and practical applicability.

By integrating quantitative and qualitative information within a single instrument, the proposed approach enables researchers to examine workload fluctuations over the course of a task while simultaneously documenting the participant-perceived factors contributing to workload changes. Furthermore, the visual workload profile facilitates the derivation of various workload metrics, including instantaneous, average, peak, and cumulative workload, and supports integration with complementary data sources such as physiological measures or behavioral observations.

The proposed method is intended as a flexible research instrument for laboratory and simulator studies as well as other application domains in which understanding the dynamics of mental workload is of interest. Future work should focus on evaluating its psychometric properties, including reliability and validity, as well as investigating its usability across different tasks, user groups, and implementation formats. Such studies will be essential to establish the method's applicability and to determine its value relative to existing workload assessment techniques.

Overall, the presented method provides a structured framework for capturing the dynamic and multidimensional nature of mental workload and offers a foundation for future methodological and empirical research in human factors and ergonomics.

References

Babaei, E., Dingler, T., Tag, B., & Velloso, E. (2025). Should we use the NASA-TLX in HCI? A review of theoretical and methodological issues around Mental Workload Measurement. *International Journal of Human-Computer Studies, 201*, 103515. https://doi.org/10.1016/j.ijhcs.2025.103515

Bartolomei, M., Gervasi, R., Acconito, C., Angioletti, L., Cannizzaro, D., Balconi, M., Mastrogiacomo, L., & Franceschini, F. (2026). Evaluating mental workload: A taxonomic approach to evaluation tools based on ISO 10075. *Applied Ergonomics, 130*, 104659. https://doi.org/10.1016/j.apergo.2025.104659

Boumann, H., Hamann, A., Biella, M., Carstengerdes, N., & Sammito, S. (2023). Suitability of Physiological, Self-report and Behavioral Measures for Assessing Mental Workload in Pilots. In D. Harris & W.-C. Li (Eds), *Engineering Psychology and Cognitive Ergonomics* (pp. 3–20). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-35392-5_1

Brandenburger, N., Dressler, A., & Grippenkoven, J. (2023). *DLR Workload Assessment Tool (DLR- WAT) – Official English Version*.

Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: A systematic review. *Applied Ergonomics, 74*, 221–232. https://doi.org/10.1016/j.apergo.2018.08.028

Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers, 21*(1–2), 133–145. https://doi.org/10.1016/j.intcom.2008.10.011

Gopher, D., & Donchin, E. (1986). Workload: An Examination of the Concept. *Handbook of Perception and Human Performance, 2*, 1–49.

Grippenkoven, J., Rodd, J., & Brandenburger, N. (2018). DLR-WAT: Ein Instrument zur Untersuchung des optimalen Beanspruchungsniveaus in hochautomatisierten Mensch-Maschine-Systemen. *AAET Automatisiertes & Vernetztes Fahren*, 199–213. https://elib.dlr.de/119443/

Guastello, S. J. (2017). Nonlinear dynamical systems for theory and research in ergonomics. *Ergonomics, 60*(2), 167–193. https://doi.org/10.1080/00140139.2016.1162851

Hamann, A., & Carstengerdes, N. (2026). Windows into the mind: From neurophysiological to multimodal measurements in the cockpit. *Current Research in Behavioral Sciences, 11*, 100223. https://doi.org/10.1016/j.crbeha.2026.100223

Hancock, P. A. (1989). A Dynamic Model of Stress and Sustained Attention. *Human Factors, 31*(5), 519–537. https://doi.org/10.1177/001872088903100503

Hart, S. G. (1986, January 1). *NASA Task Load Index (TLX)*. https://ntrs.nasa.gov/citations/20000021487

Longo, L., Wickens, C. D., Hancock, G., & Hancock, P. A. (2022). Human Mental Workload: A Survey and a Novel Inclusive Definition. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.883321

Mansikka, H., Simola, P., Virtanen, K., Harris, D., & Oksama, L. (2016). Fighter pilots' heart rate, heart rate variation and performance during instrument approaches. *Ergonomics, 59*(10), 1344–1352. https://doi.org/10.1080/00140139.2015.1136699

Martins, A. P. G. (2016). A review of important cognitive concepts in aviation. *Aviation, 20*(2), 65–84. https://doi.org/10.3846/16487788.2016.1196559

Tattersall, A. J., & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics, 39*(5), 740–748. https://doi.org/10.1080/00140139608964495

Thelen, E. (2005). Dynamic Systems Theory and the Complexity of Change. *Psychoanalytic Dialogues, 15*(2), 255–283. https://doi.org/10.1080/10481881509348831

Tsang, P. S., & Velazquez, V. M. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics, 39*(3), 358–381. https://doi.org/10.1080/00140139608964470

Wickens, C. D. (1981). *Processing Resources in Attention, Dual Task Performance, and Workload Assessment.* (EPL-81-3/0NR-81-3; Technical Report). University of Illinois. http://archive.org/details/DTIC_ADA1027192

Wickens, C. D. (2008). Multiple Resources and Mental Workload. *Human Factors, 50*(3), 449–455. https://doi.org/10.1518/001872008X288394

Xie, B., & Salvendy, G. (2000). Review and reappraisal of modelling and predicting mental workload in single- and multi-task environments. *Work & Stress, 14*(1), 74–99. https://doi.org/10.1080/026783700417249

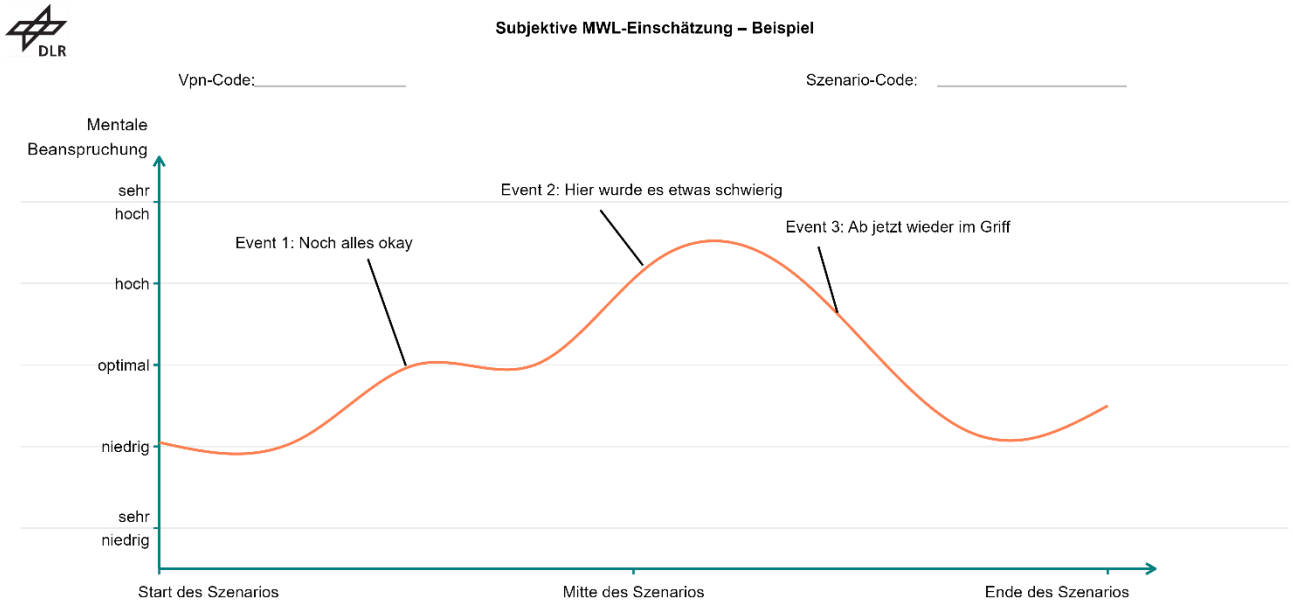
Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics, 58*(1), Article 1. https://doi.org/10.1080/00140139.2014.956151

Zak, Y., Parmet, Y., & Oron-Gilad, T. (2020). Subjective Workload Assessment Technique (SWAT) in Real Time: Affordable Methodology to Continuously Assess Human Operators' Workload. *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2687–2694. https://doi.org/10.1109/SMC42975.2020.9283168

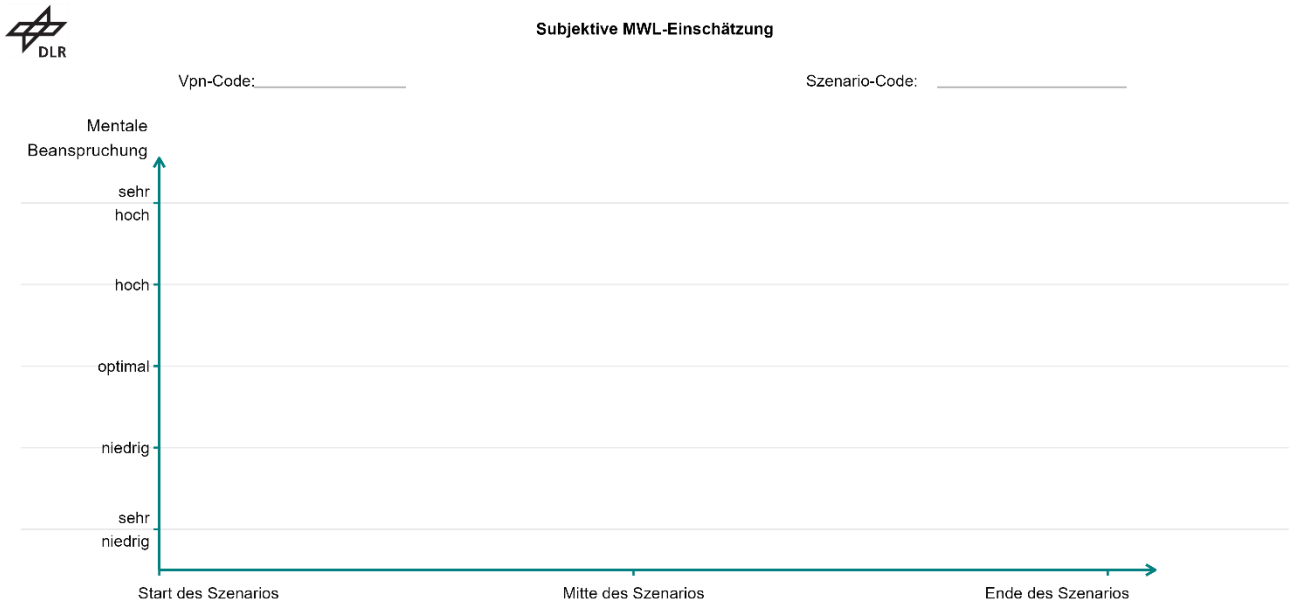
Zhao, Y., Zhu, K., Guo, W., Xu, H., Zhang, J., Zou, J., Li, R., & Wang, L. (2025). Differentiating pilot distress and eustress via multimodal physiology: Towards enhanced human-system integration in intelligent cockpits. *Ergonomics, 69*(10), 2079–2101. https://doi.org/10.1080/00140139.2025.2553131

Appendix

German Example and Template

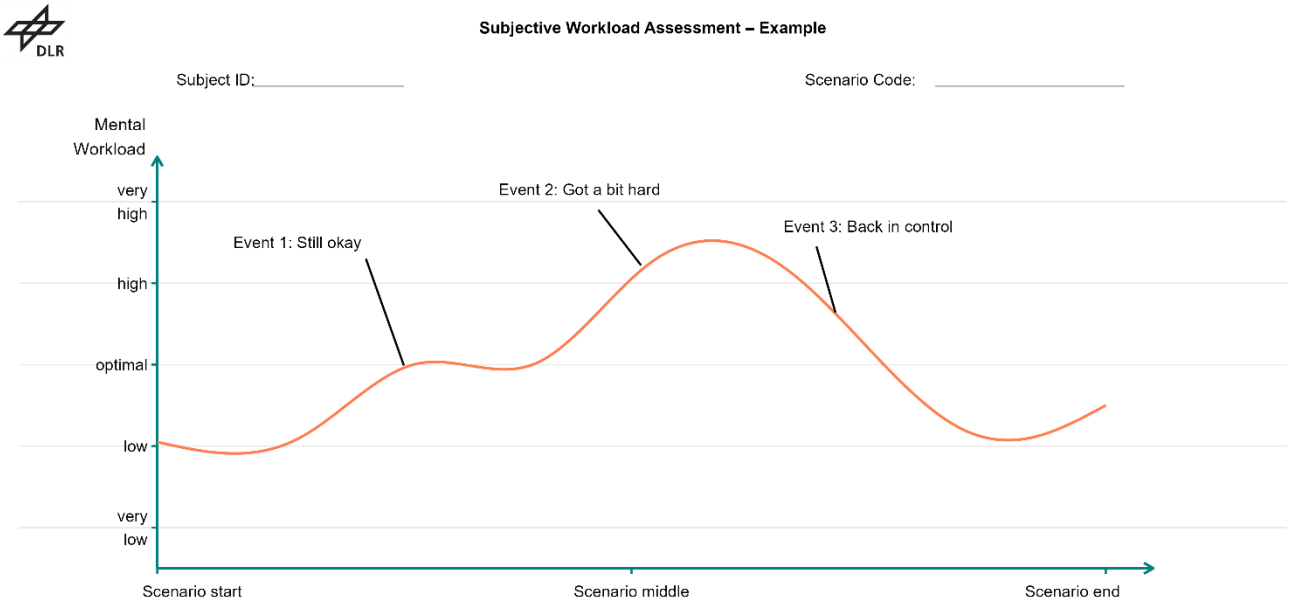


Instruktion: Bitte zeichne eine Linie ein, die Deiner wahrgenommenen mentalen Beanspruchung während des Szenarios entspricht. Relevante Ereignisse kannst Du dabei mit einem Pfeil kennzeichnen und beschriften.

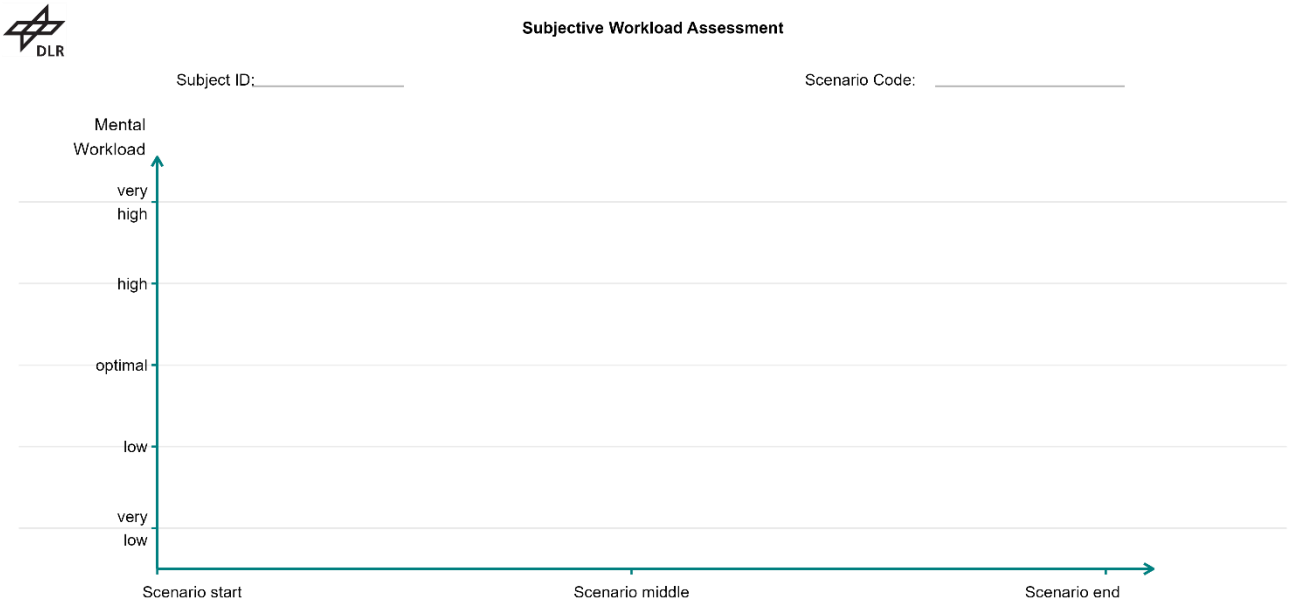


Instruktion: Bitte zeichne eine Linie ein, die Deiner wahrgenommenen mentalen Beanspruchung während des Szenarios entspricht. Relevante Ereignisse kannst Du dabei mit einem Pfeil kennzeichnen und beschriften.

English Example and Template



Instruction: Please draw a line that matches the mental workload you perceived during the scenario.
You can mark relevant events with an arrow and a short description.



Instruction: Please draw a line that matches the mental workload you perceived during the scenario.
You can mark relevant events with an arrow and a short description.