Forschungsbericht 2021-28

The Manometer Task – A Task to Induce Steady Mental Load

Hans-Jürgen Hörmann, Sarah Piechowski, Bernd Johannes

Deutsches Zentrum für Luft- und Raumfahrt
Institute of Aerospace Medicine

Hamburg, Cologne / Germany

DLR Deutsches Zentrum für Luft- und Raumfahrt

Forschungsbericht 2021-28

The Manometer Task – A Task to Induce Steady Mental Load

Hans-Jürgen Hörmann, Sarah Piechowski,
Bernd Johannes

Deutsches Zentrum für Luft- und Raumfahrt
Institute of Aerospace Medicine
Hamburg, Cologne / Germany

51 Seiten
7 Bilder
10 Tabellen
18 Literaturstellen

DLR Deutsches Zentrum
für Luft- und Raumfahrt

Herausgeber:

Deutsches Zentrum
für Luft- und Raumfahrt e. V.
Wissenschaftliche Information
Linder Höhe
D-51147 Köln

ISSN 1434-8454
ISRN DLR-FB-2021-28

DOI: https://doi.org/10.57676/t4j0-5e50

Mental load, stress tolerance, adaptive test, performance degradation, space crew

Hans-Jürgen Hörmann[1)] Sarah Piechowski[2)] and Bernd Johannes[2)]
DLR, Institute of Aerospace Medicine, [1)]Department of Aviation and Space Psychology, [2)]Muscle and Bone Metabolism, Hamburg/Cologne

The Manometer Task – A Task to Induce Steady Mental Load

Research project Psychophysiological Assessment of Strain in the SIMULATOR-Test
DLR-Forschungsbericht 2021-28

With the Manometer Task (MNT) cognitive performance can be examined under conditions of constantly elevated mental load. An adaptive algorithm continuously adjusts the task difficulty in accordance to the previous individual performance. The task concept is described and the psychometric properties are reported for the recommended MNT-configuration with data of N = 2084 candidates. The results showed that the mean reaction times are unrelated to the performance quality of the respective individual. Although, the individual performance quality showed moderate stability across the three levels of task complexity (.32 < r < .34), we did not observe substantial correlations to other tests of cognitive performance. Gender differences were insignificant. Also, age correlations are small (r < .10). Several examples of different practical applications showed how the MNT can be applied to monitor individual performance under the influence of different external stressors.

Mentale Belastung, Stresstoleranz, adaptiver Test, Leistungsabbau, Raumfahrtbesatzung

(Published in English)

Hans-Jürgen Hörmann[1)] Sarah Piechowski[2)] und Bernd Johannes[2)]
DLR, Institut für Luft- und Raumfahrtmedizin, [1)]Abt. Luft- und Raumfahrtpsychologie, [2)]Abt. Muskel- und Knochenstoffwechsel, Hamburg/Köln

Die Manometer Aufgabe – Eine Aufgabe zur Induzierung anhaltender mentaler Belastung

Forschungsprojekt Erprobung psychophysiologischer Beanspruchungsmessung im SIMULATOR-Test
DLR-Forschungsbericht 2021-28

Mit der Manometeraufgabe (MNT) läßt sich die kognitive Leistung unter anhaltend hoher mentaler Belastung untersuchen. Dabei verändert ein adaptiver Algorithmus die Aufgabenschwierigkeit in Relation zur vorangegangenen individuellen Leistung. Das Aufgabenkonzept und die psychometrischen Eigenschaften werden für die empfohlene Konfiguration des MNT anhand von Daten mit N = 2084 Personen beschrieben. Es zeigte sich, dass die mittleren Reaktionszeiten unabhängig von der individuellen Leistungsqualität sind. Obwohl die individuelle Leistungsqualität sich recht stabil über die drei Komplexitätsniveaus verhielt (.32 < r < .34), korrelierte sie nur geringfügig mit anderen kognitive Leistungstests. Geschlechtsunterschiede waren nicht signifikant. Alterskorrelationen waren niedrig (r < .10). Verschiedene Anwendungsbespiele zeigen, wie der MNT eingesetzt wurde, um individuelle Leistung unter dem Einfluss verschiedener externer Stressoren eingesetzt wurde.

Forschungsbericht 2021-28

The Manometer Task - A Task to Induce Steady Mental Load

Hans-Juergen Hoermann, Sarah Piechowski, & Bernd Johannes

German Aerospace Center (DLR)
Institute of Aerospace Medicine
Cologne, Germany

DLR Deutsches Zentrum für Luft- und Raumfahrt

Contents

Tables

Figures

Abstract

With the Manometer Task (MNT), cognitive performance can be examined under conditions of constantly elevated mental load. An adaptive algorithm continuously adjusts the task difficulty in accordance to the previous individual performance. The task concept is described and the psychometric properties are reported for the recommended MNT-configuration with data of $N = 2084$ candidates. The results showed that the mean reaction times are unrelated to the performance quality of the respective individual. Although the individual performance quality showed moderate stability across the three levels of task complexity ($.32 < r < .34$), we did not observe substantial correlations to other tests of cognitive performance. Gender differences were insignificant. Also, age correlations were small ($r < .10$). Several examples of different practical applications show how the MNT can be applied to monitor individual performance under the influence of different external stressors.

1. Introduction

The working environment of key operators in high-risk industries can be characterized by skill-based and rule-based routine activities over longer time periods, which can unexpectedly be interrupted by sudden phases of significantly elevated stress. For example, if an airline pilot performs a manual landing in bad weather on an unfamiliar airport after an eventless long-haul flight, she/he will be highly activated because even small deviations from the desired flight path could lead to catastrophic consequences. For space crews a situation can be similarly stressful when a life-threatening emergency occurs (e.g., a leakage in the outer shell of the spacecraft) and the repair procedure has to be carried out in the correct sequence and under high time-pressure in an adverse environment. In both examples the crew members have to be resilient against possible performance degradations, which could result from high levels of physical and mental load.

In order to examine the performance of human operators under such working conditions, a low-cost standardized scenario is required, which allows to induce high levels of mental load to the candidate while she/he is performing a cognitive task. Such a scenario is provided by the Manometer Task (MNT), which was developed by the third author Bernd Johannes and his colleagues at the Max-Delbrück-Center for Molecular Medicine in Berlin-Buch in the early 1990ies (e.g., Johannes et al., 1995).

The MNT is a cognitive task where the task difficulty is continuously adjusted by an adaptive algorithm in accordance to the current performance level of the respective candidate (see Chapter 2). By means of this test principle, constant mental pressure is induced in relation to the individual's performance across various levels of task complexity. While the individual candidate is kept at this high level of activation, psychophysiological monitoring can take place in order to investigate how the regulatory mechanism of this candidate is coping with the situation. In addition, this steady state of elevated activation can be used to perform the individual calibration of physiological parameters (see Chapter 5).

The purpose of this document is to describe the administrative procedure, the cognitive task, the scoring and the psychometric properties of the MNT. The MNT had been evaluated in the context of the pilot selection process for a large European airline at the DLR Department of Aviation and Space Psychology in Hamburg (see Chapter 3). The findings are reported in this document and calibrated distribution scores are provided in Chapter 4 and in the Appendix. In the final Chapter 5, some practical examples of several MNT applications are presented.

2. Task description

2.1 Concept

The MNT is a computerized task where the individual candidate has to monitor the health status of a fictitious system. The system status is indicated on a certain number of pressure gauges and a target arrow at the top of the computer screen. If the displayed system status is acceptable within the tolerances, the candidate has to respond with the voice command “OK” or by pressing the key for “OK”. If the system is showing a critical deviation from its proper status, the candidate has to respond with the voice command “Error” or by pressing the key for “Error”. The decision about the health status of the system is based on the observation whether all pressure gauges are pointing in the same direction as the target arrow on top of the screen. The direction of the target arrow can be any of the four cardinal semi-circulars: “up”, “down”, “left”, or “right”. Only if all indicators of all gauges are lined up towards the same cardinal semi-circular as the target arrow, the system is working properly. An example is shown in Figure 1, where all three pressure gauges are pointing upwards as the target arrow. If one or more of the gauge pointers are showing a deviation of over ±90° from the direction of the target arrow, the system status is incorrect. In Figure 2 for example, the target arrow is pointing downwards, but the bottom left instrument is pointing upwards. All items of the MNT have a clearly recognizable solution without ambiguities. Tasks with a correct system status and tasks with an incorrect system status are distributed randomly. When the candidate has entered his/her decision, the present task disappears immediately, an acoustic feedback about the correctness is provided, and the next task appears. Subsequent input corrections are not possible.

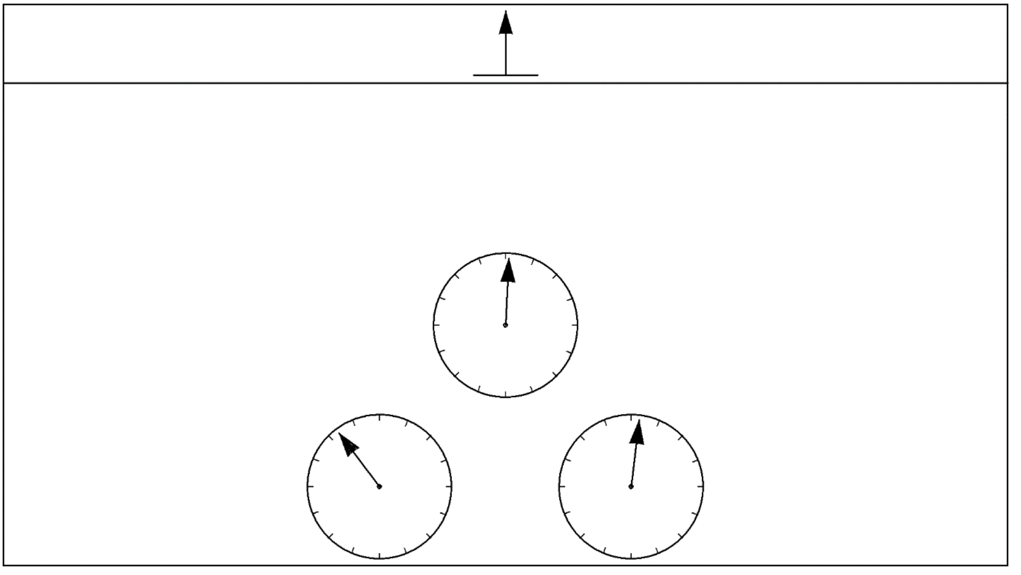


Figure 1: MNT with correct system status. All three gauges are lined up with target arrow on top of the screen

The main idea of the MNT is to induce a consistent level of elevated mental load while the individual candidate is monitoring and responding to more or less complex visual information. The mental load is instigated by varying degrees of complexity of the visual stimuli and by constant time pressure. By means of an adaptive algorithm for each presented task, the display time of the visual image varies in correspondence to the candidate's current performance. If a response was correct, the presentation time for the subsequent task is shortened by 20%. In case of an incorrect response the presentation time for the subsequent task is extended by 25%. In other words, the task difficulty follows dynamically the individual performance curve, and by doing so the mental load is kept at a relatively steady high level throughout the entire test run. In addition to the presentation time, the number of simultaneously displayed gauges varies between reasonable numbers. In the standard configuration we recommend an increasing number from five via seven and up to nine gauges at the same time for each item of that respective level. The first item of each level is presented for the maximum time period of 5 seconds. As soon as the candidate responds either orally or by pressing a response key, the item disappears and this first reaction time is adopted as the initial presentation time. Some key settings of the MNT can be defined via the configuration page (see Appendix 7.1).

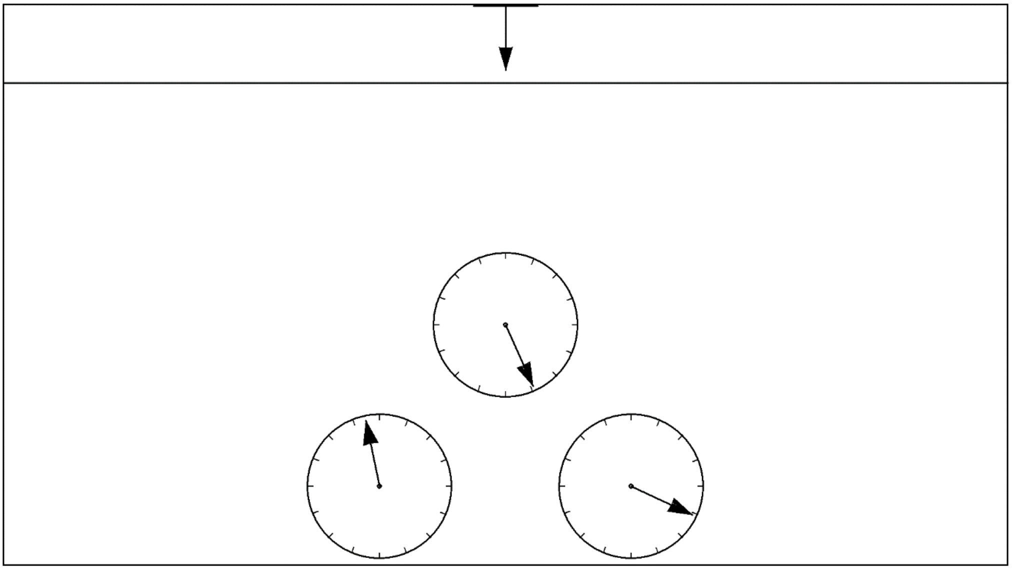


Figure 2: MNT with incorrect system status. The left gauge deviates from the target arrow.

The input of the candidate’s decision can either be done by pressing one of the two response keys on the keyboard or mouse or by voice commands. If voice inputs are chosen for the responses, the computer has to be trained in advance for individual voice recognition of the two possible commands "OK", if the system is in a correct condition, and "ERROR", if the system is in an incorrect condition. The MNT software includes this learning function. In several of the example applications described briefly in Chapter 5, the MNT was utilized to create a persisting high workload situation for the candidate, which served to calibrate the individual voice pitch and other psychophysiological parameters for a subsequent experiment (e.g., Johannes et al., 2019; Johannes, Salnitski, Gunga & Kirsch, 2000; Johannes et al., 1995).

2.2 Administration and material

The MNT is a computerized task which starts with an oral instruction of the candidate, preferably via a noise-cancelling headset (see Appendix 7.2). After the instruction, the computer can optionally be trained to recognize the individual voice commands "OK" and "ERROR". Altogether this procedure takes about 2 minutes including the instruction. In order to prevent other external variables to interfere with the experimental condition, it is recommended to administer the MNT in a quiet and dazzle-free environment without any extra distractions. However, this does not preclude administration in smaller groups. If an experimental setting is chosen, the environmental conditions can be varied according to the investigated research question. Depending on the selected number of items and levels of complexity, the entire MNT can be completed in about 5 – 8 minutes.

2.2.1 System requirements

In its standard configuration the MNT requires for each candidate a regular Windows computer with soundcard, keyboard, headset, and a monitor of at least 11". The MNT computer program installs some software including a MySQL database for storing the individual data to the mass storage device of the computer. The headset should have a microphone if voice inputs are required. Otherwise, the candidates can respond to each item by pressing two different keys on the keyboard or by using the two mouse buttons.

2.2.2 Standard configuration

For the administration of the MNT to adults in good physical and mental condition, it is recommended to adjust the test parameters via the program configuration page as follows (see Appendix 7.1):

- Levels of complexity:
  - First 5, then 7, then 9 pressure gauges presented simultaneously
- Number of items per level:
  - 25 items for 5 gauges, 20 items for 7 gauges and also 20 for 9 gauges
- Initial presentation time per item:
  - will automatically be set to a maximum of 5000 ms, depending on the first response
- Timeout if no response:
  - 5000 ms

With the configuration page the MNT can be adjusted for special research conditions. For example, it is possible to define a static instead of the dynamically changing presentation time, if the MNT tasks have to be presented simultaneously to other stimuli. Certainly, the MNT can also be administered to other populations, age-groups or in different experimental settings. However, the distribution scores provided in this document represent healthy young subjects and standard MNT-settings. Oral and written instructions are available in English, German, Russian, Bulgarian, French, Spanish, and Italian language.

2.3 Scoring of individual performance

For each MNT item on each level, the computer program records the candidate’s response (“OK” or “ERROR”), the response time, and the presentation time of each displayed task into the MYSQL database. A response from the candidate can be either right or wrong; aside from no response before timeout, which would also be counted as wrong. In the terminology of the signal detection theory (e.g., Swets, 1964), the MNT can be regarded as a two-choice cognitive task for which the responses can be sorted across four categories as shown in

Table 1. Correct responses are made up of "hits" and "correct rejections". Incorrect responses are made up of "misses" and "false alarms". The signal which should be detected in the sense of the signal detection theory, is the system state "OK" in accordance to the correct direction of the gauge pointers and the target arrow.

Table 1: Possible response categories for MNT items

	Response "OK"	**Response "Error"**
Presented system status "OK"	Hit	False alarm
Presented system status "Error"	Miss	Correct rejection

In order to reflect the individual performance, a number of variables are usually calculated from the recorded responses in the database. As a standard setting the following variables can be considered for each of the administered MNT levels:

1. Mean of response times (MRT)
2. Standard deviation of response times (SRT)
3. Mean of presentation times (MPT)
4. Standard deviation of presentation times (SPT)
5. Percentage of hits for the "OK" status items (HIT%)
6. Percentage of correct rejections for the "Error" status items (COR%)
7. Percentage of false alarms for the "OK" status items (FAA%)
8. Percentage of undetected "Error" status items (MIS%)
9. Percentage of correct responses for all items (sum of #5 and #6) (QUALI)

Besides these nine level-related variables a total score can be calculated to depict the overall performance quality:

10. Percentage of correct responses across all levels (overall rate of correct repsones) (O-QUALI)

Not all of these variables are independent from each other. For example, the percentages of the response categories "Hits" and "False Alarms" add up to 100%, as do the response categories "Misses" and "Correct Rejections". However, we kept these redundant variables for specific analyses of error types. We propose to use the relative frequencies in comparison to the absolute frequencies of responses, because in the original version of the MNT software (which were applied in this documentation) the presented system states ("OK" and "Error") were not equally distributed. In the current MNT version the frequencies are balanced.

Some of the performance variables are negatively scored, especially the time measures (MRT, MPT) and the errors (FAA%, MIS%). The standard deviations of the time measures (SRT and SPT) can be seen as parameters of performance volatility, which is also regarded as negative. That means, for these variables higher scores are related to lower performance. Positive performance measures are HIT%, COR%, QUALI, and O-QUALI, with higher scores reflecting better performance.

The MNT program has also the option to calculate presentation and response times separately for each of the four response categories. We have used these variables in Section 3.4.1.

2.4 Calibration study

A rather homogeneous sample was recruited to calibrate the single performance parameters and to serve as a basis for the empirical evaluation of the MNT's psychometric properties. For this purpose, a sample of young adults (N = 2084) was recruited alongside the psychological selection process for the pilot training program of a major European Airline. All participants were high school graduates and had been applying for pilot training in the airline. The MNT was administered in smaller groups with up to ten participants during the final selection stages. As response modality the voice input option was chosen. Three levels of complexity were presented subsequently with 5, 7, and 9 simultaneously displayed pressure gauges. The number of tasks for the three levels was 25 (level 5), 20 (level 7), and 20 (level 9). Demographic information for the entire sample can be found in Table 2. The distribution statistics for all MNT performance scores are reported in Appendix 7.3.

Data from all aptitude tests of the pilot selection process were available for this sample. These aptitude tests, which are described in further detail in Section 0, consisted of cognitive ability tests, knowledge tests, a flight simulator test, and personality assessments by a questionnaire (TSS) and assessment center exercises. Due to the multi-hurdle selection procedure the sample size can vary.

Table 2: Demographic information about the subjects in the MNT calibration sample

Sample	Sample size	*N* = 2084
High school graduates	Age	Mean = 21.1, SD = 2.3 range 17 to 29 years
	Gender	1901 males, 183 females
	Nationality	All German-speaking Europeans, 97% Germans

The combined data sets of the calibration study with 2084 subjects were used in this report to evaluate the psychometric properties of the MNT. After an inspection of all frequency distributions, the elimination of outliers was discarded, because just less than 1% of the sample displayed slight anomalies especially in the response time measures. Practical implications seem to be negligible. Nevertheless, the sample size varies slightly for the cognitive tests between $N = 1995$ and $N = 2084$ due to incomplete data. $N = 1295$ candidates also did a work sample test called the flight simulator test.

3. Evaluation

3.1 Objectivity

The instruction of the MNT was presented fully standardized via a recorded voice file from the computer (see Appendix 7.2). Influences of the test instructor were reduced to answering remaining questions. Task performance was measured automatically via response times, presentation times, number of correct responses and number of incorrect responses. Therefore, subjective influences on the results of the MNT can assumed to be minimal, if at all.

3.2 Reliability

Because of restrictions due to the speeded character of the MNT items, reliability could not be assessed item-wise via the common Cronbach's Alpha statistics. Therefore, we adopted the quality scores of the three levels as quasi-items and conducted a reliability analysis as if the MNT had only these three items. Cronbach's Alpha resulted in α = .60, which we regard as a conservative reliability estimate for the MNT performance quality variables on this aggregated level.

In addition, a factor analysis was executed with the independent MNT performance variables for the three levels. The scree plot indicated a four-factor solution as suitable (see Appendix 7.4). After the extraction of four factors with the Principal Axis Factor analyses, the communalities h^2 were taken as reliability estimates of the MNT variables. Under the assumption of the classical test theory, that influences of errors are uncorrelated (i.e., unsystematic), the common variance of the single variables, which are explained by the extracted factors, can be regarded as an estimate for the reliability. The calculated communalities h^2 are shown in Table 3. The rotated factor solution itself is described in Section 3.4.2.

Both the response times and the presentation times appear to have a higher reliability than the quality variables (percentages of correct responses). In order to avoid redundancy, the numbers of incorrect responses were omitted from the analyses. Because of the linear dependencies, reliabilities of the error variables would be the same as for the correct responses. Altogether the findings for the reliability of the MNT are regarded as being acceptable. The range is between .60 (for the item-analysis of the three quality vaiables) and .93 (for the mean response time at level7).

Table 3: Reliability estimates for the MNT performance variables.

MNT-variable	h^2
MRT.5	.64
MRT.7	.93
MRT.9	.72
MPT.5	.89
MPT.7	.87
MPT.9	.88
QUALI.5	.76
QUALI.7	.69
QUALI.9	.69

Note: MNT-variables are abbreviated as listed in Section 2.3. Numbers indicate the MNT level.

3.3 Stability

Since no data for a complete test-retest situation of the entire MNT were available in this calibration study, the Pearson correlations between corresponding MNT variables on the three different levels were calculated as time-related stability coefficients. These coefficients show the interindividual stability of performance across the three levels of complexity (Table 4).

The mean response times show the highest stability coefficients for all levels. Even the standard deviations for the response times have a significant stability across the levels. The stabilities of the presentation times are somewhat lower than those of the response times. It is remarkable that the stability coefficients of the standard deviations of the presentation times (SPT) are, in cases where the initial level 5 is involved, close to zero or even slightly negative. This is probably related to the dynamic adjustment of the presentation times (too short presentation times are extended and too long presentation times are shortened). Also, the relative frequencies of the four response categories have quite low stabilities around r_{tt} ~ .20. Slightly higher again is the stability of the performance quality across the levels.

Another observation is that the interindividual differences between level 7 and 9 seem to be more closely related than between the first level 5 and these two subsequent task levels. A reason can be that candidates had to acquire a stable working style during the initial level 5. The few examples during the instruction before the test started were not enough to stabilize their performance. Therefore, the first level has 25 items instead of 20. For these reasons the participants' performance during level 5 probably reflects slightly different skills and response strategies. Since the available data set only included the average scores per level, we could not exclude the initial five items from the analysis.

According to these findings, the MNT performance variables (especially the response categories and standard deviations) seem to be closer related to current individual states than to personal traits of the participants. The only exception are the average response times and the average presentation times with higher cross-level stabilities.

Table 4: Stability correlation coefficients across the three levels of MNT complexity

MNT-variable	Level 5 -- level 7	Level 5 -- level 9	Level 7 -- level 9
MRT	.75**	.64**	.78**
SRT	.34**	.31**	.59**
MPT	.49**	.38**	.57**
SPT	-.05*	-.04*	.29**
HIT%	.30**	.19**	.19**
COR%	.19**	.19**	.21**
FAA%	.21**	.19**	.19**
MIS%	.19**	.19**	.21**
QUALI	.34**	.33**	.32**

Note: $N = 2084$, * $p < .05$, ** $p < .01$

3.4 Validity

In terms of validity information, the calibration study provides some plausibility indications based on the pattern of intercorrelations between the different MNT performance measures (see Section 3.4.2). In addition, a factor solution is described to better understand the sources of variance embedded in the MNT measures. Finally, the pattern of relations with other cognitive tests from the pilot selection battery is reported in this chapter.

3.4.1 Analysis of response categories

First, we analysed how often the different response categories occurred. Since the distribution of “OK” and “ERROR” system states varied across participants and levels, the absolute frequencies are shown in Figure 3 in comparison to the relative frequencies in Figure 4.

Note, that level 5 had 25 items whereas levels 7 and 9 had only 20 items. The figures confirm the average performance quality of the subjects which is around 66% to 70%. Hits are slightly more frequent than correct rejections. Altogether the average performance quality is rather stable across the levels. It drops just slightly from 70% (level 5) via 68% (level 7) to 66% (level 9). This is in line with the adaptive test principle where weak performance is followed by presenting the next task with an extended presentation time. The error bars represent 95% confidence intervals. We abstained from tests of significance here, because of the large sample size even small differences would be statistically significant.

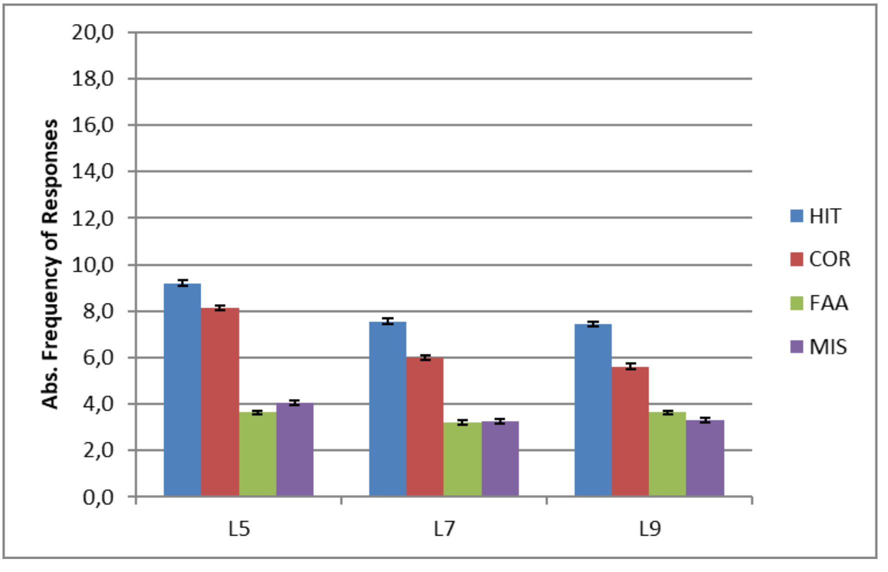


Figure 3: Absolute frequencies of MNT responses categories

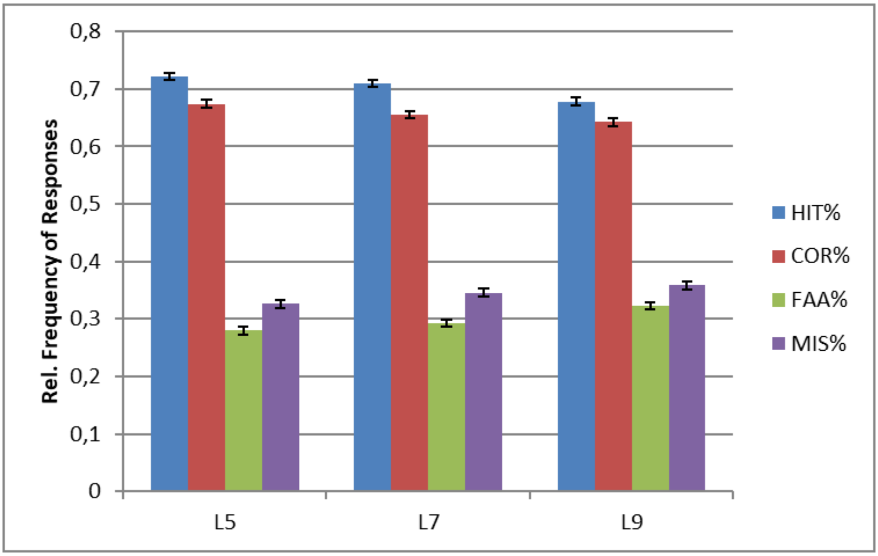


Figure 4: Relative frequencies of MNT responses categories

The average response times for the response categories are shown in Figure 5. On level 5, correct responses were about 300ms faster than wrong ones. However, this difference disappears for the other more complex levels. Only the category of overlooked "ERROR" states (misses) is associated with the slowest response times across all three levels. It is remarkable that correct rejections took a bit longer than the responses with hits. Actually, the recognition process of an error state can be abbreviated by identifying only a single gauge which is not aligned with the target arrow (regardless of all other gauges). In order to confirm a correct system status, all gauges have to be inspected for possible deviations, which should take more time. However, this effect is not clearly reflected by the data. Only with nine gauges the responses for hits and correct rejections occurred with nearly the same speed. Overall, the fastest responses occurred on level 7, which could be explained with an established response style after some rounds of practicing. Again, the error bars represent the 95%

confidence intervals.

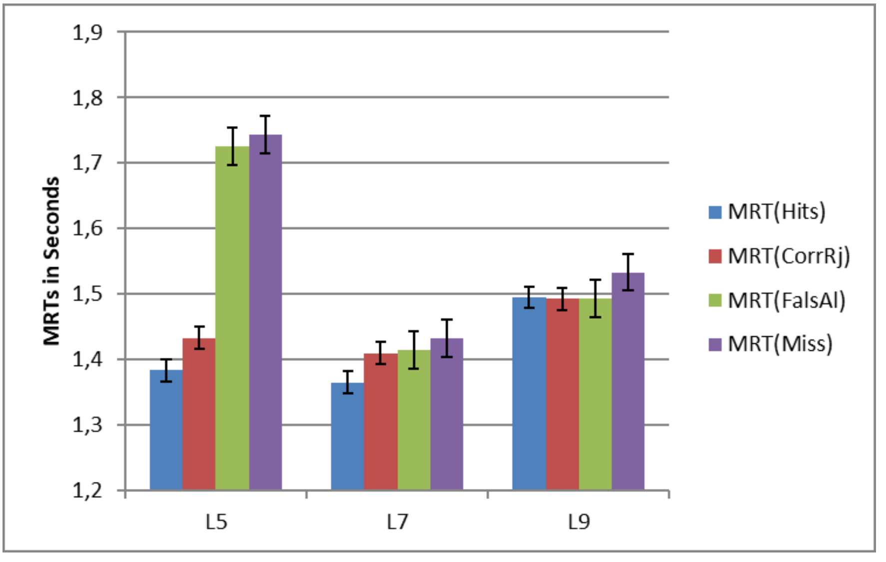


Figure 5: Average response times for the MNT response categories

Figure 6 displays the average presentation times, which also dropped substantially from level 5 to level 7 and then increased slightly for level 9. This is in line with the response times as discussed above. It should be reminded that the length of the presentation times always depends on the correctness of the previous responses and does therefore not clearly represent the respective response category.

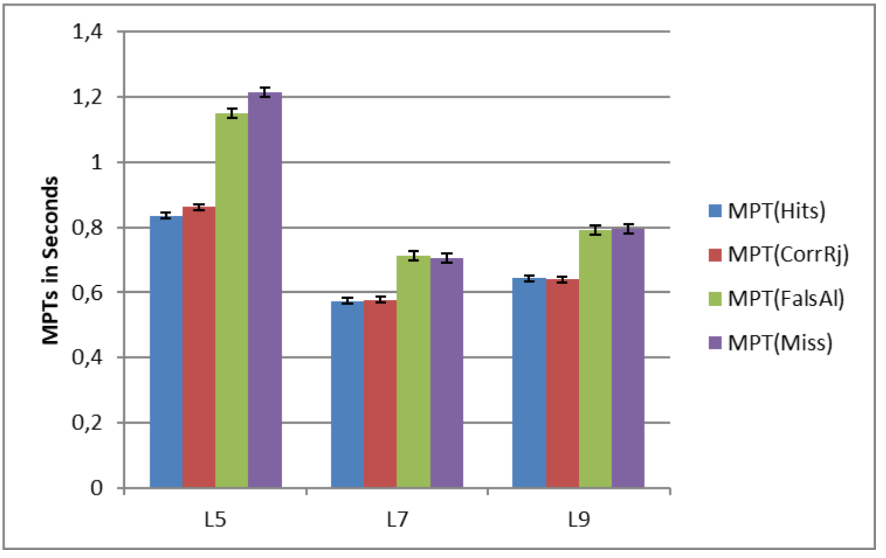


Figure 6: Average presentation times for the MNT response categories

3.4.2 Factor analysis and intercorrelations of MNT variables

With the following intercorrelation analyses we analysed the strength of relations between the single-level related performance scores with the total performance quality score (O-QUALI) for the MNT. As is shown in Table 5, the mean response times (MRT) are unrelated to the overall performance quality (O-QUALI), while all other measures, including the standard deviations of the time variables, are related as expected. All time variables and errors show negative correlations with the overall score, percentages of hits and correct rejections show positive correlations. It is confirmed that fluctuating response and presentation times correspond to lower performance quality. This analysis can be regarded as a plausibility check to verify the test concept.

Table 5: Correlations of single MNT variables with overall performance quality

	MRT	SRT	MPT	SPT	HIT%	COR%	FAA%	MIS%	QUALI
					-- Level 5 --				
O-QUALI	.02	-.20**	-.68**	-.12**	.55**	.55**	-.55**	-.55**	.79**
					-- Level 7 --				
O-QUALI	-.01	-.25**	-.72**	-.44**	.45**	.47**	-.45**	-.47**	.72**
					-- Level 9 --				
O-QUALI	-.06*	-.27**	-.73**	-.48**	.42**	.43**	-.42**	-.43**	.71**

Note: *N* between 1995 and 2084, * $p < .05$, ** $p < .01$

A factor analysis was conducted in order to explore the different sources of variance embedded in the MNT scores. To avoid any duplication of variables, only the percentages of correct responses, the response times and the presentation times for each level were included in the analyses. According to the scree plot (see Appendix 7.4), four factors were extracted with the Principal Axis Factor Analysis. The four factors account for 88% of the MNT variance. As can be seen in Table 6, each MNT level has its own factor, which is loaded by the presentation time together with the performance quality of that level. The first factor however, is composed by the response times across the three levels. This factor solution confirms that the response speed in the MNT is completely unrelated to the demonstrated performance quality of the subjects.

Table 6: VARIMAX rotated four-factor solution for the MNT variables

	Factor 1	Factor 2	Factor 3	Factor 4
MRT.5	.79			
MPT.5		.91		
MRT.7	.95			
MPT.7		.32	.85	
MRT.9	.83			
MPT.9			.38	-.83
QUALI.5		-.82		
QUALI.7			-.78	
QUALI.9				.80

Note: N = 2000; all factor loadings > |.30| are included

3.4.3 Correlations with external reference tests

Within the pilot selection process for the airline, several aptitude tests were administered, which can serve as convergent or divergent criteria for the performance scores of the MNT. The pilot aptitude test battery is described in several publications (Hoermann & Damos, 2019; Hoermann, Noser & Stelling, 2018; Hoermann & Maschke, 1996). We expect higher correlations for the convergent criteria than for the divergent criteria. Aptitude tests, which are highly speeded and presented with figural material, should be closer related to the MNT scores than a knowledge test of English language, memory, or mental arithmetic, because the latter cognitive functions are not essential for the MNT performance. The MNT is expected to be related to personality traits such as emotional stability, because this construct corresponds to general resilience against stress. It is also expected that motivational factors determine MNT performance. Therefore, we included a personality scale of achievement motivation in this analysis. On the other side, the personality trait extraversion is expected to be unrelated to the MNT.

All aptitude tests are positively scored. That means, the higher the score, the better the performance. For the MNT, only the performance quality scores (QUALI-5, QUALI-7, QUALI-9, O-QUALI) are positively scored. The time-related variables are negatively scored. For the latter variables, higher scores reflect worse performance. Consequently, the direction of the scoring has to be considered when interpreting the correlations with convergent or divergent criteria. Altogether we included eleven aptitude tests (eight as convergent and three as divergent criteria) and three personality scales (two as convergent and one as divergent criteria).

List of convergent criteria:

- Perceptual speed
    - SKT – Mental concentration test
    - OWT – Optical perception test
- Instrument comprehension and multitasking
    - MIC – Monitoring and instrument control
    - SIM – Instrument flight simulation
- Speed of work

- DCT – Dyadic coordination working speed (solo version)
- Spatial abilities
  - ROT – Rotations (acoustic)
  - PPT – Dice rotation (visual)
- Mechanical comprehension
  - TVT – Mechanical comprehension
- Emotional stability
  - TSS-EMO – Emotionality (personality questionnaire)
  - TSS-ACH – Achievement motivation (personality questionnaire)

List of divergent criteria:

- English comprehension
  - ENS – English knowledge
- Short-term memory
  - RMS – Running memory span
- Mental arithmetic
  - KRN – Mental arithmetic
- Extraversion
  - TSS-EXT – Extraversion (personality questionnaire)

Table 7 and Table 8 the correlations between the MNT variables and the reference tests are reported. All correlations are surprisingly low. One reason could be that the variation of the aptitude tests is somewhat restricted by the preceding selection steps. However, we did not correct the correlations for neither range restriction nor for possible errors of measurement. Since the included aptitude tests are trait-related measures, the low correlations imply that there is very little evidence of trait variance in the MNT variables. Only a few of the convergent validity coefficients seem to be consistent. The highest significant correlation is between the MNT performance quality variables and a visual spatial orientation test, the PPT. Also, the Dyadic Cooperation Test correlated significantly with several MNT variables. This test requires a good situational overview, quick setting of priorities and fast calculations. Additionally, some of the time-related MNT-variables are also negatively correlated with some of the reference tests.

With respect to the personality scales of emotional instability and achievement motivation, we only found a small negative relationship to achievement motivation. It seems that self-induced pressure might have slightly interfered with the MNT performance.

In summary, it seems that the MNT scores are better suited to measure intraindividual changes in terms of states than interindividual differences in terms of traits. At least within the scope of the pilot selection tests, evidence for common trait variance within the MNT was low. Reportedly, variations of the test motivation could have interfered. The candidates were aware that the MNT performance had no further consequences for being selected for pilot training. It is recommended to carry out some further validation studies based on experimental settings with various degrees of situational stressors. Other measures of resistance to stress should be included as well as repeated measurements.

Table 7: Correlations of MNT with convergent and divergent criteria (Levels 5 and 7)

	Level 5					Level 7				
	MRT	**SRT**	**MPT**	**SPT**	**QUALI**	**MRT**	**SRT**	**MPT**	**SPT**	**QUALI**
Convergent criteria:										
SKT	.06**	.04	-.02	.02	-.04	.01	.01	-.06**	-.03	.01
OWT	.02	.02	.01	-.05*	.01	-.01	.01	-.02	-.03	.03
MIC	-.01	-.03	-.04	-.04	.03	.01	.02	-.03	-.03	.03
SIM	.00	.03	.02	-.01	.02	-.02	-.00	.02	-.02	.06*
DCT	-.05*	-.04*	-.05*	-.09**	.05*	-.02	-.00	-.01	-.03	.06**
ROT	-.01	.01	.03	-.02	.03	-.03	-.01	-.00	-.02	.06**
PPT	-.06**	-.02	-.00	-.05*	.08**	-.07**	-.06**	-.01	-.02	.08**
TVT	-.07**	-.04	-.02	-.05*	.07**	-.04	-.02	.00	-.02	.06**
TSS-EMO	.05*	.03	.03	.01	-.03	.01	.01	.03	.03	.04
TSS-ACH	.04*	.01	.00	.04	-.07**	.02	.00	-.04	.02	-.03
Divergent criteria:										
ENS	-.08**	-.04*	-.08**	-.04	.06*	-.04	-.02	-.06*	-.04*	.05*
KRN	-.05*	-.02	.01	-.08**	.03	-.01	.02	.02	-.03	.02
RMS	.01	.01	.03	-.05*	-.01	-.01	-.02	-.00	-.03	.01
TSS-EXT	.03	.02	.01	-.01	-.03	.03	.02	.00	.03	-.02

Note: *N* between 1995 and 2084, for SIM *N* = 1295; * *p* < .05, ** *p* < .01

Table 8: Correlations of MNT with convergent and divergent criteria (Levels 9 and overall)

	Level 9					Overall
	MRT	SRT	MPT	SPT	QUALI	QUALI
Convergent criteria:						
SKT	.01	.03	-.08**	-.05*	.01	-.01
OWT	.02	.03	-.01	-.03	.01	.02
MIC	.01	.01	-.03	-.04	.02	.04
SIM	-.02	-.01	.01	-.00	.04	.06**
DCT	-.05*	-.04	-.04	-.04	.06**	.06**
ROT	-.02	-.00	.00	.00	.04	.06**
PPT	-.04	-.02	.01	-.02	.07**	.10**
TVT	-.02	-.04	.02	.02	.01	.05*
TSS-EMO	-.01	.03	.04	.04	.04	-.00
TSS-ACH	.05*	.06*	.00	.02	-.03	-.06**
Divergent criteria:						
ENS	-.03	-.04	-.03	-.02	.01	.05*
KRN	-.00	.01	.01	-.02	.02	.03
RMS	.03	.04	.00	-.03	.01	-.00
TSS-EXT	.03	.01	-.01	.01	-.04	-.04

Note: *N* between 1995 and 2084, for SIM *N* = 1295; * $p < .05$, ** $p < .01$

4. Standard Scores

This chapter reports the results of the analyses to generate norm tables for the core variables of the MNT. It should be emphasized again that the distribution scores represent the standard MNT-settings and a population of healthy young adults.

Initially, it will be explored whether any significant effects of gender or age on the MNT-variables can be revealed. These results are important for the choice of either one general norm or differential norms for subgroups.

4.1 Effects of gender

With *t*-tests the effects of gender on the MNT performance variables were examined. Since the percentage of female participants is only around 10%, the results do not have a representative nature yet. Several *t*-tests were conducted for each level and for the overall performance. However, only one significant difference could be detected. On level 9 male participants detected a higher percentage of system errors than females (65% for male and 61% for female participants, Cohen's d = .01). Correct rejections on the other levels do not differ significantly. Also, the average response and presentation times were equal. Therefore, we included only the performance quality variables in Table 9. Cohen's d was calculated for the effect sizes.

Table 9: T-tests for gender differences in the MNT performance variables

MNT-variable	Mean	StDev	***t*-value**	**Significance**	**Effect size**
N_1 = 183 female – N_2 = 1901 male			*df* = 2082	*p*	*d*
QUALI.5					
female	0.69	0.10	-0.71	.48	0.00
male	0.69	0.10			
QUALI.7					
female	0.67	0.11	-0.83	.40	0.00
male	0.68	0.10			
QUALI.9					
female	0.65	0.10	-0.20	.85	0.00
male	0.65	0.10			
O-QUALI					
female	0.67	0.08	-0.84	.40	0.01
male	0.68	0.08			

Altogether it can be stated that the gender differences are either very small or insignificant. Male and female participants performed equally well in the MNT and therefore norm tables do not need to be distinguished for gender.

4.2 Effects of age

The age distributions are shown in **Fehler! Verweisquelle konnte nicht gefunden werden.** 8. The total age range is between 17 and 29 years. However, the majority of participants (70.6 %) were between 19 and 22 years old at the time of the data collection.

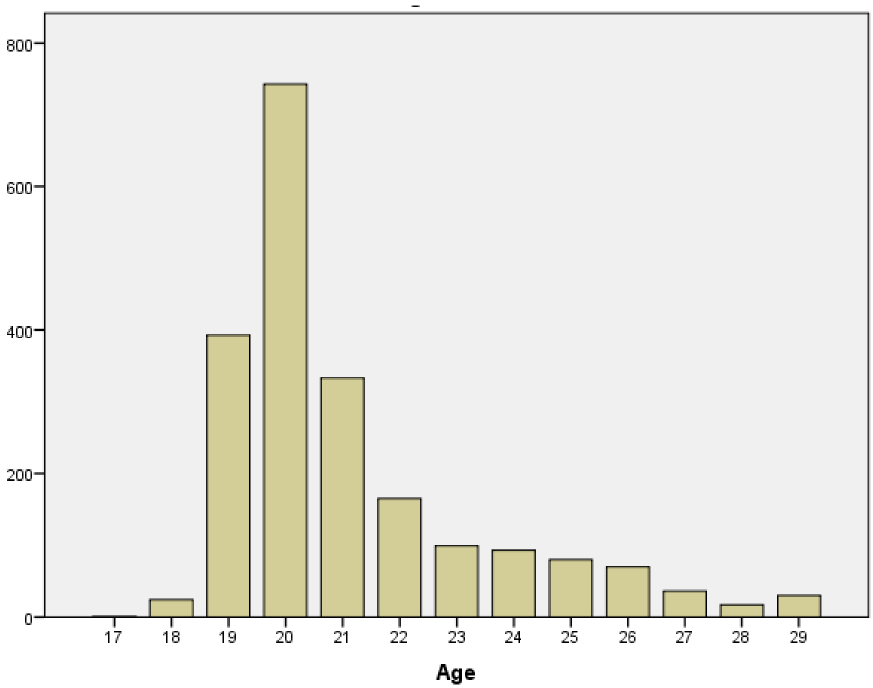


Figure 7: Age distribution in the MNT sample

To determine any significant effects of age on the MNT variables, correlation analyses were conducted with all MNT performance scores. Altogether only two of the calculated correlation coefficients were significant at $\alpha < .05$. Older subjects had a small tendency for higher standard deviations of the reaction times on level 7 and on level 9. On level 7 the correlation of age with SRT.7 is r = .05*. Also, on level 9 the correlation of age with SRT.9 is r = .05*. Based on these findings it is not indicated to calculate age-specific norms for the MNT, at least not for age groups between 17 and 29 years.

4.3 Norm tables

In this section of the document, norm tables for the MNT variables of performance quality are provided. By means of this table, the raw MNT-scores can be converted into standard scores on a nine-point scale, which reflect the relative position of a value in relation to the distribution of the *N* = 2096 subjects of the calibration study. All quality variables are positively scored. Norm tables for the other level-specific variables can be found in the Appendix 7.5. The Stanine distribution was utilized to generate the norm scores. By this method a normalisation of the standard scores is also achieved. When using the norm tables with error variables or time variables, care has to be taken not to mix up the polarity. Errors and usually also time variables are negatively poled in comparison to the performance quality variables.

The advantages of this normalisation procedure are to obtain performance scores with better distribution characteristics and to provide a standard measurement framework, which can be used to benchmark the MNT performance in other groups or situations. This could for example be useful to examine effects of psychoactive substances, external stressors or fatigue on the cognitive performance in the MNT.

Table 10: Stanine table for the MNT variables for performance quality in percent

Stanine	QUALI.5	QUALI.7	QUALI.9	O-QUALI
1 (4%)	0 – 49.6	0 – 50.0	0 – 45.0	0 – 52.0
2 (7%)	49.7 – 56.0	50.1 – 55.0	45.1 – 55.0	52.1 – 58.0
3 (12%)	56.1 – 64.0	55.1 – 60.0	55.1 – 60.0	58.1 – 63.0
4 (17%)	64.1 – 68.0	60.1 – 65.0	60.1 – 65.0	63.1 – 66.0
5 (20%)	68.1 – 72.0	65.1 – 70.0	65.1 – 70.0	66.1 – 71.0
6 (17%)	72.1 – 76.0	70.1 – 75.0	70.1 – 74.9	71.1 – 74.0
7 (12%)	76.1 – 80.0	75.1 – 80.0	75.0 – 75.0	74.1 – 75.0
8 (7%)	80.1 – 84.0	80.1 – 85.0	75.1 – 80.0	75.1 – 80.0
9 (4%)	> 84.0	> 85.0	> 80.0	> 80.0

5. Example Applications

In this section, several practical examples of the administration of the MNT in different applied settings are briefly reported.

5.1 Stress susceptibility in hypertensive patients

The MNT was developed to evoke stable reactions of the arterial blood pressure. This diagnostic approach was used to distinguish between individuals with essential hypertension (elevated blood pressure reactivity without recognizable organic reasons) and individuals with renal hypertension (high blood pressure caused by renovascular disease, which results in a systemic constriction of the blood vessels). The MNT appeared to be suitable as a stable stress stimulus (Johannes, Eichhorn & Fischer, 1994; Johannes et al., 1995).

5.2 Monitoring of Autonomic Response Pattern during long-term confinement and other space analogues

This diagnostic procedure was foreseen for space applications. Therefore, numerous terrestrial analogue studies were used to test hard- and software for this application. It should provide response pattern specific calibrated values for the psychophysiological reactivity assessed during the training of manual docking of a spacecraft to a space station. Specifically, the psychophysiological arousal value (PAV) was attempted to assess the arousal level of astronauts as indicators of their effort. The very first study was run as a 135-Days ESA / IBMP-Study 1994/95 in Moscow, called “HUBES”. Thereafter followed the 90-days study “EKOPSI-95”, “SFINCSS”, and “MARS500”, all taking place in Moscow (Johannes, Salnitski, Lukjanuk, Gunga, & Kirsch, 2001). At DLR in Germany bedrest studies were conducted, the last up to now is called “AGBRESA”.

5.3 Monitoring of Autonomic Response Pattern during long-term space flights

Beginning with the MIR station, the MNT was involved in numerous space experiments, e.g. named “Regulation” (Johannes et al., 1998; Johannes, Salnitski, Polyakov, & Kirsch K., 2003). On the ISS, these experiments were continued (Johannes et al., 2016).

5.4 MNT as a central component of psychophysiological calibration procedures

The MIR/ISS-experiment "Regulation" was carried out as psychophysiological calibration for the load assessment during docking training. The MNT is a central component in this calibration procedure. It appeared to be a very reliable and practical stimulus to evoke an elevated psychophysiological response in all repetitions (Johannes, Salnitski, Gunga, & Kirsch, 2000; Johannes et al., 1998).

5.5 MNT as a central component of voice calibration procedures

For the assessment of mental load during manual docking training, analysis methods for voice and integrated psychophysiological arousal assessment were developed. In this context, the MNT served as a central component of the voice calibration procedure (Wittels, Johannes, Enne, Kirsch, & Gunga, 2002). Further examples can be found in Johannes et al. (2007) or Johannes et al. (2019).

5.6 MNT response patterns as indicators of effort

In this context, a cluster analysis was carried out to determine whether individuals differ with respect to their response styles when performing the MNT tasks. According to this analysis, subjects could be distinguished into three clusters representing different performance patterns. The main difference was the frequency of correct identifications of "Error" system states (correct rejections). Subjects of only one response cluster performed with equal quality when identifying "OK" and "Error" system states. In the other two clusters especially the correct identification of "Error" states was poorer (Johannes, Bronnikov, et al., 2017).

5.7 MNT as reference task for further psychophysiological monitoring methods

The set of psychophysiological methods for the objective, unobtrusive monitoring of effort during any relevant behaviour of astronauts and pilots is still under development and will remain in the focus of further research. A continuation of the assessment of interindividually comparable psycho-physiological arousal (PAV) as well as latest methods to include EEG parameters is work in progress (Johannes & Gaillard, 2014; Johannes, Rothe, et al., 2017). Since 2015, these new methods are applied in space as project NEUROLAB2010 on ISS (Johannes et al., 2021).

6. References

Hoermann, H.-J. & Damos, D.L. (2019). The use of a perceptual speed test in civilian pilot selection. *Proceedings of the 20th International Symposium on Aviation Psychology*. Wright State University, Dayton/OH, May 7-10.

Hoermann, H.-J., Noser, P. & Stelling, D. (2018). A psychometrician's dream: Test validities in an unrestricted sample of future Swiss Air Force pilots. *Proceedings of the 33rd EAAP Conference (pp 205-212)*. Dubrovnik/Croatia, September 24-28.

Hoermann, H.-J. & Maschke, P. (1996). On the relationship between personality and job performance of airline pilots. *The International Journal of Aviation Psychology*, *6*, 171-178.

Johannes, B., Bronnikov, S., Bubeev, Y.A., Dudukin, A.V., Hoermann, H.-J., Frett, T., Rittweger, J., & Gaillard, A.W.K. (2017). A tool to facilitate learning in a complex manual control task. *International Journal of Applied Psychology, 7(4*), 79-85.

Johannes, B., Bronnikov, S.V., Bubeev, Y.A., Kotrovskaya, T.I., Shastlivtseva, D.V., Piechowski, S., Hoermann, H.-J., Rittweger, J., & Jordan, J. (2019). Operational and experimental tasks, performance, and voice in space. *Aerospace Medicine and Human Performance*, *90(7),* 624-631.

Johannes, B., Bubeev, Y.A., Kotrovskaya, T.I., Bronnikov, S.V., Piechowski, S., Elmenhorst, E-M., Rittweger, J., Jordan, J., & Gaillard, A.W.K. (2021). Assessing cognitive capacity by P3 during a complex manual control task. *Journal of Psychophysiology*, *35(1),* 43-50.

Johannes, B., Eichhorn, C. & Fischer, F. (1994). A complex experimental assessment for objective hierarchical description of hierarchical psychophysiological behavior as human regulatory phenotype. *Journal of Gravitational Physiol*ogy,*1(1)*, 73-74.

Johannes, B. & Gaillard, A.W.K. (2014). A methodology to compensate for individual differences in psychophysiological assessment. *Biological Psychology*, *96*, 77-85.

Johannes, B., Rothe, S., Gens, A., Westphal, S., Mulder, E., Rittweger, J., & Ledderhos, C. (2017). Psychophysiological assessment in pilots performing challenging simulated and real flight maneuvers. *Aerospace Medicine and Human Performance*, *88(9),* 834-840.

Johannes, B., Salnitski, V.P., Dudukin, A.V., Shevchenko, L. G., Shebuchev, A. E., & Bronnikov, S. V. (2016). Performance Assessment in the Experiment PILOT on-board Space Stations MIR and ISS. *Aerospace Medicine and Human Performance, 87(6),* 534-544.

Johannes, B., Salnitski, V.P., Gunga, H.C., & Kirsch, K. (2000). Voice stress monitoring in space- possibilities and limits. Workshop Human factors in space, Tokyo, 7.-9.7.1999, *Aviation, Space and Environmental Medicine*, *71(9)*, Section II, A58-A65, September 2000.

Johannes, B., Salnitski, V.P., Haller, H., Wilke, D., Fischer, F., & Schlykova, L. (1995). Comparison of voice stress reactivity under psychological stress test and simulated MIR-docking-maneuver. *Journal of Gravitational Physiology, 2(1),* 107-108.

Johannes, B., Salnitski, V.P, Korsun, V., Kaleri, A.J., Kirsch, K., Schewtschenko, O., Fischer, F., Dudukin A.V, & Gunga, H.-C. (1998). Psychophysiological reactivity under MIR-simulation and real micro-G. *Acta Astronautica, 42(1-8),* 215-218.

Johannes, B., Salnitski, V.P., Lukjanuk, V.V., Gunga, H.-C. & Kirsch, K.A. (2001). Monitoring of Autonomic Outlet Types during long-term confinement. In: V.M. Baranov (Ed) *Simulation of extended isolation: Advances and problems*. Slovo, Moscow, pp 51–60.

Johannes, B., Salnitski, V.P., Polyakov, V.V., & Kirsch K. (2003). Changes in the autonomic reactivity pattern to psychological load under long-term microgravity - Twelve men during 6-month spaceflights. *Aviakosmicheskaia i ekologicheskaia meditsina (Aerospace and Environmental Medicine), 37(3),* 6-16.

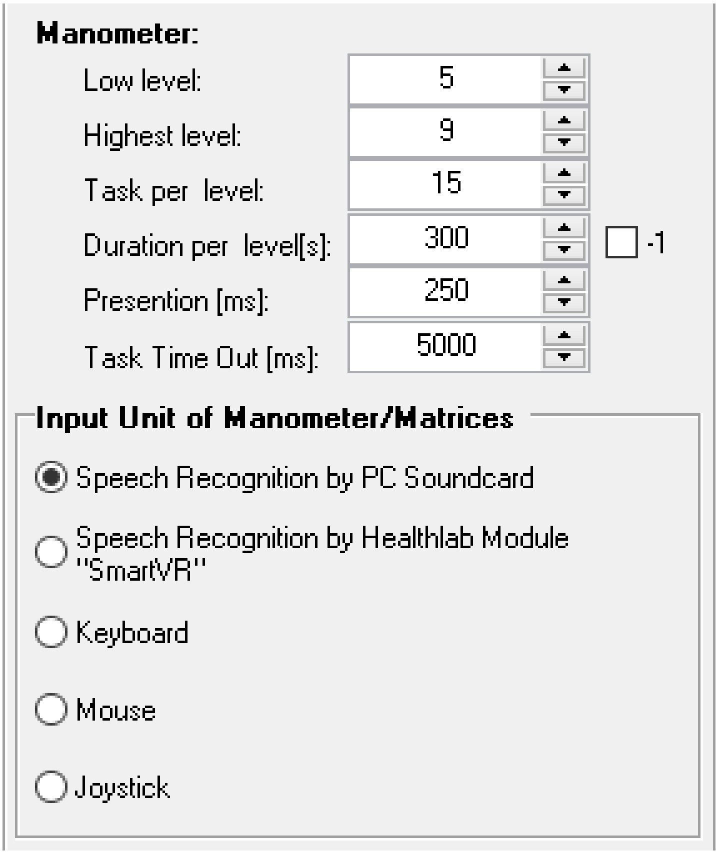
Johannes, B., Wittels, P., Enne, R., Eisinger, G., Castro, C., Thomas, J., Adler, A., & Gerzer, R. (2007). Non-linear function model of voice pitch dependency on physical and mental load. *European Journal of Applied Physiology, 101*, 267-276.

Swets, J.A. (Ed.) (1964). *Signal detection and recognition by human observers*. New York: Wiley.

Wittels, P., Johannes, B., Enne, R., Kirsch, K., & Gunga, H.-C. (2002). Voice monitoring to measure emotional load during short-term stress. *European Journal of Applied Physiology, 87*, 278-282.

7. Appendix

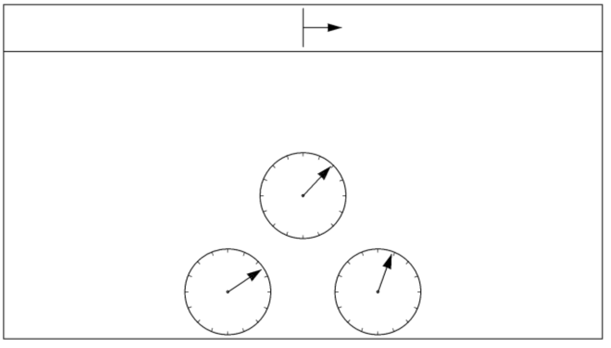
7.1 Configuration page of the MNT software



This figure shows the MNT configuration page with the standard values. It can be used to adapt the stimulus presentation parameters for special purposes. If the box "-1" is checked, the test will not adopt a dynamic presentation time. In this case, a static value can be entered for the presentation time (in the example 250ms).

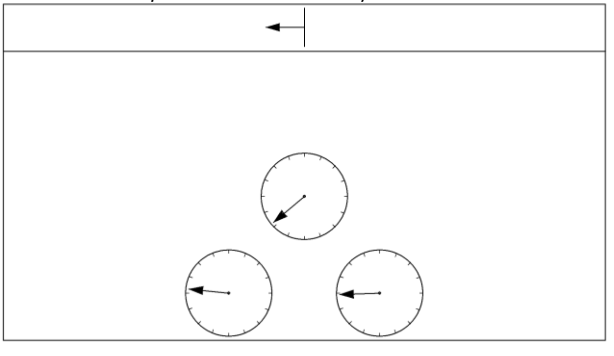
7.2 MNT Instruction text

INTRODUCTION:

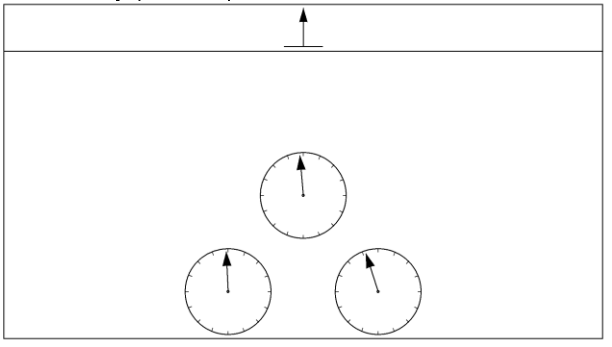


On the computer screen varying numbers of such pressure gauges will appear. The arrow shown in the upper part of the screen, indicates the direction for all indicators on the other pressure gauges. This direction will change from trial to trial. In this trial all indicators point to the right.

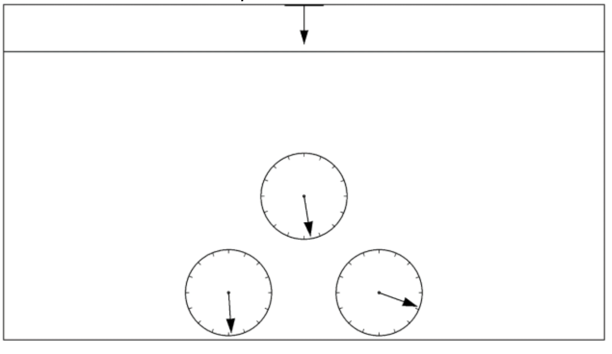
In this example all indicators point to the left.



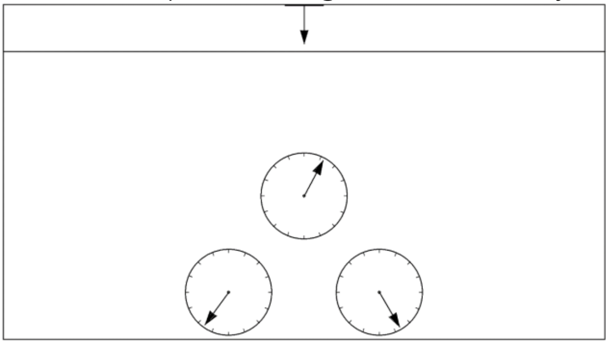
Here, they point upwards.



Here, all indicators point downwards.



When all indicators point in the given direction, the system is working properly. For each trial the indicators will point in a different direction. Should at least one indicator not point in the given direction, as you can see here …



… there is an error in the system. The arrow points downwards, but at least one indicator of the pressure gauges points upwards.

Your task is to continuously check the pictures for errors in the system and tell the computer whether the system is OK, or whether there is an Error in the system. You will receive feedback about your performance. If your input is correct, you will get < low beep>, if your input is wrong, you will get < high beep >.

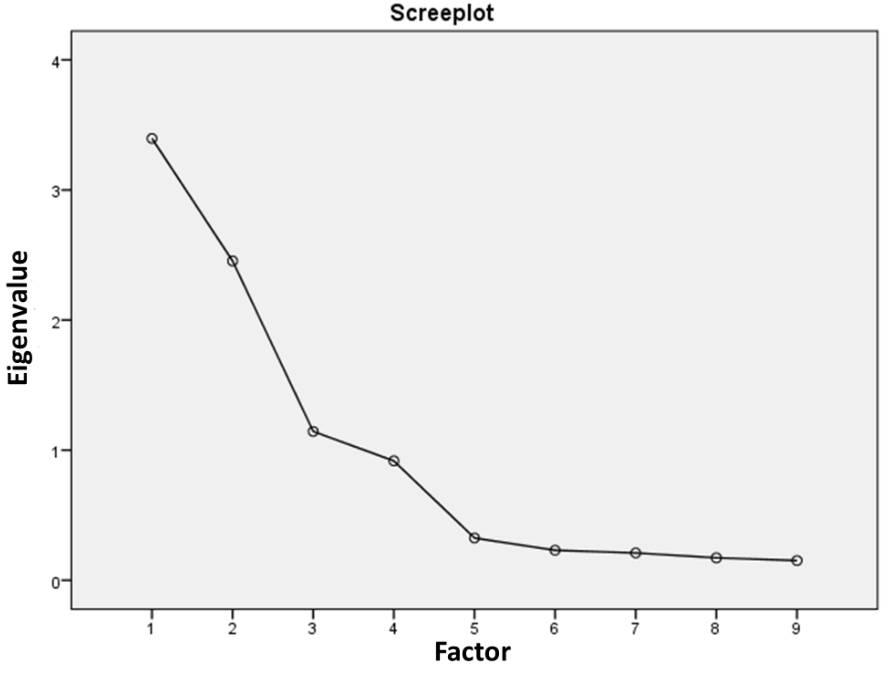
PRACTICE:
Attention, now you will be given a demonstration of the task.
Now, trials with given solutions will be presented. This is done for practice and comparison. Trails will alternate constantly and you have to answer "OK" or "ERROR". The correct solution for the trials will be presented in the upper right corner of the screen. Please, speak you commands in a loud and clear manner. The feedback tone has no meaning during this part of the exercise.

MAIN PART:
The main part of the test begins now. The trials will alternate in shorter time intervals. The higher tone will be presented after incorrect answers.

7.3 Distribution statistics for the MNT variables in the calibration study (N = 2084)

	Level 5		Level 7		Level 9	
Variable	Mean	StDev	Mean	StDev	Mean	StDev
MRT	1.51	0.37	1.41	0.31	1.51	0.31
SRT	0.81	0.30	0.35	0.18	0.37	0.17
MPT	0.96	0.33	0.50	0.22	0.55	0.25
SPT	0.88	0.26	0.27	0.10	0.27	0.11
HIT%	0.72	0.14	0.71	0.15	0.68	0.15
COR%	0.67	0.15	0.66	0.17	0.64	0.17
FAA%	0.28	0.14	0.29	0.15	0.32	0.15
MIS%	0.33	0.15	0.35	0.17	0.36	0.17
QUALI	0.69	0.10	0.68	0.10	0.65	0.10
O-QUALI		Mean = 0.68			StDev = 0.08	

7.4 Scree plot of a Principal Axis Factor Analysis of MNT performance variables (N = 2084)



7.5 Norm tables for level-specific MNT variables

	Stanines								
MNT	1	2	3	4	5	6	7	8	9
MRT.5	0.92	1.14	1.29	1.42	1.57	1.74	1.91	2.11	>2.11
SRT.5	0.35	0.44	0.54	0.74	0.87	1.00	1.17	1.36	>1.36
MPT.5	0.57	0.63	0.73	0.84	0.97	1.10	1.30	1.61	>1.61
SPT.5	0.51	0.58	0.63	0.77	0.97	1.09	1.22	1.36	>1.36
HIT%.5	0.45	0.56	0.63	0.70	0.77	0.82	0.88	0.93	>0.93
COR%.5	0.41	0.50	0.57	0.64	0.71	0.79	0.85	0.91	>0.91
FAA%.5	0.07	0.12	0.18	0.23	0.30	0.38	0.44	0.55	>0.55
MIS%.5	0.09	0.15	0.21	0.29	0.36	0.43	0.50	0.59	>0.59
QUALI.5	0.50	0.56	0.64	0.68	0.72	0.76	0.80	0.84	>0.84
MRT.7	0.87	1.01	1.15	1.30	1.46	1.61	1.80	2.02	>2.02
SRT.7	0.15	0.19	0.24	0.29	0.36	0.48	0.63	0.86	>0.86
MPT.7	0.26	0.30	0.36	0.43	0.55	0.74	1.10	1.72	>1.72
SPT.7	0.16	0.18	0.21	0.23	0.29	0.36	0.46	0.64	>0.64
HIT%.7	0.44	0.53	0.60	0.67	0.75	0.82	0.89	0.94	>0.94
COR%.7	0.33	0.45	0.55	0.62	0.70	0.78	0.86	0.92	>0.92
FAA%.7	0.06	0.11	0.18	0.25	0.33	0.40	0.47	0.56	>0.56
MIS%.7	0.08	0.14	0.22	0.30	0.38	0.45	0.55	0.67	>0.67
QUALI.7	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	>0.85
MRT.9	0.88	1.08	1.25	1.40	1.56	1.72	1.92	2.19	>2.19
SRT.9	0.17	0.21	0.25	0.31	0.40	0.51	0.63	0.90	>0.90

MPT.9	0.28	0.33	0.39	0.48	0.62	0.85	1.22	1.85	>1.85
SPT.9	0.16	0.18	0.20	0.23	0.29	0.38	0.48	0.71	>0.71
HIT%.9	0.40	0.50	0.57	0.64	0.71	0.79	0.86	0.92	>0.92
COR%.9	0.33	0.43	0.50	0.60	0.69	0.77	0.83	0.99	>0.99
FAA%.9	0.08	0.14	0.21	0.29	0.36	0.43	0.50	0.60	>0.60
MIS%.9	0.00	0.17	0.23	0.31	0.40	0.50	0.57	0.67	>0.67
QUALI.9	0.45	0.55	0.60	0.65	0.70	0.74	0.75	0.80	>0.80