

# Effects of intraoperative breaks on mental and somatic operator fatigue: a randomized clinical trial

Carsten Engelmann · Mischa Schneider ·  
Clemens Kirschbaum · Gudela Grote ·  
Jens Dingemann · Stefan Schoof · Benno M. Ure

Received: 10 May 2010 / Accepted: 18 August 2010 / Published online: 11 September 2010  
© Springer Science+Business Media, LLC 2010

## Abstract

**Background** Intermittent work breaks are common in fields with high workload but not yet for surgeons during operations. We evaluated the effects of intraoperative breaks during complex laparoscopic surgery (5 min every half hour) on the surgeon.

**Methods** Fifty-one operations were randomized to a scheme with intraoperative breaks and release of the pneumoperitoneum (intermittent pneumoperitoneum (IPP)) or conventional conduct (CPP). Stress hormones and  $\alpha$ -amylase were determined in the surgeon's saliva pre-, intra-, and postoperatively. Mental performance and error scores, musculoskeletal strain, and continuous ECG were secondary endpoints.

**Results** Regular intraoperative breaks did not prolong the operation (IPP vs. CPP group: 176 vs. 180 min,  $p > 0.05$ ). The surgeon's cortisol levels during the operation were reduced by  $22 \pm 10.3\%$  in the IPP vs. the CPP group ( $p < 0.05$ ). There were significantly fewer ( $p < 0.05$ ) intraoperative events in the IPP vs. the CPP group, which

yielded higher  $\alpha$ -amylase peaks. The pre- to postoperative increase in the error rates of the bp-concentration test was fourfold reduced in the IPP group ( $p = 0.052$ ). The relevant locomotive strain-scores were grossly reduced by IPP ( $p < 0.001$ ).

**Conclusions** Our data support the idea that work breaks during complex laparoscopic surgery can reduce psychological stress and preserve performance without prolongation of the operation time compared with the traditional work scheme.

**Keywords** Intermittent pneumoperitoneum · Work breaks · Complex laparoscopic surgery · Stress hormones · Concentration–error performance

In many work environments where a high performance has to be delivered over a sustained time period, short, intermittent breaks are used to control workload and reduce error. Air traffic controllers or U.N.-simultaneous translators are supposed to have breaks every 20 min [1, 2]. This is not yet common practice for long operations. The image of the infatigable surgeon has powerful sociocultural roots that extend from historical novels to contemporary broadcasts [3].

However, the demands to the surgeon have changed since video techniques have added a high spatial recognition task load to the classical demands of persistence and precision [4]. It results from the constant switch between a two-dimensional televised feedback and three-dimensional instrument operation. The visual input is further restricted by smoke from diathermy and oligochrome color definition. Processing the incoming sensory information, real-time decision making [5], and the required motor output may add up to a maximal recruitment of operator

---

C. Engelmann (✉) · M. Schneider · J. Dingemann · B. M. Ure  
Department of Pediatric Surgery, Hannover Medical School,  
30625 Hannover, Germany  
e-mail: ceng@bdc.de

C. Kirschbaum  
Department of Biopsychology, Technical University Dresden,  
01062 Dresden, Germany

G. Grote  
Department of Management Technology and Economics,  
ETH Zurich, Zurich 8032, Switzerland

S. Schoof  
Department of Pediatric Cardiology, Hannover Medical School,  
30625 Hannover, Germany

resources. An experienced onlooker may recognize this “work load red line” [6, 7]; the operator perseveres or proceeds to a false move.

Incident prevention is one, fatigue the other, less spectacular issue of high workload: the effort has to be raised to maintain performance constant during an entire procedure [8, 9]. Many will confirm the observation that the development of fatigue is correlated with difficulty [10]: i.e., after an hour of difficult laparoscopy one may feel “drained of energy” but may yet have barely started on the day’s operating list.

Breaks are the oldest method to counteract fatigue. Alpinists discuss break schemes as a key success factor for the historic first climb on Mt. Everest; traditional climbing involved a long break every 4 h. Contrastingly, the successful party in 1953 moved according to the so-called “Sherpa scheme,” which means 50 min fast climb followed by a mandatory break of 10 min (Reinhold Messner, e-mail 10/2007).

Inspired by this scheme, we tested the hypothesis that breaks during long duration laparoscopic operations lower the surgeon’s stress response and improve performance and well-being [11].

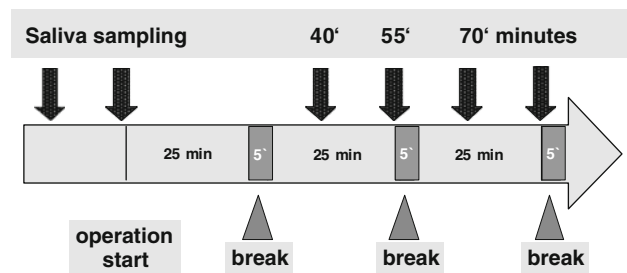
## Methods

We designed a randomized, controlled trial in which complex laparoscopic operations (Table 1, inclusion criteria) in children were subjected to a break scheme consisting of 25-min work periods followed by 5-min unstructured breaks (Fig. 1), during which the pneumoperitoneum was released (IPP group). The control group consisted of operations performed without breaks, termed CPP group (continuous pneumoperitoneum).

**Table 1** Cohort data

	IPP	CPP
Patients		
Total	26	25
Age < 1 yr	11	10
Procedure		
Hepatojejunostomy/Kasai	3	8
Duodenal reconstruction	1	0
(redo) Fundoplication	8	6
Partial nephrectomy	7	5
Pyeloplasty	7	6
Past history		
Previous abdominal surgery	7	8

IPP intermittent pneumoperitoneum, CPP conventional conduct



**Fig. 1** IPP break scheme protocol. Operations were structured into successive 30-min blocks. These consisted each of 25 min of work followed by a 5-min break with release of the pneumoperitoneum. Saliva was sampled just before and 10 min after the break. Preoperative samples were taken at 7.30 a.m. and before the incision was made

## Surgeon endpoints

Surgeons ( $n = 7$ , 6 man/1 woman), all with >300 laparoscopic procedures experience), were monitored biometrically, by task measures, and by use of behaviorally anchored self-report questionnaires. The primary endpoint was the hormonal stress response of the operating surgeon. The strain markers, cortisol, amylase (served as surrogate parameter for adrenal activation [12]), and testosterone, including the control dehydroepiandrosterone (DHEA), were measured from the surgeon’s saliva [13]. Sample collection was performed by passing unused straws behind the surgeon’s mask. Sample moments were placed after the breaks/events according to the physiological kinetics of hormones in saliva (Fig. 1).

Steroids were determined using competitive chemiluminescence immunoassays (IBL, Hamburg, Germany [www.ibl-international.com](http://www.ibl-international.com)). Alpha-amylase was assayed by ELISA (Roche Diagnostics, Mannheim/Germany; [www.roche.com](http://www.roche.com)). Detailed time-resolved mission protocols from each operation scenario were recorded [14]. Analysis of the results was focused on the interval where fatigue first becomes relevant (beginning 1 h after the start of the operation).

Secondary endpoints for the surgeon included:

1. Biometry by continuous ECG (Getemed Cardio Memory software, Berlin-Teltow, Germany). Sudden intraoperative “events” were defined by a heart-rate increase >30% or >20 bpm *plus* a specified context from the mission scenario record (loss of exposure, bleeding, difficulties in performing a motor action, such as dissection or knot tying).
2. Psychometric studies: concentration and performance were tested pre- and postoperatively for 3 min each (bp-test [15]). In the bp-test the error rate (concentration) is defined as the percentage of false out of the

total number of completed markings. The test is designed as supramaximal task (the number of randomly compiled rows of b and p signs is higher than the tested subject can conclude in the given time), such as the overall capacity can also be determined.

Self-ratings of own satisfaction, performance and fatigue were measured with Likert item rating scales using questionnaires derived from the NASA Task load index [16]. To eliminate possible selection bias, scales were behaviorally anchored (scale plus explicit verbally descriptive answer) where appropriate [17].

3. Musculoskeletal system (MSS) and ophthalmologic strain: wear-off effects of the relevant MSS elements were evaluated in detail on rating scales [6] ranging from 1–10 (1 = uninfluenced; 10 = maximum fatigue).

General:

The trial was approved by Hannover Medical School ethical committee (No. 4165; 3-2-2006) and registered (ClinicalTrials.gov; NCT 01009372). Informed, written consent was obtained from parents and the operators. Randomization was done in the operating room before the incision by a research fellow uninvolved with clinical duties by picking one sealed envelope containing the group information.

The insufflation pressure for the pneumoperitoneum was  $p = 8$  mm mercury provided by an Electronic Endoflator (www.karlstorz.de). Operations started in the morning (10:29 a.m.  $\pm$  1:44 (IPP) vs. 10:38 a.m.  $\pm$  1:44 (CPP)). The duration of the operations (incision–skin closure) was corrected for their complexity by multiplication with a correction factor<sup>1</sup> assigned to each type of procedure (e.g., fundoplication vs. choledochal cyst).

Statistics and computing: The sample size was calculated for the primary endpoint saliva steroids to  $n = 21$  patients per group based on preliminary experiments with five operations by IPP and CPP (software: nQuery, http://www.statsol.ie). The source database was closed February 2008. The independent *t* test and for small samples the Mann–Whitney test were performed using SPSS (www.spss.com). Correlations were determined according to Pearson’s algorithm. For integrations (determinations of

the area under the curve, AUC), we employed Microcal origin version 7.5. (www.microcal.com).

Results

IPP and CPP operations were equally distributed among the surgeons ( $n_{IPP}/n_{CPP}$ : A = 11/12, B = 6/4, C = 5/4, D = 2/1, E = 0/1, F = 1/3); also the groups were comparable regarding other human factors (tobacco/coffee/tea-consumption, meals; data not shown).

Surgeon’s primary endpoints: saliva stress hormones

The saliva cortisol levels of surgeons operating without breaks were significantly higher by 22% than those of surgeons operating with breaks (from 55 to 100 min after the start of the operation,  $p < 0.05$ ; Fig. 2).

DHEA levels were not altered significantly (AUC:  $30.3 \pm 12$  (IPP) vs.  $35.2 \pm 18$  (CPP),  $p > 0.05$ ). Corrected for interindividual baseline variations testosterone showed a graph comparable to cortisol (AUC  $21 \pm 11$  (IPP) vs.  $26 \pm 13$  (CPP),  $p > 0.05$ ). For the female surgeon, there were highly significant differences in favor of the IPP scheme (AUC:  $33.1 \pm 9$  (IPP) vs.  $88 \pm 6$  (CPP),  $n = 7$ ,  $p < 0.001$ ).

Saliva hormones and intraoperative events: There were significantly more intraoperative events in the CPP group ( $n_{IPP} = 64$  vs.  $n_{CPP} = 91$ ,  $p < 0.05$ ) and event-related alpha amylase peaks were higher in the IPP group ( $164 \pm 73$  u/l vs.  $152 \pm 69$  u/l (CPP),  $p > 0.05$ ). Inversely, the event-related cortisol concentrations in saliva

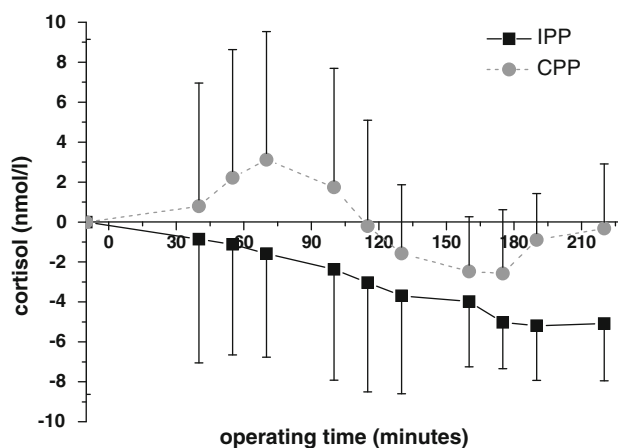


Fig. 2 Saliva cortisol graphs over time for the intermittent (IPP) and continuous (CPP) pneumoperitoneum groups. The start of the operation was normalized to zero. The area under the curve in the interval from 55 to 100 min after the start of the operation decreased by 22% ( $AUC_{t = +55 \text{ min} - t125 \text{ min}} = 57.8 \pm 29$  (IPP) vs.  $70.4 \pm 37$ ;  $p < 0.05$ )

<sup>1</sup> Correction factor for complexity (cf): Mean results for every procedure were referenced to the most abundant operation in this series (pyeloplasty, which was assigned the factor 1.0).

$$Cf = \frac{\text{over all mean duration}_{(\text{operation type})}}{\text{over all mean duration}_{\text{pyeloplasty}}}$$

$$\Sigma cf_{IPP \text{ or } CPP\text{-group}} / n_{\text{Operations IPP or CPP group}}$$

$$= \text{mean } cf_{IPP \text{ or } CPP \text{ group}}$$

$$\text{mean } cf_{IPP \text{ or } CPP \text{ group}} \times \text{mean duration}_{IPP \text{ or } CPP \text{ group}}$$

$$= \text{corrected duration}$$

**Table 2** Summary of saliva results

	IPP	CPP	<i>p</i> Value
Cortisol	▼	▲	<0.05
DHEA	▼	▲	NS
Testosterone	▼	▲	<0.001
Amylase	▲	▼	NS

Arrows indicate pre- to postoperative evolution of the parameter; ▲ increase, ▼ decrease

NS not significant

were lower in the IPP group ( $87 \pm 39$  mmol/l (IPP) vs.  $103 \pm 59$  mmol/l (CPP),  $p < 0.05$ ). The first derivation of the amylase time–concentration graph (marker for steadiness in the graph/operative conduct) has more zero-crossings in the CPP than in the IPP group ( $n = 4.5 \pm 1.8$  (CPP) vs.  $n = 3.7 \pm 1.9$  (IPP),  $p = 0.06$ ). The saliva results are summarized in Table 2.

### Surgeon's secondary endpoints

For the biometry by continuous ECG we found a significantly higher start–end variability in the conventional group ( $0.9 \pm 9$  (IPP) vs.  $6 \pm 10.7$  (CPP),  $p < 0.05$ ). Other parameters did not significantly differ among groups.

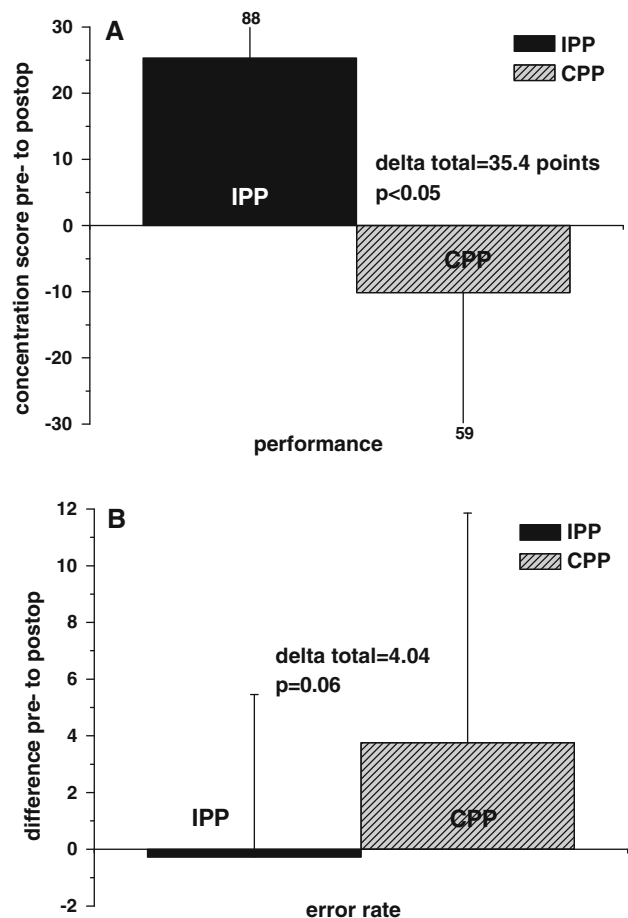
The concentration and performance test results (d2-test) in the CPP group from pre- to postoperatively have a decreased performance score and increased error rate. In contrast, the IPP break scheme maintained the overall performance and error rates. The intergroup differences for “performance” were significant ( $p < 0.05$ ), and those for “error” were at the threshold to significance ( $p = 0.06$ ) because of one tested individual's problem solving behavior, which markedly traded off accuracy for speed [18]. However, the IPP error rate was threefold lower than the CPP error rate (Fig. 3).

Instantaneous self assessment scales (Fig. 4) showed a lesser increase in fatigue from pre- to postoperation in the IPP group ( $+15\% \pm 9\%$  (IPP) vs.  $+28\% \pm 17\%$  (CPP),  $p < 0.005$ ). The perceived impairment by fatigue was lower in the IPP group. There also was a decrease in perceived stress.

For the musculoskeletal strain and pain scores, significant improvements ( $p < 0.001$ ) were recorded throughout the upper extremities locomotive and the trunk's static elements (summary scores in Fig. 5). The fatigue score for the eyes improved by 50% from  $1.92 \pm 2.48$  (CPP) to  $0.96 \pm 1.27$  (IPP) ( $p = 0.09$ ).

### General data

The mean complexity factors of the surgical procedures ( $0.98 \pm 0.98$  (IPP break scheme) vs.  $1.01 \pm 0.33$  (CPP))

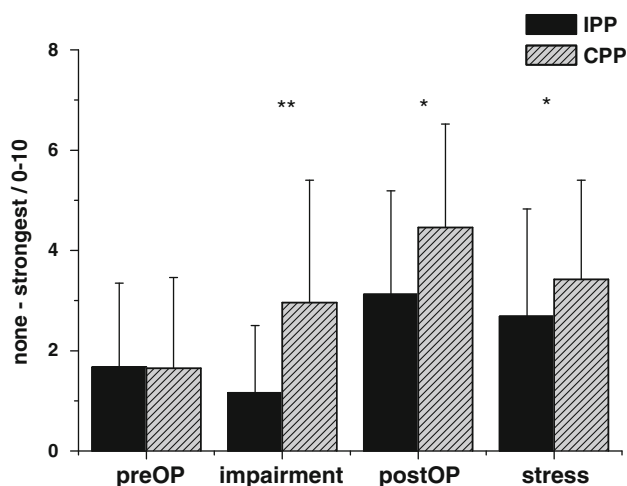


**Fig. 3** Pre- to postoperative results of the d2 concentration–performance test. **A** Pre- to postoperative changes in the concentration score in the IPP vs. CPP group. **B** Pre- to postoperative change of the error rate referred to the baseline. *p* values are stated for the intergroup difference CPP vs. IPP

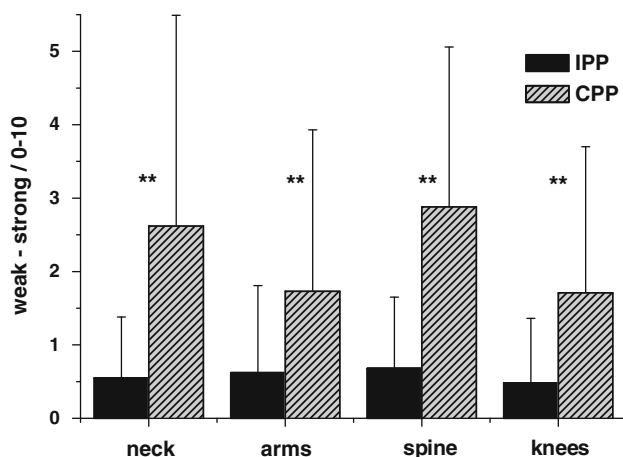
were comparable ( $p > 0.05$ ) between groups. The mean operation time corrected for complexity was  $176 \pm 45$  min in the IPP versus  $180 \pm 49$  min in the CPP group ( $p > 0.05$ ). There were two postoperative complications (none of it infectious) and two conversions to open surgery in each group.

### Discussion

The work routine under review in this paper is nothing less than the traditional surgical operation. Our data provide for the first time a rational basis for compulsory breaks *within* surgical operations. A part of the biometry relies on saliva hormone and enzyme data. In the past the majority of saliva measurements were performed only *after* the task was completed [19]. The present samples were obtained *during* the operation with a straw from behind the surgeon's mask, which proved to be a low intrusion on the



**Fig. 4** Self-rated psychometric data on the influence of the IPP break scheme on fatigue and stress. PreOP = preoperatively perceived fatigue, impairment: perceived intraoperative impairment by fatigue, postOP = postoperatively perceived fatigue, stress = rating for the maximal intraoperative stress at the resented most precarious moment; \*  $p < 0.05$ ; \*\*  $p < 0.001$



**Fig. 5** Intensity of musculoskeletal system pain and strain. Scores from locomotive units relevant for laparoscopy. Displayed scores are summary data calculated from detailed evaluation of the muscle groups. Example: Arms = scores for shoulder, elbow, wrist, hand, fingers (left and right, respectively). Spine: thoracic and lumbar spine. \*\*  $p < 0.001$

primary task. In all instances, the surgeon continued to observe the monitor while being sampled. This method may be recommended for future use.

Cortisol, the classical hormone involved in stress [20], is liberated into the saliva independently of the saliva flow rate [21]. Our results were corrected for the circadian effects of different starting times. They demonstrate for the group making regular breaks significantly lower cortisol liberation from approximately 1 h after the start of the operation and onwards—the time interval where fatigue effects commonly occur. The differences persist to the end of the operation, even though the statistical significance is

lost due to fluctuating CPP/IPP group sizes because—with time—some operations are finished, whereas others continue.

Besides its proverbial role in the achievement of complex tasks, testosterone is increasingly regarded as a valid workload-marker, especially in women [22, 23]. We confirmed our cortisol results; there was a strong trend toward lower intraoperative testosterone levels in those surgeons operating according to the IPP break scheme. For the female values, not impeded by the strong interindividual baseline variations of the males, this difference in favor of the IPP scheme was highly significant.

The time–concentration graph of dehydroepiandrosterone (DHEA) termed “anti-stress” hormone or a “functional” negative control did not differ significantly between the IPP and CPP groups [24].

We next looked at sudden threatening “events” defined by the context and a significant HR increase. Many readers will confirm from their own experience that such “events” do induce changes in saliva composition; perceptions range from “bitterness in the mouth” to “thickened saliva.” These subjective sensations corresponded to measurable changes: (1) in the IPP group, the total number of events was significantly lower than in the conventional group; and (2) in the amylase time–concentration graph, there were fewer slope changes pointing to a “quieter” conduct of the intervention. Mean cortisol after events was higher in the CPP group. However, rather surprisingly the amylase peaks (a surrogate parameter for adrenergic activation) associated to events were higher ( $p = 0.06$ ). This is a new observation and may indicate that an operator working according to a break scheme retains a higher capacity for adrenergic reaction to unexpected adversities than a fatigued one.

The continuous trace ECG showed that there was a significantly less heart rate increase from start to the end of the operation in those participating in the IPP break scheme [25]. Because the ECG is prone to contamination by anxiety and physical activity we may only conclude, very broadly, that IPP decreased both physical and mental fatigue [26].

Postoperative performance–concentration measurements indicated a conservation of performance and significantly lower error rates in those surgeons adhering to the IPP scheme. The importance of this IPP effect on the objective performance markers is underlined by the fact that its statistical significance was not ablated by the interindividual differences in the d2 test problem-solving approaches among the tested surgeons, which increased data variance [18]. The importance of the concentration–performance study results is underlined by the fact that surgery, compared with many other fields, is a domain with very high penalty for error in instantaneous signal detection/vigilance tasks on sensory channels [27].

Subjective performance markers, postoperative satisfaction with the procedure, fatigue, and well-being were evaluated on behaviorally anchored rating scales. In our study preoperative values were identical among both groups, whereas postoperatively in the IPP group fatigue and perceived impairment by fatigue were significantly reduced. The maximal perceived stress also was decreased, which can be explained by the fact that fatigue raises the workload [6, 28].

The surgeon's pain and strain scores of the muscular skeletal system after IPP operations were highly significantly lower than those after CPP surgery. This was a consistent finding even from those operators who disliked the IPP break scheme in whom an emotional contamination toward the opposite direction could be expected.

It was finally of major concern whether the benefits for the surgeon were at the cost of the patient. Surprisingly, the breaks did not prolong the operating time at all.

In conclusion, the IPP break scheme decreased the surgeon's stress hormones, the count of intraoperative events, and the objective error–performance scores. It increased the surgeon's well-being without any disadvantageous prolongation of the operation time.

The effect sizes reported were all seen only for the surgeon. We are currently evaluating the effect of breaks and the interruption of the pneumoperitoneum on pediatric patients.

**Acknowledgments** We thank D. Auge and A. Osthaus, MHH for expert help with ECG systems and cardiac output determination, Dr. Judith Orsanu from NASA's Ames research center in California for discussion, and Prof. R. Dietrich, Humboldt University Berlin/Germany for information on simultaneous translation and Mr. Bill Groundwater, Kirkwall/Orkney for help with the manuscript.

**Disclosures** This study was financed from the Hannover Medical School general health care budget. Drs. Engelmann, Schneider, Dingemann, Schoof, Ure, and Professors Kirschbaum and Grote have no conflicts of interest or financial ties to disclose.

## References

- Farmer E, Brownson A (2003) Review of workload measurement, analysis and interpretation methods (CARE-Integra-TRS-130-02-WP2). 1130 Brussels, Belgium, Rue de la Fusée 96: Eurocontrol
- AICC. Guidelines for simultaneous translation. [http://www.aiic.de/english/suche\\_arten2.php](http://www.aiic.de/english/suche_arten2.php). Accessed 2008
- Kilian H (1957) *Hinter uns steht nur der Herrgott* ["Behind us only god"]. Kindler Publishers, München, pp 1–385
- Brookhuis KA, de Waard D (2002) On the assessment of (mental) workload and other subjective qualifications. *Ergonomics* 45:1026–1030, discussion 1042–1046
- Janssen WH (2004) Route planning and route guidance: a review of the literature (Report IZF 1979-C13). Instituut voor Zintuigfysiologie, Soesterberg
- O'Donnel RD, Eggemeier FT (1986) Workload assessment methodology. In: Boff KR, Kaufmann L, Thomas JP (eds) *Handbook of perception and human performance*. Volume II, cognitive processes and performance. Wiley, New York, pp 42/1–42/49
- DT (determination) test [www.schuhfried.at](http://www.schuhfried.at). Accessed 2008
- Wickens CD (1991) Processing resources and attention. In: Damos DL (ed) *Multiple-task performance*. Taylor & Francis, London, pp 3–34
- Mulder G, Mulder LJM (1980) Coping with mental workload. In: Levine S, Ursin H (eds) *Coping and health*. Plenum Press, New York, pp 233–258
- Leung AW, Chan CC, Ng JJ, Wong PC (2006) Factors contributing to officers' fatigue in high-speed maritime craft operations. *Appl Ergon* 37:565–576
- Wickens CD (1992) *Engineering psychology and human performance*. Harper Collins, New York
- Nater UM, La Marca R, Florin L, Moses A, Langhans W, Koller MM, Ehlert U (2006) Stress-induced changes in human salivary alpha-amylase activity—associations with adrenergic activity. *Psychoneuroendocrinology* 31:49–58
- Holder G (2006) Measurement of glucocorticoids in biological fluids. *Methods Mol Biol* 324:141–157
- U.S. Department of Defense (1999) *Human engineering program process and procedures (MIL-HDBK-46855A)*
- Brickenkamp R, Yillmer E (1998) *The d2 test of attention*, 1st U.S. edn. Hogrefe & Huber, Seattle
- Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) *Human mental workload*. North-Holland, Amsterdam, pp 139–183
- Weston-Lovelock K, Abram A (1996) Establishing the reliability of behaviorally anchored rating scales at the RCB (report no. PLSD/CHS/HS3/CR96055/1.0). DERA, Farnborough
- Fitts PM (1966) Cognitive aspects of information processing: III. Set for speed versus accuracy. *J Exp Psychol* 71:849–857
- Damos DL (1996) *Multiple-tasks performance*. Taylor and Francis Ltd, London, pp 328–360
- Kirschbaum C, Hellhammer DH (1994) Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology* 19:313–333
- Rohleder N, Wolf JM, Malsonado EF, Kirschbaum C (2006) The psychosocial stress-induced increase in salivary alpha-amylase is independent of saliva flow rate. *Psychophysiology* 43:645–652
- Kantowitz BH (1987) Mental workload. In: Hancock PA (ed) *Human factors psychology*. North-Holland, Amsterdam, pp 81–112
- Cook NJ (1992) Salivary cortisol and testosterone as markers of stress in normal subjects in abnormal situations. In: Kirschbaum C (ed) *Assessment of hormones and drugs in saliva in biobehavioral research*. Hogrefe & Huber, Seattle
- McCarty R (1998) The impact of a new emotional self-management program on stress, emotions, heart rate variability, DHEA and cortisol. *Integr Physiol Behav Sci* 33:151–170
- Gartner WB, Murphy MR (1979) Concepts of workload. In: Hartman BO, McKenry RE (eds) *Survey of methods to assess workload*; AGARDograph no. 246. AGARD, Neuilly-sur-Seine
- Hankins TC, Wilson GF (1998) A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviat Space Environ Med* 69:360–367
- Egan JP (1975) *Signal detection theory and ROC analysis*. Academic Press, New York
- Meister D (1976) *Behavioral foundation of system development*. Wiley, New York