

# Clinical Work Intensity Among Physician Specialties: How Might We Assess It?

## What Do We Find?

Ronnie D. Horner, PhD,\*† Jerzy P. Szaflarski, MD, PhD, FAAN,†‡ C. Jeffrey Jacobson, PhD,§¶ Nancy Elder, MD, MSPH,¶ Shannon Bolon, MD, MPH,¶ Gerald Matthews, PhD,‡ Jun Ying, PhD,\* Karthikeyan Meganathan, MS,\* and Marc Raphaelson, MD||

**Background:** The level of work intensity associated with patient encounters has implications for quality of care, patient safety, practice management, and reimbursement. The utility of available instruments for clinical work intensity assessment is unknown.

**Objective:** We assessed, in the clinical setting, the performance of existing measures of work intensity that are valid for nonclinical contexts.

**Research Design:** A cross-sectional, multimeasure design involving work intensity assessments for the last patient encounter and for an entire half-day clinic session.

**Subjects:** A convenience sample of 14 providers from the following 4 specialties: family medicine, general internal medicine, neurology, and surgery.

**Measures:** Perceived clinical work intensity was measured by the following 3 instruments: National Aeronautic and Space Administration-Task Load Index, Subjective Workload Assessment Technique, and Multiple Resources Questionnaire; stress was measured by the Dundee Stress State Questionnaire. Convergent validity was assessed by correlation among the instruments.

**Results:** For the last patient encounter, there was a moderate to high correlation between the work intensity instruments' scores (Pearson's  $r$  ranged from 0.41 to 0.73) and low to moderate correlation with the distress subscale of the Dundee Stress State Questionnaire (Pearson's  $r$  ranged from -0.11 to 0.46), reflecting their stress dimension. Provider personality was associated with reported levels of work intensity and stress. Similar results were obtained when the entire clinic session was the unit of reference.

**Conclusion:** Existing measures of work intensity and stress appear to be valid for use in the clinical setting to generate evidence on

perceived intensity and stress experienced by providers in the performance of medical services.

**Key Words:** medical specialties, personal health services, workload, stress, NASA-TLX, SWAT, MRQ, DSSQ

(*Med Care* 2011;49: 108–113)

Performance of a medical service (eg, evaluation and management [E&M] or a procedure) involves a number of factors, including technical skill and physical effort, clinical judgment or mental effort, stress, and time or temporal demand.<sup>1</sup> These factors have implications for health care delivery, including quality of care, patient safety, practice organization and management (eg, clinical workloads, staffing), and level of reimbursement. This is especially important now as the nation enters a new era in healthcare delivery with the passage of the Patient Protection and Affordable Care Act in March 2010. For example, the demand for health care, and especially primary care, is anticipated to increase substantially. However, the supply of primary care providers is unlikely to be sufficient to meet that demand.<sup>2</sup> Therefore, knowledge of the level of work intensity associated with various medical services will become central to designing strategies to increase the supply and productivity of primary care providers whether the strategies involve greater use of limited licensure providers within a practice, improved reimbursement for primary care services or other approaches.<sup>3</sup> Despite the importance of work intensity for health care delivery, there remains uncertainty as to how to measure it in the clinical setting.

Widely accepted measures of perceived work intensity exist but, with one exception, all were developed outside the clinical context.<sup>4–22</sup> Although an increasing number of reports describe the use of these instruments to assess clinical work intensity, limitations include a focus on a single measurement instrument, specialty, or type of medical service (eg, E&M or procedural services).<sup>23,24</sup> As an initial step in a larger study to address these limitations and gaps in knowledge, we assessed among 4 medical specialties the relevance (face and content validity), feasibility of use, and perfor-

From the Departments of \*Public Health Sciences, †Neurology, ‡Psychology, §Anthropology, and ¶Family and Community Medicine, University of Cincinnati, Cincinnati, OH; and ||The American Academy of Neurology, Chicago, IL.

Supported by the American Academy Neurology (AAN) and the American Academy of Dermatology (AAD).

The views expressed are those of the authors and do not necessarily reflect those held by the AAN or AAD.

Reprints: Ronnie D. Horner, PhD, Department of Public Health Sciences, University of Cincinnati, Stetson Building—Suite 4200, 260 Stetson Place, Cincinnati, OH 45219. E-mail: Ronnie.Horner@uc.edu.

Copyright © 2010 by Lippincott Williams & Wilkins  
ISSN: 0025-7079/11/4901-0108

mance of a set of widely used measures of work intensity along with a measure of stress in office-based and surgical (operating room) settings. In this brief report, we present results on instrument performance (convergent or concurrent validity). Qualitative assessments of face and content validity are presented in a companion report.<sup>25</sup>

## METHODS

### Study Design

In this cross-sectional study, participating providers were interviewed regarding work intensity and, if agreeable, were observed during a typical clinic session at which they completed a set of work intensity measures and a measure of stress with regard to the entire half-day clinic and the last patient encounter of the clinic. Only E&M services were evaluated in the office setting. The surgeons could elect to be observed during a session in the operating room wherein the measures were completed with regard to the surgical procedure performed; if more than one operation was scheduled during the session, the last operation was used. The assessments were completed promptly after clinic or procedure to avoid any effects of delayed reporting.<sup>26</sup> The number of patients seen at the half-day clinics varied among the providers, ranging from 3 to 19 patients with a median of 7 patients. In the operating room setting, the number of operations performed was 1 or 2. The study was approved by the institutional review board at the University of Cincinnati.

### Study Population

A convenience sample of 5 providers in family medicine, general internal medicine, and neurology and 4 in

surgery was assembled; the surgeons represented general surgery, neurosurgery, and orthopedic surgery. Of the 19 providers, 14 agreed to be observed during a typical clinic or in the operating room while interacting with their patients, including 5 family practitioners, 2 general internists, 3 neurologists, and 4 surgeons.

### Measures

The outcome measures included 3 instruments of perceived work intensity: NASA-Task Load Index (NASA-TLX), Subjective Workload Assessment Technique (SWAT), and Multiple Resources Questionnaire (MRQ), and one measure of stress: Dundee Stress State Questionnaire (DSSQ). Criteria for inclusion of an instrument were proven validity and reliability, wide acceptance as a standard measure, and brevity. Table 1 summarizes the characteristics of these instruments and typical contexts in which each has been used. Reports on the use of these instruments in the clinical setting with actual patients are limited to the NASA-TLX.<sup>23,24</sup>

### Covariates

Because these instruments measure perceived level of clinical work intensity or stress, data were collected on several factors that may modify the provider's perceptions including age, gender, race/ethnicity, years in practice, and personality, with the latter assessed by the Mini-International Personality Item Pool (mini-IPIP).<sup>28,29</sup>

### Data Analysis

This analysis focused on the convergent validity of the instruments with exploration of the role of the covariates as potential modifiers of self-reported levels of work intensity

**TABLE 1.** Characteristics of the Work Intensity and Stress Instruments

Instrument	No. Items	Time to Complete (min)	Dimensions Assessed	Psychometric Properties (See Reference)	Examples of Contexts in Which Used
NASA-TLX	6	2	Mental demands Physical demands Temporal demands Effort Performance Frustration	4, 27	Sustained attention <sup>6</sup> Simulated power plant operation <sup>7</sup> Physicians in hospital setting <sup>8</sup>
SWAT	3	1	Time load Mental effort load Stress load	9, 27	Mental arithmetic <sup>11</sup> Simulated military aviation <sup>12</sup> Vehicle driving <sup>13</sup>
MRQ	17	4	Processing Demands for: Auditory Facial Manual Memory Spatial Tactile Visual Vocal	14, 15	Dual-task performance <sup>16</sup> Simulated endoscopic surgery <sup>17</sup> Complex computer games <sup>16</sup>
DSSQ: full, short	90, 30	5	Task engagement Distress Worry	18	Tactical decision-making <sup>19</sup> Simulated vehicle driving (Matthews, personal communication) Simulated robot operation <sup>20</sup>

NASA-TLX indicates NASA-Task Load Index; SWAT, Subjective Workload Assessment Technique; MRQ, Multiple Resources Questionnaire; DSSQ, Dundee Stress State Questionnaire.

and stress. Of the surgeons, 2 completed the work intensity instruments with regard to the office setting only and 2 with regard to the operating room setting only. Parallel analyses were performed for data relating to the entire half-day clinic and those relating to the last patient encounter (or the surgical operation). Relationships between the work intensity instruments were assessed using Pearson's correlation coefficients. The association of the covariates to the work intensity scores was assessed using either Pearson's correlation coefficients or 2-sample *t* tests. Because of the small sample size, cross-sectional comparisons of work intensity scores and covariates among specialties were not performed by statistical tests; rather only descriptive statistics, that is, mean  $\pm$  standard deviation, median (range), and frequency, were used to summarize numerical and categorical variables.

Scores were calculated according to the scoring protocol of the particular scale. The overall scores from NASA-TLX, SWAT, and MRQ were transformed to the 0–100 scale for convenience of comparison, with a higher score indicating a higher level of perceived clinical work intensity.<sup>30</sup> The subscale scores from the short version of the DSSQ were transformed to the 0–100 scale with higher scores indicating a stronger experience of the particular stress state dimension. Scores on the mini-IPIP were also transformed to the 0–100 scale with higher scores indicating a stronger presence of the trait. Because the personality trait scores are interpretable only in comparison to a referent population, a referent mean score was obtained from an existing report on a large general population sample that had completed the mini-IPIP.<sup>28</sup>

## RESULTS

The typical participant was a 45-year-old white man with 10 years of practice (Table 2). In comparison to the other specialties, the participating neurologists included only women; the neurologists were also younger in age and had fewer years of practice. In relation to a general population (ie, the referent mean for each personality trait), our sample of providers tended to score more highly on agreeableness and

conscientiousness, and lower on neuroticism. Among the specialties, neurologists and general internists had notably lower scores on extroversion while neurologists had the lowest score on conscientiousness but the highest score on neuroticism.

For the last patient encounter of a clinic, there was moderate to high correlation between the scores from the various work intensity instruments with high correlations between the NASA-TLX and the SWAT, and between the SWAT and the MRQ (Table 3, upper half). The NASA-TLX had a moderate correlation with the Distress subscale of the DSSQ but low correlation with the other subscales. The SWAT, however, had low correlations with all of the DSSQ subscales. The MRQ was associated at a moderate level with both the Task Engagement and Worry subscales of the DSSQ. Findings were similar when the entire half-day clinic was the reference point (Table 3, lower half), although the correlations between the DSSQ subscale scores and the work intensity scores for the various measures tended to be stronger in the half-day context vis-à-vis the last patient encounter.

The level of perceived work intensity reported for the last patient encounter of a clinic was similar across the nonsurgical specialties and similar to those reported by the surgeons for the operating room context (Table 4, upper half). Compared with the other specialties, the surgeons' had a relatively high mean score on the SWAT and the MRQ but the lowest mean score for the NASA-TLX; the surgeons also had the highest mean score on the Task Engagement subscale of the DSSQ. When the entire half-day clinic is the context (Table 4, lower half), the surgeons had the lowest or among the lowest mean scores for perceived work intensity and stress.

The provider's age, race, gender, and number of years in practice were largely unrelated to the reported level of perceived work intensity whether for the last patient encounter or the entire clinic (data not shown). However, in general, these findings are most likely underpowered because of the small sample size of this pilot study.

**TABLE 2.** Characteristics of the Participants, by Medical Specialty (n = 14)

Characteristic	Medical Specialty				
	All (n = 14)	Family Medicine (n = 5)	General Internal Medicine (n = 2)	Neurology (n = 3)	Surgery (n = 4)
Male (n)	9	3	2	0	4
White (n)	11	3	2	3	3
Median age in years (range)	45 (33–66)	48 (33–54)	51.5 (49–54)	38 (37–40)	48.5 (39–66)
Median years in practice (range)	10 (2–27)	12 (2–23)	20 (18–22)	6 (4–8)	16.5 (4–27)
Mean mini-IPIP subscale scores (SD) and referent score					
Extroversion (Referent: 57.0)	51.4 ( $\pm$ 22.3)	62.5 ( $\pm$ 21.2)	28.1 ( $\pm$ 4.4)	28.1 ( $\pm$ 4.4)	60.9 ( $\pm$ 18.0)
Agreeableness (Referent: 75.3)	84.1 ( $\pm$ 8.3)	81.3 ( $\pm$ 10.8)	84.4 ( $\pm$ 4.4)	90.6 ( $\pm$ 13.3)	84.4 ( $\pm$ 3.6)
Conscientiousness (Referent: 75.3)	75.0 ( $\pm$ 14.2)	76.3 ( $\pm$ 12.0)	87.5 ( $\pm$ 0)	59.4 ( $\pm$ 4.4)	75.0 ( $\pm$ 18.4)
Neuroticism (Referent: 75.3)	29.8 ( $\pm$ 22.1)	27.5 ( $\pm$ 20.5)	12.5 ( $\pm$ 8.8)	65.6 ( $\pm$ 13.3)	23.4 ( $\pm$ 13.9)
Openness* (Referent: 75.3)	73.1 ( $\pm$ 12.3)	76.3 ( $\pm$ 12.0)	65.6 ( $\pm$ 4.4)	71.9 ( $\pm$ 13.3)	73.4 ( $\pm$ 17.2)

\*The Mini-IPIP subscale Openness is also referred to as openness to experience or as intellect/imagination.  
IPIP indicates International Personality Item Pool.

**TABLE 3.** Pearson Correlation Coefficient Between Measures of Work Intensity/Stress for the Last Patient Encounter of a Clinic or a Surgical Operation and an Entire Half-Day Clinic (n = 14\*)

Measure of Work Intensity or Stress	Measure of Work Intensity or Stress		
	NASA-TLX	SWAT	MRQ
Last patient encounter of the clinic (n = 12)			
SWAT	0.71 <sup>†</sup>		
MRQ	0.41	0.73 <sup>†</sup>	
Dundee stress state <sup>‡</sup>			
Engagement	0.14	0.33	0.57
Distress	0.46	0.31	−0.11
Worry	−0.07	0.21	0.44
Entire half-day clinic (n = 12)			
SWAT	0.95 <sup>†</sup>		
MRQ	0.35	0.51	
Dundee stress state <sup>§</sup>			
Engagement	0.54	0.51	0.40
Distress	0.58	0.61 <sup>†</sup>	0.09
Worry	0.35	0.59	0.77 <sup>†</sup>

\*Two surgeons completed instruments for the office context while 2 other surgeons completed instruments for the operating room context.

<sup>†</sup>Correlation coefficient statistically significant at  $P < 0.05$  or lower.

<sup>‡</sup>Short version with unadjusted scores.

<sup>§</sup>Full version with normative-adjusted scores.

NASA-TLX indicates NASA-Task Load Index; SWAT, Subjective Workload Assessment Technique; MRQ, Multiple Resources Questionnaire.

Personality, on the other hand, was found to have some association with the measures of perceived work intensity and stress. The neuroticism score related positively to the score on the DSSQ subscale of Distress ( $r = 0.63$ ), and also to the NASA-TLX workload score ( $r = 0.45$ ) in the half-day clinic context. The agreeableness score was negatively related to the MRQ and SWAT scores ( $r = -0.62$  and  $r = -0.56$ , respectively) for the last patient encounter and to the MRQ only ( $r = -0.52$ ) for entire half-day clinic context. It was negatively related to the Engagement and Worry subscale scores of the DSSQ, significantly so in the half-day clinic context ( $r = -0.48$  and  $r = -0.45$ , respectively). The conscientiousness score was positively related to MRQ score for the last patient encounter ( $r = 0.41$ ) and half-day clinic ( $r = 0.51$ ); it was negatively related to the Distress subscale score in these contexts ( $r = -0.69$  and  $r = -0.56$ , for last patient encounter and half-day clinic, respectively) but had a positive correlation with the Engagement and Worry subscales of the DSSQ.

## DISCUSSION

This is the first report on the performance of several work intensity instruments in clinical settings involving actual patient encounters for multiple medical specialties. The 3 instruments of work intensity: NASA-TLX, SWAT, and MRQ, yielded similar scores. The scores also had moderate correlation with the Distress subscale of the DSSQ as would be anticipated, given that the work intensity instruments

include a measure of stress as one of their dimensions. However, coping with stress may be an additional factor that heightens the perceived work intensity of providing health care. The consistently positive (though often statistically nonsignificant) correlations observed between work intensity and the DSSQ Task Engagement subscale suggest that mental demands may also have positive effects in maintaining the provider's energy, motivation, and concentration. In further work it will be important to more clearly elucidate the role of stress in perceived work intensity, and to evaluate the various dimensions of stress response for different medical specialists to provide a context for interpreting identified differences among specialties.

The work intensity scores reported for evaluation and management services among the 4 medical specialties were in the mid to upper range of possible scores, with average scores for a specific encounter ranging from 43.8 to 51.5. By comparison, a previous study of attending emergency medicine physicians in the emergency room setting reported an average NASA-TLX score of 50.6 for a 3-hour period, where activities varied from answering emergency service calls for which the average score was 26.0 to direct patient care activities for which the average score was 53.7.<sup>24</sup> Because we had only 2 surgeons reporting on work intensity in the operating room setting, we cannot draw conclusion regarding the perceived work intensity associated with the performance of surgical procedures.

Consistent with a previous study of providers,<sup>30</sup> we found that personality factors may relate to both perceived clinical work intensity and distress. The personality trait of neuroticism was positively correlated with perceived clinical work intensity across the measures and with the DSSQ subscale of Distress, although the correlations lacked statistical significance. Therefore, it may be prudent to control for the neuroticism personality trait when conducting interspecialty comparisons of work intensity. Moreover, as the neuroticism personality trait tends to be stronger in women,<sup>31</sup> differences in gender between specialties may also confound comparisons of clinical work intensity. The data further suggest a need to control for any interspecialty differences in conscientiousness and agreeableness, as both dimensions related to several work intensity measures and the measure of stress. Third, these data suggest that perceived work intensity and distress may be dissociated, although these 2 measures tend to be positively correlated. That is, highly conscientious doctors reported less distress, but also showed a trend toward higher work intensity on the MRQ. These individuals may be able to direct their mental effort so as to cope more effectively with the challenges of medical practice, alleviating the distress that may accompany high workload. Other potential modifiers of perceived work intensity and stress were not found to be important, including gender, race, and years in practice.

We note that work intensity scores were similar for a single patient encounter and for an entire half-day clinic. The work intensity instruments that we used were designed for assessing one specific task and are believed to be most accurate when completed within 15 minutes of task comple-



**TABLE 4.** Summary of Work Intensity/Stress, by Medical Specialty (n = 14)

Measure of Work Intensity or Stress	Medical Specialty			Surgery	
	Family Medicine (n = 5)	Internal Medicine (n = 2)	Neurology (n = 3)	Clinic (n = 2)	OR (n = 2)
Last patient encounter of the clinic (n = 12)					
NASA-TLX	51.5 (±14.2)	44.4 (±21.0)	43.8 (±19.8)	NA	40.7 (±13.1)
SWAT	80.0 (±16.5)	66.7 (±31.4)	74.1 (±25.7)	NA	83.3 (±7.9)
MRQ	58.0 (±21.9)	43.4 (±13.9)	39.5 (±14.5)	NA	59.4 (±0.0)
Dundee stress state <sup>†</sup>				NA	
Engagement	81.3 (±6.3)	85.9 (±15.5)	62.5 (±3.1)		92.2 (±6.6)
Distress	25.0 (±4.9)	17.2 (±2.2)	29.2 (±7.9)		25.0 (±8.8)
Worry	40.0 (±5.6)	10.9 (±2.2)	22.9 (±7.9)		39.1 (±28.7)
Entire half-day clinic (n = 12)					
NASA-TLX	51.7 (±18.4)	52.8 (±NA)*	63.6 (±3.5)	35.2 (±31.4)	NA
SWAT	73.3 (±20.2)	83.3 (±23.6)	81.5 (±6.4)	50.0 (±23.6)	NA
MRQ	65.2 (±24.1)	62.9 (±24.5)	48.7 (±4.6)	54.7 (±3.3)	NA
Dundee stress state <sup>‡</sup>					NA
Engagement	1.2 (±0.4)	1.5 (±0.1)	1.4 (±0.2)	1.4 (±NA)*	
Distress	−0.5 (±0.8)	−1.2 (±0.5)	−0.1 (±0.5)	−1.9 (±NA)*	
Worry	−0.4 (±0.9)	−0.3 (±1.1)	−0.7 (±0.3)	−0.9 (±NA)*	

\*n = 1.

<sup>†</sup>Short version with unadjusted scores.<sup>‡</sup>Full version with normative-adjusted scores.

NASA-TLX indicates NASA-Task Load Index; SWAT, Subjective Workload Assessment Technique; MRQ, Multiple Resources Questionnaire.

tion.<sup>26</sup> Although using these instruments to rate average intensity of visits during a 4-hour clinic may be valid, the perceived intensity of the final visit may affect the physician's rating of prior visits or the experience in earlier visits may affect the rating of the final visit. Context is known to affect reporting of the level of work intensity in other circumstances.<sup>26,32</sup> Other limitations of this study include the small number of respondents and the use of convenience sampling. We also did not control for patient factors that might influence the intensity experienced by the provider in the performance of the evaluation and management activities or the surgical procedures. One might reasonably anticipate that the level of intensity for a well-patient visit would differ from that for the diagnosis of a new disease or condition, although both represent an evaluation and management service.

We further caution interpretation of our findings by noting that the instruments measure self-rated work intensity and stress. We did not assess work intensity and stress through more objective measures. However, although objective physiologic measures such as heart rate, pupil dilatation, blood pressure, or cortisol level can indicate stress reactions, they correlate poorly with measures of work intensity (p. 990).<sup>33</sup> These parameters are affected by physical activity as well as mental work, and may be difficult to measure in the clinical environment. Although intensity and stress are related and confounding factors in health care service, future research should define, evaluate, and measure them separately.

Despite their preliminary nature, we believe these findings support the use of these existing work intensity instru-

ments to measure perceived clinical work intensity. If confirmed in larger studies, these instruments will provide a means of generating comparable information regarding the level of work intensity and stress associated with the performance of various medical services. In turn, such information could serve as a rational basis by which to improve health care delivery, whether the goal is greater efficiency in practice organization and management, higher quality of care, greater patient safety, or creating physician incentives proportional to work.

## REFERENCES

1. Hsiao W, Braun P, Yntema D, et al. Estimating physician's work for a resource-based relative value scale. *N Engl J Med*. 1988;319:835–841.
2. Doherty RB. The certitudes and uncertainties of health care reform. *Ann Intern Med*. 2010;152:679–682.
3. Kane GC, Grever MR, Kennedy JJ, et al. The anticipated physician shortage: meeting the nation's need for physician services. *Am J Med*. 2009;122:1156–1162.
4. Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, eds. *Human Mental Workload*. Oxford, United Kingdom: North-Holland; 1988:139–183.
5. O'Donnell RD, Eggemeier FT. Workload assessment methodology. In: Boff KR, Kaufmann L, Thomas J, eds. *Handbook of Perception and Human Performance*. New York, NY: Wiley; 1986:2:42/1–42/49.
6. Temple JG, Warm JS, Dember WN, et al. The effects of signal salience and caffeine on performance, workload and stress in an abbreviated vigilance task. *Hum Factors*. 2000;42:183–194.
7. Hwang S, Yau Y, Lin Y, et al. Predicting work performance in nuclear power plants. *Saf Sci*. 2008;46:1115–1124.
8. Zwaan L, Thijs A, Wagner C, et al. Design of a study on suboptimal cognitive acts in the diagnostic process, the effect on patient outcomes and the influence of workload, fatigue and experience of physician. *BMC Health Serv Res*. 2009;9:65.

9. Reid GB, Nygren TE. The subjective workload assessment technique: a scaling procedure for measuring mental workload. In: Hancock PS, Meshkati N, eds. *Human Mental Workload*. Amsterdam, the Netherlands: Elsevier; 1988.
10. Tsang PS, Vidulich MA. Mental workload and situation awareness. In: Salvendy G, ed. *Handbook of Human Factors and Ergonomics*. 3rd ed. Hoboken, NJ: Wiley; 2006;417–449.
11. Luximon A, Goonetilleke S. Simplified subjective workload assessment technique. *Ergonomics*. 2001;44:229–243.
12. Colle HA, Reid B. Estimating a mental workload redline in a simulated air-to-ground combat mission. *Int J Aviat Psychol*. 2005;15:303–319.
13. Verwey WB, Veltman A. Detecting short periods of elevated workload: a comparison of nine workload assessment techniques. *J Exp Psychol*. 1996;2:270–285.
14. Boles DB, Adair LP. The Multiple Resources Questionnaire (MRQ). *Proc Hum Factors Ergon Soc*. 2001;45:1790–1794.
15. Boles DB, Adair LP. Validity of the multiple resources questionnaire (MRQ). *Proc Hum Factors Ergon Soc*. 2001;45:1795–1799.
16. Boles DB, Bursk JH, Phillips JB, et al. Predicting dual-task performance with the Multiple Resources Questionnaire (MRQ). *Hum Factors*. 2007;49:32–45.
17. Klein MI, Warm JS, Riley MA, et al. Perceptual distortions produce multidimensional stress profiles in novice users of an endoscopic surgery simulator. *Hum Factors*. 2008;50:291–300.
18. Matthews G, Campbell SE, Falconer S, et al. Fundamental dimensions of subjective state in performance settings: task engagement, distress and worry. *Emotion*. 2002;2:315–340.
19. Hudlicka E, Matthews G. Affect, risk and uncertainty in decision-making: an integrated computational-empirical approach. Air Force Office of Scientific Research Technical Report, Number A291505; 2009.
20. Guznov S, Matthews G, Warm JS. Team member personality, performance and stress in a Roboflag synthetic task environment. *Proc Hum Factors Ergon Soc*. In press.
21. Bertram DA, Hershey CO, Opila DA, et al. A measure of physician mental work load in internal medicine ambulatory care clinics. *Med Care*. 1990;28:458–467.
22. Bertram DA, Opila DA, Brown JL, et al. Measuring physician mental workload: reliability and validity assessment of a brief instrument. *Med Care*. 1992;30:95–104.
23. Tawil AZ, Thomas G, Lee JD, et al. What's so hard about bronchoscopic surgery? Identifying the perceptual and cognitive difficulties of bronchoscopic surgery. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society; 1999, pp 845–849.
24. Levin S, France DJ, Hemphill R, et al. Tracking workload in the emergency department. *Hum Factors*. 2006;48:526–539.
25. Jacobson CJ, Bolon S, Elder N, et al. Temporal and subjective work demands in office based patient care: an exploration of the dimensions of physician work intensity. *Med Care*. 2011;49:52–58.
26. Moroney WF, Biers DW, Eggemeier FT. Some measurement and methodological considerations in the application of subjective workload measurement techniques. *Int J Aviat Psychol*. 1995;5:87–106.
27. Rubio S, Diaz E, Martin J, et al. Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Appl Psychol*. 2004;53:61–86.
28. Donnellan MB, Oswald FL, Baird BM, et al. The Mini-IPIP Scales: tiny-yet-effective measures of the big five factors of personality. *Psychol Assess*. 2006;18:192–203.
29. Matthews G, Deary IJ, Whiteman MC. *Personality Traits*. 3rd ed. Cambridge, United Kingdom: Cambridge University Press; 2009.
30. Cosway R, Endler NS, Sadler AJ, et al. The coping inventory for stressful situations: factorial structure and associations with personality traits and psychological health. *J Appl Biobehav Res*. 2000;5:121–143.
31. Costa P, Terracciano A, McCrae RR. Gender differences in personality traits across cultures: robust and surprising findings. *J Pers Soc Psychol*. 2001;81:322–331.
32. Colle HA, Reid GB. Context effects in subjective mental workload ratings. *Hum Factors*. 1998;40:591–600.
33. Cegarra J, Chevalier A. The use of Tholos software for combining measures of mental workload: toward theoretical and methodological improvement. *Behav Res Methods*. 2008;40:988–1000.