

Human Factors – rok 2017, roč. 59

Číslo 4 (June)



SPECIAL SECTION: 2016 HUMAN FACTORS PRIZE FOR EXCELLENCE IN HUMAN FACTORS/ERGONOMICS RESEARCH: BIG DATA

HUMAN FACTORS PRIZE WINNER

Andrew J. Hampton, Valerie L. Shalin. *Sentinels of Breach: Lexical Choice as a Measure of Urgency in Social Media*. pp. 505–519.

Objective: This paper identifies general properties of language style in social media to help identify areas of need in disasters. **Background:** In the search for metrics of need in social media data, much of the existing literature ignores processes of language usage. Psychological concepts, such as narrative breach, Gricean maxims, and lexical marking in cognition, may assist the recovery of disaster-relevant metrics from altered patterns of word prevalence. **Method:** We analyzed several hundred thousand location-specific microblogs from Twitter for Hurricane Sandy, Oklahoma tornadoes, and the Boston Marathon bombing along with a fantasy football control corpus, examining the relative frequency of words in 36 antonym pairs. We compared the ratio of words within these pairs to the corresponding ratios recovered from an online word norm database. **Results:** Partial rank correlation values between observed antonym ratios demonstrate consistent patterns across disasters. For Hurricane Sandy data, 25 antonym pairs have moderate to large effect sizes for discrepancies between observed and normative ratios. Across disasters, 7 pairs are stable and meet effect size criteria. Sentiment analysis, supplementary word frequency counts with respect to disaster proximity, and examples support a “breach” account for the observed results. **Conclusion:** Lexical choice between antonyms, only somewhat related to sentiment, suggests that social media capture wide-ranging breaches of normal functioning. **Application:** Antonym selection contributes to screening tools based on language style for identifying relevant content and quantifying disruption using social media without the *a priori* specification of content keywords.

HUMAN FACTORS PRIZE FINALIST

William J. Irwin, Saul D. Robinson, Stephen M. Belt. *Visualization of Large-Scale Narrative Data Describing Human Error*. pp. 520–534.

Objective: Introduced is a visual data exploration technique for compiling, reducing, organizing, visually rendering, and filtering text-based narratives for detailed analysis. **Background:** The analysis of data sets provides an increasingly difficult problem. The method of visual representation is considered an effective tool in many applications. The focus of this study was to determine if a latent semantic analysis-based projection of narrative data into a geographic information systems software program provided a useful tool for reducing and organizing large sums of narrative data for analysis. **Method:** This approach utilizes latent semantic analysis to reduce narratives to a high-dimensional vector, truncates the vector to a two-dimensional projection through application of isometric mapping, and then visually renders the result with geographic information systems software. This method is demonstrated on aviation self-reported safety narratives sourced from the Aviation Safety Reporting System. **Results:** Thematic regions from the corpus are illustrated along with the first five topics identified. **Conclusion:** Shown is the ability to assimilate a large number of narratives, identify contextual themes, recognize common events and outliers, and organize resultant topics. **Application:** Large narrative-based data sets present in aviation and other domains may be visualized to facilitate efficient analysis, enhance comprehension, and improve safety.

BIOMECHANICS, ANTHROPOMETRY, WORK PHYSIOLOGY

Thomas Karakolis, Brittany A. Sinclair, Alison Kelly, Phil Terhaar, Linda L. M. Bossi. *Determination of Orientation and Practice Requirements When Using an Obstacle Course for Mobility Performance Assessment.* pp. 535–545.

Objective: Determine effect of orientation (introduction and familiarization) and practice (repeated performance) on human performance under various load conditions as assessed by an obstacle course. **Background:** Obstacle courses are commonly used as screening tools by military, police, and firefighters or to assess human capabilities and the effect of wearing personal protective equipment (PPE) and other occupationally necessary equipment on mobility task performance. Unfortunately, little is formally documented about the effect of orientation and practice on performance outcomes of obstacle or mobility courses being used. **Method:** Forty-eight participants were recruited from the Canadian Army Infantry and Combat Engineer population. Participants either received regular or extensive orientation of the course before completing it. Following orientation, participants completed the course five consecutive times while wearing their PPE with full fighting order (FFO) and five consecutive times while wearing no PPE and non-FFO across a five-day period (maximum two runs per day), with ensemble presentation order counterbalanced. Total course completion time and individual obstacle completion times were measured for each run of the course. **Results:** While wearing FFO, participants continued to decrease the time required for completing the course; however, while wearing non-FFO, time to course completion did not significantly change over the five runs. There were no differences in course completion times for the regular and extensive course orientation groups. **Conclusions:** Considerations required to mitigate orientation and practicing effects can differ depending on type or complexity of load condition. While wearing FFO, practicing effects can introduce undesired confounding factors into data collection. **Application:** Any practice runs on an obstacle course prior to its use as an assessment tool should focus on the loaded (e.g., FFO) condition because improvement on loaded runs is likely transferred to unloaded, but this does not apply in the reverse.

Amanda M. Farias Zuniga, Julie N. Côté. *Effects of Dual Monitor Computer Work Versus Laptop Work on Cervical Muscular and Proprioceptive Characteristics of Males and Females.* pp. 546–563.

Objective: The effects of performing a 90-minute computer task with a laptop versus a dual monitor desktop workstation were investigated in healthy young male and female adults. **Background:** Work-related musculoskeletal disorders are common among computer (especially female) users. Laptops have surpassed desktop computer sales, and working with multiple monitors has also become popular. However, few studies have provided objective evidence on how they affect the musculoskeletal system in both genders. **Methods:** Twenty-seven healthy participants (mean age = 24.6 years; 13 males) completed a 90-minute computer task while using a laptop or dual monitor (DualMon) desktop. Electromyography (EMG) from eight upper body muscles and visual strain were measured throughout the task. Neck proprioception was tested before and after the computer task using a head-repositioning test. EMG amplitude (root mean square [RMS]), variability (coefficients of variation [CV]), and normalized mutual information (NMI) were computed. **Results:** Visual strain ($p < .01$) and right upper trapezius RMS ($p = .03$) increased significantly over time regardless of workstation. Right cervical erector spinae RMS and cervical NMI were smaller, while degrees of overshoot (mean = 4.15°) and end position error (mean = 1.26°) were larger in DualMon regardless of time. Effects on muscle activity were more pronounced in males, whereas effects on proprioception were more pronounced in females. **Conclusion:** Results suggest that compared to laptop, DualMon work is effective in reducing cervical muscle activity, dissociating cervical connectivity, and maintaining more typical neck repositioning patterns, suggesting some health-protective effects. **Application:** This evidence could be considered when deciding on computer workstation designs.

COGNITION

Robert R. Hoffman, P. A. Hancock. *Measuring Resilience*. pp. 564–581.

Objective: As human factors and ergonomics (HF/E) moves to embrace a greater systems perspective concerning human-machine technologies, new and emergent properties, such as resilience, have arisen. Our objective here is to promote discussion as to how to measure this latter, complex phenomenon. **Background:** Resilience is now a much-referenced goal for technology and work system design. It subsumes the new movement of resilience engineering. As part of a broader systems approach to HF/E, this concept requires both a definitive specification and an associated measurement methodology. Such an effort epitomizes our present work. **Method:** Using rational analytic and synthetic methods, we offer an approach to the measurement of resilience capacity. **Results:** We explicate how our proposed approach can be employed to compare resilience across multiple systems and domains, and emphasize avenues for its future development and validation. **Conclusion:** Emerging concerns for the promise and potential of resilience and associated concepts, such as adaptability, are highlighted. Arguments skeptical of these emerging dimensions must be met with quantitative answers; we advance one approach here. **Application:** Robust and validated measures of resilience will enable coherent and rational discussions of complex emergent properties in macrocognitive system science.

HEALTH CARE/HEALTH SYSTEMS

Rocio Garcia-Retamero, Edward T. Cokely. *Designing Visual Aids That Promote Risk Literacy: A Systematic Review of Health Research and Evidence-Based Design Heuristics*. pp. 582–627.

Background: Effective risk communication is essential for informed decision making. Unfortunately, many people struggle to understand typical risk communications because they lack essential decision-making skills. **Objective:** The aim of this study was to review the literature on the effect of numeracy on risk literacy, decision making, and health outcomes, and to evaluate the benefits of visual aids in risk communication. **Method:** We present a conceptual framework describing the influence of numeracy on

risk literacy, decision making, and health outcomes, followed by a systematic review of the benefits of visual aids in risk communication for people with different levels of numeracy and graph literacy. The systematic review covers scientific research published between January 1995 and April 2016, drawn from the following databases: Web of Science, PubMed, PsycINFO, ERIC, Medline, and Google Scholar. Inclusion criteria were investigation of the effect of numeracy and/or graph literacy, and investigation of the effect of visual aids or comparison of their effect with that of numerical information. Thirty-six publications met the criteria, providing data on 27,885 diverse participants from 60 countries. **Results:** Transparent visual aids robustly improved risk understanding in diverse individuals by encouraging thorough deliberation, enhancing cognitive self-assessment, and reducing conceptual biases in memory. Improvements in risk understanding consistently produced beneficial changes in attitudes, behavioral intentions, trust, and healthy behaviors. Visual aids were found to be particularly beneficial for vulnerable and less skilled individuals. **Conclusion:** Well-designed visual aids tend to be highly effective tools for improving informed decision making among diverse decision makers. We identify five categories of practical, evidence-based guidelines for heuristic evaluation and design of effective visual aids.

Farzan Sasangohar, Birsen Donmez, Anthony C. Easty, Patricia L. Trbovich. *Effects of Nested Interruptions on Task Resumption: A Laboratory Study With Intensive Care Nurses.* pp. 628–639.

Objective: Interruptions to secondary tasks resulting in multiple tasks to resume may tax working memory. The objective of this research is to study such interruptions experienced by intensive care unit (ICU) nurses. **Background:** ICU nurses are frequently interrupted, resulting in a switch from primary to secondary tasks. In two recent studies, we observed that some of these secondary tasks also get interrupted, resulting in multiple tasks that have to be resumed, a phenomenon we refer to as nested interruptions. Although completing multiple secondary tasks in a serial fashion during an interruption period can create context-switching costs, we hypothesize that nested interruptions tax the working memory even more than just performing multiple secondary tasks sequentially because the nurse would have to encode in working memory the resumption goals for both the primary and the interrupted secondary tasks. **Method:** We conducted a laboratory study with 30 ICU nurses, who performed an electronic order-entry task under three interruption conditions: (a) baseline—no secondary task during the interruption period; (2) serial—performance of two tasks one after the other during the interruption period; and (3) nested—performance of two tasks during the interruption period, one of which was also interrupted. **Results:** Nested interruptions resulted in significantly longer primary-task resumption lag and less accurate task resumption compared with both the serial interruption and baseline conditions. **Conclusion:** The nested nature of interruptions adds to the resumption lag and diminishes resumption accuracy by likely populating the working memory with goals associated with interrupted secondary tasks.

HUMAN-COMPUTER INTERACTION, COMPUTER SYSTEMS

Aiping Xiong, Robert W. Proctor, Weining Yang, Ninghui Li. *Is Domain Highlighting Actually Helpful in Identifying Phishing Web Pages?* pp. 640–660.

Objective: To evaluate the effectiveness of domain highlighting in helping users identify whether Web pages are legitimate or spurious. **Background:** As a component of the URL, a domain name can be overlooked. Consequently, browsers highlight the domain name to help users identify which Web site they are visiting. Nevertheless, few studies have assessed the effectiveness of domain highlighting, and the only formal study confounded highlighting with instructions to look at the address bar. **Method:** We

conducted two phishing detection experiments. Experiment 1 was run online: Participants judged the legitimacy of Web pages in two phases. In Phase 1, participants were to judge the legitimacy based on any information on the Web page, whereas in Phase 2, they were to focus on the address bar. Whether the domain was highlighted was also varied. Experiment 2 was conducted similarly but with participants in a laboratory setting, which allowed tracking of fixations. **Results:** Participants differentiated the legitimate and fraudulent Web pages better than chance. There was some benefit of attending to the address bar, but domain highlighting did not provide effective protection against phishing attacks. Analysis of eye-gaze fixation measures was in agreement with the task performance, but heat-map results revealed that participants' visual attention was attracted by the highlighted domains. **Conclusion:** Failure to detect many fraudulent Web pages even when the domain was highlighted implies that users lacked knowledge of Web page security cues or how to use those cues. **Application:** Potential applications include development of phishing prevention training incorporating domain highlighting with other methods to help users identify phishing Web pages.

PHYSIOLOGICAL AND PSYCHOLOGICAL CONDITIONS ("INTERNAL ENVIRONMENT")

Lucia Arsintescu, Jeffrey B. Mulligan, Erin E. Flynn-Evans. [Evaluation of a Psychomotor Vigilance Task for Touch Screen Devices](#). pp. 661–670.

Objective: Our goals were to compare three techniques for performing a psychomotor vigilance task (PVT) on a touch screen device (fifth-generation iPod) and to determine the device latency. **Background:** The PVT is a reaction-time test that is sensitive to sleep loss and circadian misalignment. Several PVT tests have been developed for touch screen devices, but unlike the standard PVT developed for laboratory use, these tests allow for touch responses to be recorded at any location on the device, with contact from any finger. In addition, touch screen devices exhibit latency in processing time between the touch response and the time registered by the device. **Method:** Thirteen participants completed a 5-min PVT on a touch screen device held in three positions (on a table with index finger, handheld portrait with index finger, handheld landscape with thumb). We compared reaction-time outcomes in each orientation condition using paired t tests. We recorded the first session using a high-speed video camera to determine the latency between the touch response and the documented response time. **Results:** The participants had significantly faster reaction times in the landscape-oriented position using the thumb, compared with the portrait-oriented position using the index ($M = 224.13$ and $M = 244.26$, $p = .045$). Using data from 1,241 unique touch events, we found a mean device latency of 68.53 ms that varied highly between individuals. **Conclusion:** Device orientation and device latency should be considered when using a touch screen version of a PVT. **Application:** Our findings apply to researchers administering touch screen versions of the PVT.

SURFACE TRANSPORTATION

Kathryn G. Tippet, Elayaraj Sivaraj, Thomas K. Ferris. *Driving While Interacting With Google Glass: Investigating the Combined Effect of Head-Up Display and Hands-Free Input on Driving Safety and Multitask Performance*. pp. 671–688.

Objective: This study evaluated the individual and combined effects of voice (vs. manual) input and head-up (vs. head-down) display in a driving and device interaction task. **Background:** Advances in wearable technology offer new possibilities for in-vehicle interaction but also present new challenges for managing driver attention and regulating device usage in vehicles. This research investigated how driving performance is affected by interface characteristics of devices used for concurrent secondary tasks. A positive

impact on driving performance was expected when devices included voice-to-text functionality (reducing demand for visual and manual resources) and a head-up display (HUD) (supporting greater visibility of the driving environment). **Method:** Driver behavior and performance was compared in a texting-while-driving task set during a driving simulation. The texting task was completed with and without voice-to-text using a smartphone and with voice-to-text using Google Glass's HUD. **Results:** Driving task performance degraded with the addition of the secondary texting task. However, voice-to-text input supported relatively better performance in both driving and texting tasks compared to using manual entry. HUD functionality further improved driving performance compared to conditions using a smartphone and often was not significantly worse than performance without the texting task. **Conclusion:** This study suggests that despite the performance costs of texting-while-driving, voice input methods improve performance over manual entry, and head-up displays may further extend those performance benefits. **Application:** This study can inform designers and potential users of wearable technologies as well as policymakers tasked with regulating the use of these technologies while driving.

Alexander Eriksson, Neville A. Stanton. *Takeover Time in Highly Automated Vehicles: Noncritical Transitions to and From Manual Control.* pp. 689–705.

Objective: The aim of this study was to review existing research into driver control transitions and to determine the time it takes drivers to resume control from a highly automated vehicle in noncritical scenarios. **Background:** Contemporary research has moved from an inclusive design approach to adhering only to mean/median values when designing control transitions in automated driving. Research into control transitions in highly automated driving has focused on urgent scenarios where drivers are given a relatively short time span to respond to a request to resume manual control. We found a paucity in research into more frequent scenarios for control transitions, such as planned exits from highway systems. **Method:** Twenty-six drivers drove two scenarios with an automated driving feature activated. Drivers were asked to read a newspaper, or to monitor the system, and to relinquish, or resume, control from the automation when prompted by vehicle systems. **Results:** Significantly longer control transition times were found between driving with and without secondary tasks. Control transition times were substantially longer than those reported in the peer-reviewed literature. **Conclusion:** We found that drivers take longer to resume control when under no time pressure compared with that reported in the literature. Moreover, we found that drivers occupied by a secondary task exhibit larger variance and slower responses to requests to resume control. Workload scores implied optimal workload. **Application:** Intra- and interindividual differences need to be accommodated by vehicle manufacturers and policy makers alike to ensure inclusive design of contemporary systems and safety during control transitions.