

Human Factors – rok 2016, roč. 58

Číslo 3 (May)



AUTOMATION, EXPERT SYSTEMS

Kristin E. Schaefer, Jessie Y. C. Chen, James L. Szalma, and P. A. Hancock. *A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems*. Pages 377-400.

Objective: We used meta-analysis to assess research concerning human trust in automation to understand the foundation upon which future autonomous systems can be built. **Background:** Trust is increasingly important in the growing need for synergistic human-machine teaming. Thus, we expand on our previous meta-analytic foundation in the field of human-robot interaction to include all of automation interaction. **Method:** We used meta-analysis to assess trust in automation. Thirty studies provided 164 pairwise effect sizes, and 16 studies provided 63 correlational effect sizes. **Results:** The overall effect size of all factors on trust development was $\bar{g} = +0.48$, and the correlational effect was $\bar{r} = +0.34$, each of which represented medium effects. Moderator effects were observed for the human-related ($\bar{g} = +0.49$; $\bar{r} = +0.16$) and automation-related ($\bar{g} = +0.53$; $\bar{r} = +0.41$) factors. Moderator effects specific to environmental factors proved insufficient in number to calculate at this time. **Conclusion:** Findings provide a quantitative representation of factors influencing the development of trust in automation as well as identify additional areas of needed empirical research. **Application:** This work has important implications to the enhancement of current and future human-automation interaction, especially in high-risk or extreme performance environments.

- **Keywords:** human-automation interaction, human-robot interaction, meta-analysis, trust

AVIATION AND AEROSPACE

Joseph E. Mercado, Michael A. Rupp, Jessie Y. C. Chen, Michael J. Barnes, Daniel Barber, and Katelyn Procci. *Intelligent Agent Transparency in Human-Agent Teaming for Multi-UxV Management*. Pages 401-415.

Objective: We investigated the effects of level of agent transparency on operator performance, trust, and workload in a context of human-agent teaming for multirobot management. **Background:** Participants played the role of a heterogeneous unmanned

vehicle (UxV) operator and were instructed to complete various missions by giving orders to UxVs through a computer interface. An intelligent agent (IA) assisted the participant by recommending two plans—a top recommendation and a secondary recommendation—for every mission. **Method:** A within-subjects design with three levels of agent transparency was employed in the present experiment. There were eight missions in each of three experimental blocks, grouped by level of transparency. During each experimental block, the IA was incorrect three out of eight times due to external information (e.g., commander's intent and intelligence). Operator performance, trust, workload, and usability data were collected. **Results:** Results indicate that operator performance, trust, and perceived usability increased as a function of transparency level. Subjective and objective workload data indicate that participants' workload did not increase as a function of transparency. Furthermore, response time did not increase as a function of transparency. **Conclusion:** Unlike previous research, which showed that increased transparency resulted in increased performance and trust calibration at the cost of greater workload and longer response time, our results support the benefits of transparency for performance effectiveness without additional costs. **Application:** The current results will facilitate the implementation of IAs in military settings and will provide useful data to the design of heterogeneous UxV teams.

- **Keywords:** intelligent agent transparency, human-agent teaming, multi-UxV management

DISPLAYS AND CONTROL

Michael A. Nees, Benji Helbein, and Anna Porter. *Speech Auditory Alerts Promote Memory for Alerted Events in a Video-Simulated Self-Driving Car Ride*. Pages 416-426.

Objective: Auditory displays could be essential to helping drivers maintain situation awareness in autonomous vehicles, but to date, few or no studies have examined the effectiveness of different types of auditory displays for this application scenario. **Background:** Recent advances in the development of autonomous vehicles (i.e., self-driving cars) have suggested that widespread automation of driving may be tenable in the near future. Drivers may be required to monitor the status of automation programs and vehicle conditions as they engage in secondary leisure or work tasks (entertainment, communication, etc.) in autonomous vehicles. **Method:** An experiment compared memory for alerted events—a component of Level 1 situation awareness—using speech alerts, auditory icons, and a visual control condition during a video-simulated self-driving car ride with a visual secondary task. The alerts gave information about the vehicle's operating status and the driving scenario. **Results:** Speech alerts resulted in better memory for alerted events. Both auditory display types resulted in less perceived effort devoted toward the study tasks but also greater perceived annoyance with the alerts. **Conclusion:** Speech auditory displays promoted Level 1 situation awareness during a simulation of a ride in a self-driving vehicle under routine conditions, but annoyance remains a concern with auditory displays. **Application:** Speech auditory displays showed promise as a means of increasing Level 1 situation awareness of routine scenarios during an autonomous vehicle ride with an unrelated secondary task.

- **Keywords:** auditory displays, autonomous vehicles, human-automation interaction, situation awareness, auditory icons, speech displays, dual-task performance, memory

HEALTH CARE/HEALTH SYSTEMS

David P. Azari, Carla M. Pugh, Shlomi Laufer, Calvin Kwan, Chia-Hsiung Chen, Thomas Y. Yen, Yu Hen Hu, and Robert G. Radwin. *Evaluation of*

Simulated Clinical Breast Exam Motion Patterns Using Marker-Less Video Tracking. Pages 427-440.

Objective: This study investigates using marker-less video tracking to evaluate hands-on clinical skills during simulated clinical breast examinations (CBEs). **Background:** There are currently no standardized and widely accepted CBE screening techniques. **Methods:** Experienced physicians attending a national conference conducted simulated CBEs presenting different pathologies with distinct tumorous lesions. Single hand exam motion was recorded and analyzed using marker-less video tracking. Four kinematic measures were developed to describe temporal (time pressing and time searching) and spatial (area covered and distance explored) patterns. **Results:** Mean differences between time pressing, area covered, and distance explored varied across the simulated lesions. Exams were objectively categorized as either sporadic, localized, thorough, or efficient for both temporal and spatial categories based on spatiotemporal characteristics. The majority of trials were temporally or spatially thorough (78% and 91%), exhibiting proportionally greater time pressing and time searching (temporally thorough) and greater area probed with greater distance explored (spatially thorough). More efficient exams exhibited proportionally more time pressing with less time searching (temporally efficient) and greater area probed with less distance explored (spatially efficient). Just two (5.9 %) of the trials exhibited both high temporal and spatial efficiency. **Conclusions:** Marker-less video tracking was used to discriminate different examination techniques and measure when an exam changes from general searching to specific probing. The majority of participants exhibited more thorough than efficient patterns. **Application:** Marker-less video kinematic tracking may be useful for quantifying clinical skills for training and assessment.

- **Keywords:** hands-on clinical examination, tactile inspection, medical simulation, examination technique

Antje C. Venjakob, Tim Marnitz, Peter Phillips, and Claudia R. Mello-Thoms. Image Size Influences Visual Search and Perception of Hemorrhages When Reading Cranial CT: An Eye-Tracking Study. Pages 441-451.

Objective: The aim of this study was to explore reader gaze, performance, and preference during interpretation of cranial computed tomography (cCT) in stack mode at two different sizes. **Background:** Digital display of medical images allows for the manipulation of many imaging factors, like image size, by the radiologists, yet it is often not known what display parameters better suit human perception. **Method:** Twenty-one radiologists provided informed consent to be eye tracked while reading 20 cCT cases. Half of these cases were presented at a size of 14 × 14 cm (512 × 512 pixels), half at 28 × 28 cm (1,024 × 1,024 pixels). Visual search, performance, and preference for the two image sizes were assessed. **Results:** When reading small images, significantly fewer, but longer, fixations were observed, and these fixations covered significantly more slices. Time to first fixation of true positive findings was faster in small images, but dwell time on true findings was longer. Readers made more false positive decisions in small images, but no overall difference in either jackknife alternative free-response receiver operating characteristic or reading time was found. **Conclusion:** Overall performance is not affected by image size. However, small-stack-mode cCT images may better support the use of motion perception and acquiring an overview, whereas large-stack-mode cCT images seem better suited for detailed analyses. **Application:** Subjective and eye-tracking data suggest that image size influences how images are searched and that different search strategies might be beneficial under different circumstances.

- **Keywords:** radiology and medical imaging, eye movements, tracking, visual search, computer interface

MOTOR BEHAVIOUR

Alexander Woodham, Mark Billingham, and William S. Helton. *Climbing With a Head-Mounted Display: Dual-Task Costs*. Pages 452-461.

Objective: We explored the dual-task costs of climbers performing a visual communication task using a head-mounted display (HMD) while simultaneously climbing along a vertical surface. **Background:** Climbing is affected by secondary auditory cognitive tasks, and climbing impairs later recall of secondary task information; the effects of visually presented tasks are less clear. Given that HMDs are projected to be adopted into emergency response work, questions are raised about the effects of HMD use during climbing or other physical tasks. **Method:** Climbers performed five conditions—a climbing-only condition, two dual-task climbing conditions (words presented on the HMD with and without auditory warnings while climbing), and two seated control conditions (words presented on the HMD with and without auditory warnings)—in a repeated-measures design. Motion data were also collected to examine participant motion around word presentation. **Results:** We found a decrease in both climbing performance and word recall under dual-task conditions, paralleling results found in previous research using auditory tasks. Participants slowed around word presentations on the HMD. Additional comparisons to previous research indicate that physical tasks may be more detrimental to word recall than are seated tasks and that visual stimuli might hinder climbing performance more than do audible stimuli. **Conclusion:** Complex physical activity, like climbing, is disruptive to memory rehearsal and later recall, and cognitive tasks disrupt physical performance. **Application:** Avoiding cognitive HMD tasks requiring later recall during complex physical activity is advisable. However, these systems may be developed to provide intelligent assistance, or memory augmentation, in these settings.

- **Keywords:** communication, exercise, head-mounted display, resource theory, physical tasks

Yusuke Yamani and Jason S. McCarley. *Workload Capacity: A Response Time-Based Measure of Automation Dependence*. Pages 462-471.

Objective An experiment used the workload capacity measure $C(t)$ to quantify the processing efficiency of human-automation teams and identify operators' automation usage strategies in a speeded decision task. **Background** Although response accuracy rates and related measures are often used to measure the influence of an automated decision aid on human performance, aids can also influence response speed. Mean response times (RTs), however, conflate the influence of the human operator and the automated aid on team performance and may mask changes in the operator's performance strategy under aided conditions. The present study used a measure of parallel processing efficiency, or *workload capacity*, derived from empirical RT distributions as a novel gauge of human-automation performance and automation dependence in a speeded task. **Method** Participants performed a speeded probabilistic decision task with and without the assistance of an automated aid. RT distributions were used to calculate two variants of a workload capacity measure, $C_{OR}(t)$ and $C_{AND}(t)$. **Results** Capacity measures gave evidence that a diagnosis from the automated aid speeded human participants' responses, and that participants did not moderate their own decision times in anticipation of diagnoses from the aid. **Conclusion and Application** Workload capacity provides a sensitive and informative measure of human-automation performance and operators' automation dependence in speeded tasks.

- **Keywords:** human-automation system, workload capacity

NEUROERGONOMICS

Julian Lim and Kenneth Kwok. *The Effects of Varying Break Length on Attention and Time on Task*. Pages 472-481.

Objective: We aimed to discover how varying the length of task breaks would affect the time-on-task effect in subsequent testing periods. **Background:** An important means of preventing errors and accidents caused by mental fatigue and time on task is to intersperse rest intervals within long work periods. Most studies of rest pauses to date have examined their effects in real-world tasks and settings, and their subtler effects on behavior, as measurable by laboratory paradigms, are not well understood. **Method:** We studied a group of 71 participants as they completed a 1-hr auditory oddball task with two rest opportunities. Rest intervals were 1, 5, or 10 min long. **Results:** Improvements in reaction time were significantly positively associated with length of the rest break. However, longer breaks were also associated with steeper decrements in performance in the subsequent task block. Across individuals, the amount of immediate improvement correlated with the extent of later decline. **Conclusion:** Our results support a resource/effort-allocation model of fatigue, whereby longer breaks bias participants toward greater effort expenditure on resumption of the task when cognitive resources may not have been fully replenished. **Application:** These findings may have implications for the refinement of work-rest schedules in industries where time-on-task degradation in performance is an important concern.

- **Keywords:** mental fatigue, sustained attention, rest break, effort allocation, resource theory

SENSORY AND PERCEPTUAL PROCESSES

Sara Lu Riggs and Nadine Sarter. *The Development and Evaluation of Countermeasures to Tactile Change Blindness*. Pages 482-495.

Objective: The goal of the present study was to develop and empirically evaluate three countermeasures to tactile change blindness (where a tactile signal is missed in the presence of a tactile transient). Each of these countermeasures relates to a different cognitive step involved in successful change detection. **Background:** To date, change blindness has been studied primarily in vision, but there is limited empirical evidence that the tactile modality may also be subject to this phenomenon. Change blindness raises concerns regarding the robustness of tactile and multimodal interfaces. **Method:** Three countermeasures to tactile change blindness were evaluated in the context of a highly demanding monitoring task. One countermeasure was proactive (alerting the participant to a possible change before it occurred) whereas the other two were adaptive (triggered after the change upon an observed miss). Performance and subjective data were collected. **Results:** Compared to the baseline condition, all countermeasures improved intramodal tactile change detection. Adaptive measures resulted in the highest detection rates, specifically when signal gradation was employed (i.e., when the intensity of the tactile signal was increased after a miss was observed). **Conclusion:** Adaptive displays can be used to counter the effects of change blindness and ensure that tactile information is reliably detected. Increasing the tactile intensity after a missed change appears most promising and was the preferred countermeasure. **Application:** The findings from this study can inform the design of interfaces employing the tactile modality to support monitoring and attention management in data-rich domains.

- **Keywords:** change blindness, tactile information presentation, multimodal interfaces, adaptive displays, signal gradation

SIMULATION AND VIRTUAL REALITY

Jinling Wang, Amine Chellali, and Caroline G. L. Cao. *Haptic Communication in Collaborative Virtual Environments*. Pages 496-508.

Objective: To understand the interaction between haptic and verbal communication, we quantified the relative effect of verbal, haptic, and haptic-plus-verbal feedback in a collaborative virtual pointing task. **Background:** Collaborative virtual environments (CVEs) provide a medium for interaction among remote participants. Better understanding of the role of haptic feedback as a supplement to verbalization can improve the design of CVEs. **Methods:** Thirty-six participants were randomly paired into 18 dyads to complete a 2-D pointing task in a CVE. In a mixed experimental design, participants completed the task in three communication conditions: haptic only (H), verbal only (V), and haptic plus verbal (HV). The order of the conditions presented to the participants was counterbalanced. **Results:** The time to task completion, path length, overshoot, and root mean square error were analyzed. Overall, performance in the V and HV conditions was significantly better than in the H condition. H was the least efficient communication channel but elicited response with the shortest reaction time. When verbalization was not available, the use of the haptic device was more likely to be exaggerated to ensure information transmission. When verbalization was used, participants converged on the use of a Cartesian coordinate system for communicating spatial information. **Conclusion:** Haptic communication can be used to complete a collaborative virtual task but is less efficient than verbal communication. A training period may help to improve the efficiency of haptic communication. **Application:** These results can be used to design remote collaboration tasks incorporating haptic components and for improving the design of CVEs that support haptic communication.

- **Keywords:** computer-supported collaborations, team communication, multimodality, virtual environments, team collaboration

SURFACE TRANSPORTATION

Sebastian Hergeth, Lutz Lorenz, Roman Vilimek, and Josef F. Krems. *Keep Your Scanners Peeled: Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving*. Pages 509-519.

Objective: The feasibility of measuring drivers' automation trust via gaze behavior during highly automated driving was assessed with eye tracking and validated with self-reported automation trust in a driving simulator study. **Background:** Earlier research from other domains indicates that drivers' automation trust might be inferred from gaze behavior, such as monitoring frequency. **Method:** The gaze behavior and self-reported automation trust of 35 participants attending to a visually demanding non-driving-related task (NDRT) during highly automated driving was evaluated. The relationship between dispositional, situational, and learned automation trust with gaze behavior was compared. **Results:** Overall, there was a consistent relationship between drivers' automation trust and gaze behavior. Participants reporting higher automation trust tended to monitor the automation less frequently. Further analyses revealed that higher automation trust was associated with lower monitoring frequency of the automation during NDRTs, and an increase in trust over the experimental session was connected with a decrease in monitoring frequency. **Conclusion:** We suggest that (a) the current results indicate a negative relationship between drivers' self-reported automation trust and monitoring frequency, (b) gaze behavior provides a more direct measure of automation trust than other behavioral measures, and (c) with further refinement, drivers' automation trust during highly automated driving might be inferred from gaze behavior. **Application:** Potential applications of this research include the estimation of drivers' automation trust and reliance during highly automated driving.

- **Keywords:** trust in automation, eye tracking, autonomous driving, human-automation interaction, vehicle design