

2011 Human Factors Prize Winner

Automation in Surgery: The Impact of Navigated-Control Assistance on Performance, Workload, Situation Awareness, and Acquisition of Surgical Skills

Dietrich Manzey and Maria Luz, Berlin Institute of Technology, Berlin, Germany, Stefan Mueller, Andreas Dietz, Juergen Meixensberger, and Gero Strauss, University Hospital Leipzig, Germany

Objective: Human performance consequences of a new technology of image-guided navigation (IGN) support for surgeons are investigated.

Background: Navigated control (NC) represents an advancement of IGN technology. In contrast to currently available pointer-based systems, it represents a higher degree of automation that supports processes not only of information analysis and integration but also of intraoperative decision making.

Method: In the first experiment, 14 surgical novices performed a simulated mastoidectomy with and without NC support. Effects of provision of the system were analyzed with respect to different measures of surgical performance and outcome, workload, and situation awareness. In the second experiment, 21 advanced medical students were trained to perform a mastoidectomy by practicing it either with or without NC support. It was investigated to what extent the provision of the system during practice would affect the acquisition of surgical skills.

Results: The results reveal that NC support can reduce both the risk of intraoperative injuries and complications as well as the physiological effort of surgeons. "Cost effects" compared to a conventional (i.e., not supported) surgery emerged with respect to the time needed for the surgery, increased subjective workload, reduced spare capacity, and a reduced level of situation awareness. However, no significant effects on processes of skill acquisition were found.

Conclusion: NC systems can contribute to improved patient safety. Most of the cost effects seem to be related not to the basic principle of NC but to its current technological implementation.

Application: The results have consequences for the design and clinical use of automated navigation support.

Keywords: image-guided navigation, human-automation interaction, health care ergonomics

Address correspondence to Dietrich Manzey, Berlin Institute of Technology, Institute of Psychology and Ergonomics, Marchstrasse 12, Sekr. F 7, D-10587 Berlin, Germany; dietrich.manzey@tu-berlin.de.

HUMAN FACTORS

Vol. 53, No. 6, December 2011, pp. 584-599

DOI:10.1177/0018720811426141

Copyright © 2011, Human Factors and Ergonomics Society.

INTRODUCTION

Image-guided navigation (IGN) systems represent an advanced technology to support a surgeon in navigating through a patient's anatomy. The main areas of application are minimally invasive surgeries (MISs), which usually put high demands on the anatomical knowledge and spatial orientation skills of surgeons. Based on a digital computer or magnetic resonance tomography data set of the patient's anatomy, a 3-D camera, and sophisticated tracking technology, IGN systems automatically display the location of the surgical instrument in relation to the anatomical structure of a patient on a specific navigation screen. Prominent areas of application include neurosurgery, functional endoscopic sinus surgery, and other skull base surgeries (Ecke, Luebben, Maurer, Boor, & Mann, 2003; Koele, Stammberger, Lackner, & Reittner, 2002). Yet IGN systems have gained increasing acceptance in other surgical fields (e.g., orthopedics) and have also been applied for surgeries that do not represent MIS but also provide complex navigation challenges (e.g., mastoidectomy).

The first systems of this kind have been in clinical use now for more than one decade (Koele et al., 2002). These pointer-based systems enable the surgeon to recall the current position of his or her surgical instruments at discrete time points during the surgery. For this purpose, the surgeon needs to bring a specific pointer to the same anatomical site as his or her surgical instrument and then has the position displayed on the navigation screen. The principle of navigated control (NC) represents a recent advancement in this area (Strauss et al., 2005; 2007). In contrast to the pointer-based tools, it enables a continuous tracking of the

surgical instrument. Furthermore, and even more important, the computer provides autonomous support for protecting risk structures in the anatomical area where the surgeon is working. Based on a preoperative segmentation of workspace in the digital set of tomographical data (excluding the critical structures), the system continuously assesses the distance between the surgical instrument (e.g., drill) and the workspace border and automatically stops the instrument in case this border is reached.

Theoretically, IGN systems can be regarded as automation where the cognitive task of navigation is no longer performed by the surgeon alone but is shared with a machine (Manzey, Strauss, et al., 2009). The main objectives of introducing these systems are to improve patient safety by reducing the risk of surgical errors and to provide support for a particularly demanding perceptual-cognitive task, thus reducing the physiological effort and workload of the surgeon. However, as is known from other areas (e.g., aviation), the benefits of automated assistance may not always be realized but can be offset by human performance costs associated with new error sources and risks. The latter can include issues of overreliance on the automated functions, issues of workload shifts associated with managing an additional system, and issues of out-of-loop unfamiliarity, that is, difficulties in developing and maintaining an appropriate level of situation awareness because of a reduced active involvement in the automated processes (Ferris, Sarter, & Wickens, 2010).

A generic framework model of types and levels of automation that can be used to classify different automated systems and to evaluate their consequences for human performance has been introduced by Parasuraman, Sheridan, and Wickens (2000; Wickens, Huiyang, Santamaria, Sebok, & Sarter, 2010). Referring to this model, pointer-based IGN systems represent a comparatively low degree of automated assistance that just partially supports the acquisition and integration of visual information from the operative site and leaves any further decision making and implementation of actions completely with the surgeon (Manzey, Strauss, et al., 2009). In contrast, NC systems represent a significant step toward a higher degree of automation. These

systems provide not only information support but also support of intraoperative decision making (i.e., where to work and where to stop) and the implementation of appropriate actions (i.e., an automatic stop of the instrument in case it approaches the border of the work space).

Most of the available research evaluating IGN systems has addressed aspects of its technical performance. This is reflected in clinical studies investigating the suitability of these systems for certain surgeries (e.g., Reijnders et al., 2007), the performance of certain technological approaches (e.g., Casap, Wexler, & Eliashar, 2008), and the benefits provided by these systems in terms of surgical outcomes (e.g., Gong, Mohr, & Vézina, 2007). In contrast, only one study is available, thus far, that has attempted to address performance consequences of IGN systems from a human factors perspective (Manzey, Roettger, et al., 2009). In this survey study, surgeons were asked to subjectively assess the impact of pointer-based systems on different aspects of performance compared to the “gold standard,” that is, an unassisted surgery. Beneficial effects reported included an improved quality of surgical outcome, a reduced level of effort and stress during the surgery, and an increase of situation awareness with respect to a correct orientation within the operative site. However, negative side effects also were reported, including a prolonged time needed for the surgery because of the necessity of dealing with another instrument. Furthermore, some indications of overreliance were found, which seemed to affect about 20% of the surgeons participating in the survey. Finally, concerns were raised that the provision of automated navigation assistance already during surgical training might interfere with the acquisition of surgical skills that, in turn, would make surgeons more or less dependent on the availability and reliability of the systems. Although providing first insights into the performance consequences of IGN tools, these results have a conclusiveness that seems to be limited. First, all surgeons responding to the survey did use IGN systems on a regular basis, which might have biased their responses positively. Second, the systems considered in the survey represented pointer-based IGN systems, which represent a rather low degree of automation (see earlier discussion).

This raises the question of to what extent the results might be generalized to the next generation of higher automated IGN systems, that is, NC technology. This is suggested by findings that human performance consequences of automated aids are dependent on their degree of automation. Specifically, they suggest that higher degrees of automation are beneficial for the intended positive effects but may also raise the risk of loss of situation awareness, or return-to-manual performance issues that, in turn, might negatively affect performance in case of automation failures (Endsley & Kiris, 1995; Wickens et al., 2010).

In addition, another important question relates to the impact of these systems on skill acquisition of surgeons. This has been a matter of concern already with respect to the provision of pointer-based systems during surgical training (Manzey, Roettger, et al., 2009). It can be assumed that this risk even increases with the higher automated NC systems. Such effect would suggest that novice surgeons who have been trained with automation support should be certified to conduct a given surgery only if such a system is available.

The two experiments described in this article represent a first approach to investigate the human performance consequences of NC support experimentally. Participants of these experiments were clinically trained students and physicians. The model used for this research was a simulated mastoidectomy, which represents a particularly risky surgical intervention at the petrosal bone behind the ear. The task is to remove as much of the infected bone of the mastoid as possible by means of a mill (medical term: trephine) without injuring relevant risk structures present in this anatomic area. Risk structures to be protected include the dura mater, the sigmoid sinus, the auditory ossicles, the lateral semicircular canal, and—most important—the nervus facialis. This model was chosen because it represents a compact surgery that nevertheless is sufficiently complex to assess many different aspects of performance (e.g., surgical outcome, workload, and situation awareness). Furthermore, a high-fidelity device for simulating this surgery in the laboratory was available.

The first experiment investigated the effects of NC support on surgical performance, workload, and situation awareness compared to the conventional approach (unassisted surgery). Measures of surgical performance included a registration of injuries of risk structures, an evaluation of the time needed for the surgery, and expert evaluations of different aspects of the surgical outcome. Workload was assessed by means of subjective ratings, secondary task performance, and several physiological measures. Situation awareness was assessed by using the Situation Awareness Global Assessment Technique (SAGAT; see Endsley, 2000). With respect to positive effects of NC support, it was expected that the provision of NC support would enable even surgical novices to conduct the surgery without compromising patient safety. In addition, it was expected that NC support would reduce the level of physiological effort during the surgery. With respect to possible negative side effects, it was hypothesized that the demand to interact with the assistance system would increase the subjectively experienced workload and prolong the time of the surgery. This was expected based on the results of pointer-based assistance by Manzey, Roettger, et al. (2009). In addition, it was assumed that, because of the increased degree of automation, NC support would negatively affect the situation awareness of surgeons, reflected, for example, in less awareness of the distance to risk structures or specifics of a patient's anatomy.

The second experiment addressed the impact of NC support on the acquisition of surgical skills. This represents a topic that has rarely (if ever) been addressed in human-automation interaction research. For this purpose, the effectiveness and efficiency of performing a simulated mastoidectomy manually, that is, without NC support, was compared for two groups of medical students who were trained to perform this surgery with and without NC support. It was hypothesized that providing NC support already during training would interfere with the acquisition of proper surgical skills. More specifically, it was assumed that participants who were trained with NC support would be less able to protect risk structures, would experience more stress and workload and would have a

lower level of situation awareness than conventionally trained surgeons in case the NC system was not available.

EXPERIMENT 1

Method

Participants. A total of 14 advanced students of medicine and physicians who just had finished their studies participated in the experiment (5 male, 9 female; mean age = 26 years). On average the participants had their first clinical experiences from a 4.5-week clerkship (range = 0–12 weeks), and 11 of the participants already had their first surgical experiences from an internship. None of them had any experience in performing a mastoidectomy. Participants were paid € 25 as compensation for their participation in the study.

Task. The primary task consisted of a simulated mastoidectomy. Participants were instructed to perform this task as well as possible without injuring any of the risk structures that are present in this anatomical area. They were further instructed to work according to a given sequence of steps, that is, (a) break through the *corticalis*, (b) express the dura mater, (c) express the sigmoid sinus, (d) lance the antrum, (e) express the posterior wall of the auditory channel, (f) express the sinus-dura angle, and (g) express the nervus facialis. In addition to this primary task, a secondary task had to be performed. This task represented a simple probe reaction time task. The participants had to react to an acoustic signal as quickly as possible by pressing a foot pedal. The signal was presented randomly with a mean interstimulus interval of 90 s and a standard deviation of this interval of 5.5 s.

Apparatus. An artificial skull with an exchangeable petrosal bone module was used for the simulation. All modules used were designed to correspond as closely as possible to the real anatomy of the petrosal bone. These included typical levels of pneumatization of the bone material as well as a simulation of all risk structures that are represented in this anatomical area. All risk structures but the auditory ossicles were equipped with integrated sensors, which allowed for automatically logging any intrasurgical injuries of these structures. Two kinds of anatomies were used, one with normal

pneumatization and standard position of all risk structures and another one also normally pneumatized but with an exteriorized sigmoid sinus.

The NC system (KARL STORZ) used for NC-supported interventions consisted of a 3-D camera system, a trephine equipped with sensors needed for instrument tracking, a navigation display, and the control unit. To use this system for the simulation, the risk structures of the artificial petrosal bone were segmented out in the digital CT data set by one of the coauthors (S.M.). This corresponds to the usual clinical preparation needed for use of NC assistance (Strauss et al., 2007).

Physiological data (ECG, respiration) were recorded by a mobile recording device (NeXus-10, version 2008a; Mind Media B.V., Netherlands). Respiration was assessed via a respiration belt. The ECG was sampled at a rate of 256 Hz from electrodes placed on the chest. Mean heart rate and respiration rate data were derived based on the peak detection algorithms implemented in the recording system. To calculate frequency domain measures of heart rate variability, the time series of interbeat intervals was derived from the ECG by means of Nevrokard HRV Preparation software (Version 9.2.2) and further analyzed by Carspan for Windows (Version 0.0.1.26) developed by Ben Mulder (University of Groningen). Using a moving-window technique, mean normalized power values for two different frequency bands were calculated: mid-frequency band (MF; 0.07–0.14 Hz) and high-frequency band (HF; 0.15–0.40 Hz). For blood pressure recordings at discrete points of time, a standard sphygmomanometer was used.

Design. Each participant performed a complete mastoidectomy with and without NC assistance. The sequence of the two experimental runs was balanced across participants.

Dependent measures. Four different performance measures were defined to assess surgical performance: (a) the time needed to complete the mastoidectomy, defined as the sum of time (minutes) from the start of the trephine until completion of the simulated surgery, excluding times needed for the assessment of situation awareness (see the following discussion), (b) the number of injuries to risk structures, which were automatically registered by the simulation

software, (c) expert ratings of the overall quality of the surgical outcome (5-point rating scale ranging from *very good* to *very bad*), and (d) expert ratings about whether or not seven possible complications might be expected to arise from the surgical outcome (4-point rating scale with stages defined as *sure no*, *rather no*, *rather yes*, *sure yes*). All expert ratings were provided by an experienced senior surgeon and collected according to a double-blind procedure.

Three different kinds of measures were used to assess different aspects of workload (O'Donnell & Eggemeier, 1986): (a) subjective workload was assessed with the NASA Task Load Index (NASA TLX; Hart & Staveland, 1988), (b) demands on attentional resources were assessed by secondary task performance operationally defined by the mean of probe reaction time, and (c) physiological effort was assessed by mean heart rate (HR), mean respiration rate (RR), systolic blood pressure (BP), and the MF/HF ratio derived from the frequency domain measures of HRV. The latter index was chosen because it has been shown to sensitively reflect differences in "mental strain" of surgeons while conducting different kinds of surgery (Boehm, Roetting, Schwenk, Grebe, & Mansmann, 2001).

The SAGAT (Endsley, 2000) was used to assess situation awareness (SA). After a pre-defined step of the surgery had been completed (lancing of the antrum), the task was interrupted and the participants were required to answer four different sets of questions. The questions focused on (a) an estimation of the relative distances between the last position of the milling head and six defined anatomical structures at the moment of interruption, (b) a report on the already accomplished subtasks, (c) a report on specific aspects of the anatomy (i.e., position of the sigmoid sinus; level of pneumatization), and (d) the estimation of the remaining time needed for completing the surgery. The answers were validated by objective measurements (Questions a and d), observations of the experimenters (Question b), or the objective characteristics of the simulated anatomies (Question c). Four indicators of SA were derived: (a) the rank correlation between the estimated and real distances calculated across the six risk structures, which reflects how accurately participants were aware of their current

spatial position in the 3-D space of anatomy (0–1), (b) the number of correctly remembered steps (0–7), (c) the number of correct answers about the anatomical characteristics (0–3), and (d) the deviation between estimated and real remaining time (in min) needed for the surgery after the break.

Procedure. The experiment consisted of two practice sessions and two experimental runs. Before the first practice session, all participants were provided a script that familiarized them with the anatomical characteristics of the petrosal bone, the critical risk structures in this area, the different steps of a mastoidectomy, and the relevant surgical instruments. The actual task training included two sessions where the participants had to perform the mastoidectomy on normally pneumatized petrosal bone modules with standard anatomy. These runs had to be performed manually, that is, without NC assistance. This ensured that participants acquired sufficient technical skills to perform the experimental trials without the risk that basic problems of device handling would interfere with the effects of the experimental treatments.

The two experimental runs were distributed across two different days. They took place in the endoscopic operating theater of the ORL University Hospital in Leipzig. Before starting the simulated surgery, each participant had the opportunity to familiarize himself or herself with the medical instruments. In the NC condition, this included an introduction to the functions of the navigation system and its use. Then several baseline measures were collected. These included a measure of single-task response times in the secondary task (only first session), a measure of BP, and the recording of a 5-min baseline of ECG and respiration data. Participants were instructed to rest during this period.

The simulated surgery itself was performed in sitting position. The trephine was controlled by a foot pedal that was located on the right side underneath the surgery table. On the left side, another pedal, which was used to register responses to the secondary task, was located. The camera of the navigation system and the navigation screen were placed in front of the participants. In manual conditions the navigation screen was taken out of their sight.

The mastoidectomy had to be performed according to the standard protocol. However, the anatomy of the modules was varied compared to the training trials by using modules with an exteriorized sigmoid sinus. This variation was introduced to simulate the normal interindividual variety of anatomies and to elevate the risk of a serious injury of the simulated patient. After the lancing of the antrum, the participants were interrupted and asked to turn to the experimenter. Immediately after this interruption, the SA questionnaire was presented to the participants. In addition, a measure of BP was taken during the break. Thereafter, the surgery continued until the participants indicated that the intervention was completed. When completed, subjective workload ratings were collected and postoperative baseline measurements of all physiological parameters were performed.

Results

Performance. A significant difference between the two experimental conditions was found with respect to the time needed to complete the mastoidectomy. Participants needed substantially less time to perform this task under manual conditions (64 min) compared to the condition where they were supported by the NC system (100 min), $t(13) = 4.31, p = .001, \eta^2 = .59$. In contrast, the overall quality of the surgeries did not differ. This was revealed by almost equal expert ratings for modules that had been milled manually (2.57) and those milled with NC support (2.36), $t(13) = 0.68, p = .51, \eta^2 = .03$. However, the most important difference between conditions emerged with respect to the number of injuries. As expected, no injuries of risk structures were found for NC-assisted surgeries. In contrast, 21% ($n = 3$) of the surgeries performed manually resulted in an injury of the sigmoid sinus, which in reality would have had severe consequences for the patient. This advantage of NC support was also reflected in expert ratings of possible complications that might have arisen during the surgery. Averaged across all kinds of complications, this possibility was rated higher for the manual (1.44) than for the NC-assisted conditions (1.18), $t(13) = 2.62, p < .03, \eta^2 = .35$.

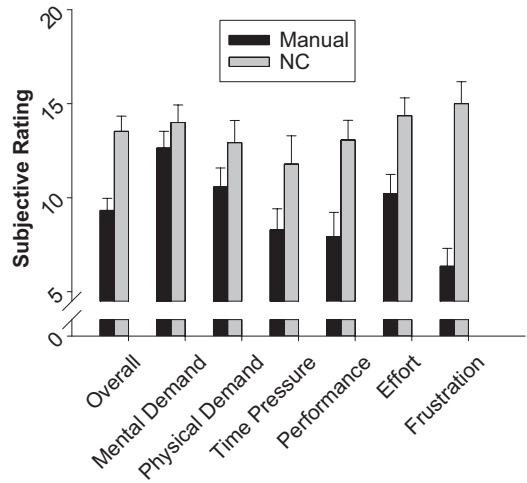


Figure 1. Effects of navigated control (NC) support on subjective workload in Experiment 1. Displayed are means and standard errors of Task Load Index ratings dependent on whether the simulated surgery had to be performed manually or with support of the NC system.

Workload. Effects of the two experimental conditions on subjective workload assessed by the NASA TLX are shown in Figure 1. A significant difference between the two experimental conditions was found for overall workload defined as the average rating across the different dimensions of TLX, $t(13) = 3.83, p < .01, \eta^2 = .53$. Participants reported higher workload in the NC condition (13.5 on the scale from 0 to 20) than in the manual condition (9.3). More detailed analyses revealed that this difference was mainly determined by a considerably higher frustration level while working with the support system (15.0) in contrast to manual milling (6.4), $t(13) = 4.82, p < .001, \eta^2 = .64$. All other dimensions of the NASA TLX pointed in the same direction, but differences did not become significant when Bonferroni adjustments of the alpha level were taken into account. Observations during the different experimental trials suggested that the different frustration levels were mainly related to a high number of “false alarms,” that is, stopping events in the NC condition that were related not to the protection mechanism but to other causes (e.g., loss

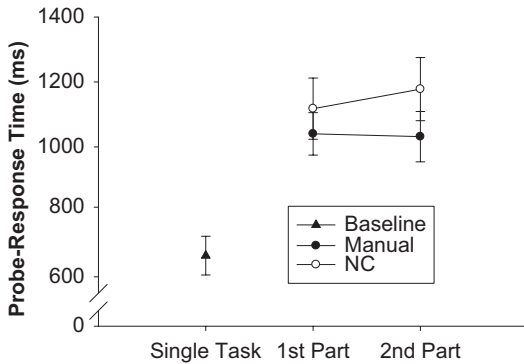


Figure 2. Effects of navigated control (NC) support on secondary task performance in Experiment 1. Displayed are means and standard errors of probe reaction times for single-task (baseline) performance and concurrent-task performance during the two parts of the simulated surgery (before and after break for situation awareness assessment).

of line of sight). Because it was not possible to immediately identify these events correctly as “false alarms,” they usually required a short cross-check of the reason by looking up from the microscope and analyzing the information on the navigation display.

This need for cross-checks in response to a significant number of stopping events might also have contributed to the raised attentional demands in the NC compared to the manual condition, which were reflected in secondary task performance. These effects are shown in Figure 2. Compared to the single-task response times ($M = 629$ ms), participants needed almost as twice as long (1,095 ms) for responding to the acoustical probe during milling, irrespective of whether or not NC support was available. However, this effect was significantly more pronounced in the NC condition. A 2 (condition) \times 2 (time of measurement, i.e., before vs. after interruption for SA assessment) ANOVA revealed that participants on average needed significantly more time to respond to the acoustic probe in the NC condition ($M = 1,151$ ms) than when milling manually (1,039 ms), $F(1, 13) = 7.94, p < .02, \eta^2 = .38$, and this difference was higher during the second as compared to the first phase of the surgery when the participants had to cope with the difficult task of

expressing the nervus facialis, $F(1, 13) = 4.99, p < .05, \eta^2 = .28$. Yet the main effect of time of measurement (ToM) did not become significant for this measure, $F < 1$.

The effects for the physiological measures are shown in Figure 3. Despite the raised level of frustration and longer reaction times in secondary task, all physiological data but the MF/HF index of HR variability pointed to a lower physiological effort level while milling with NC compared to the manual condition. HR (upper left panel) showed a stronger increase during the simulated surgery under manual conditions compared to NC, and this difference also persisted in the postbaseline measurement. This effect was analyzed by a 2 (condition) \times 4 (ToM: prebaseline, Part 1, Part 2, postbaseline) ANOVA. It revealed a main effect of condition, $F(1, 12) = 7.29, p < .02, \eta^2 = .38$, and a main effect of ToM, $F(3, 36) = 45.90, p < .001, \eta^2 = .79$. The interaction effect just failed to reach the usual level of significance, $F(3, 36) = 2.80, p = .054, \eta^2 = .19$. A similar pattern of effects emerged for RR (upper right panel). It showed a significant increase during the simulated surgery, $F(3, 36) = 77.44, p < .001, \eta^2 = .87$, and this increase was steeper in the NC than the manual condition, $F(3, 36) = 5.63, p < .01, \eta^2 = .32$. However, the main effect of condition did not become significant for this measure, $F(1, 12) = 1.70, p = .22, \eta^2 = .12$.

The effects for systolic BP were analyzed by a 2 (condition) \times 3 (ToM: prebaseline, surgery, postbaseline) ANOVA. This analysis revealed a significant main effect of ToM, $F(2, 26) = 4.13, p < .05, \eta^2 = .24$, and a significant Condition \times ToM interaction effect, $F(2, 26) = 4.68, p < .05, \eta^2 = .26$. The main effect of condition just failed to reach a conventional level of significance, $F(1, 13) = 4.37, p = .057, \eta^2 = .25$. As becomes evident from Figure 3 (lower-left panel), systolic BP was higher in the manual than the NC condition already during the preoperative baseline, remained at this level during the first phase of the simulated surgery, and showed a considerable increase during the second phase when participants had to work on the difficult task to express the nervus facialis. The latter effect was reflected in the measurement taken immediately after completion of the surgery. In the NC

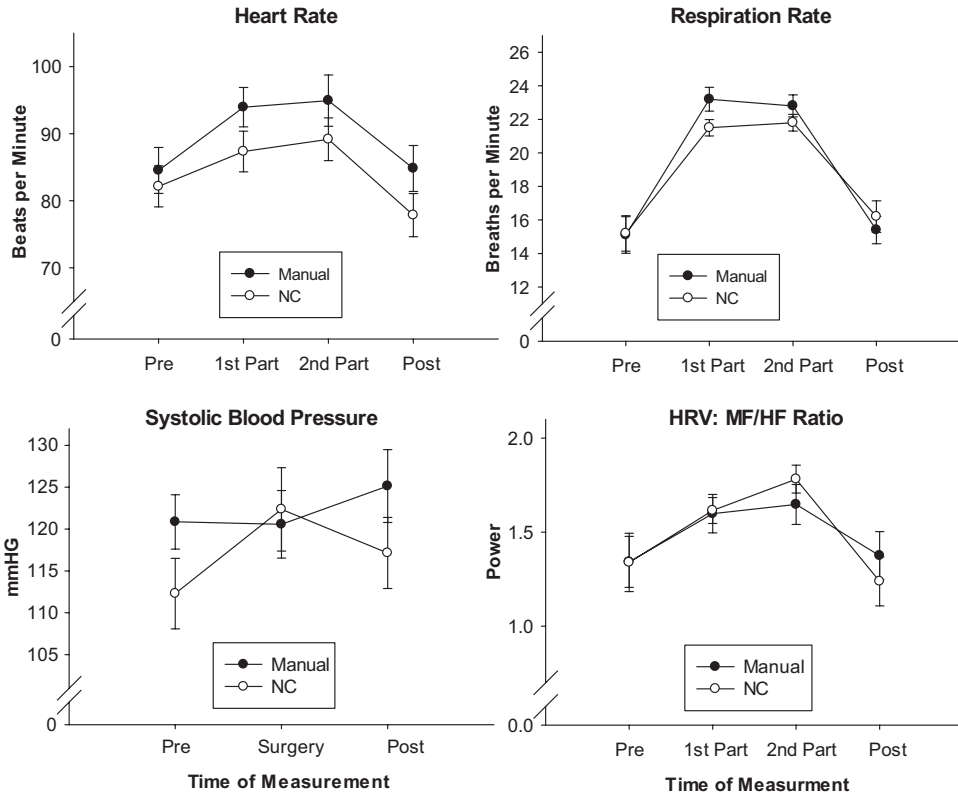


Figure 3. Effects of navigated control (NC) support on physiological effort. Displayed are means and standard errors for the different physiological measures sampled in Experiment 1 as a function of experimental condition (manual vs. NC) and time of measurement (prebaseline, the two parts of the surgery until and after the break for situation awareness assessment, postbaseline). For blood pressure, only three measurements were taken (before and immediately after the simulated surgery as well as during the break for situation awareness assessment).

condition, an increase in systolic BP was found during the first phase of the surgery. However, in contrast to the manual condition, it decreased again to the preoperative level across the second phase.

In contrast to the foregoing measures, the MF/HF index increased only during the simulated surgery but did not show any effect of whether or not the automation support was available. A 2 (condition) × 4 (ToM) ANOVA revealed a significant effect of ToM, $F(3, 36) = 5.89, p < .02, \eta^2 = .33$. Neither the effect of condition, $F < 1$, nor the interaction became significant, $F(3, 36) = 1.21, p = .32, \eta^2 = .09$.

Situation awareness. In both conditions participants were equally able to estimate the relative distances from the milling head to the risk

structures at the moment of interruption. Rank correlations between subjective estimates and real differences measured based on screenshots were $r = .47$ (manual) and $r = .51$ (NC) and did not differ significantly. Similarly, no evidence was found for the expected decline of SA in the NC condition with respect to the awareness of already accomplished subtasks and the estimation of the remaining time needed for the surgery. These directed hypotheses were tested by one-tailed t tests contrasting SA performance in the two conditions but did not reveal significant effects for either measure, $t(13) = 1.24, p = .12, \eta^2 = .11$, and $t(13) = 1.17, p = .13, \eta^2 = .09$. However, a significant difference was found with respect to the assessment of anatomical characteristics. Participants were better able to

assess specifics of the anatomy when they performed the surgery without (1.64 of 3 possible correct answers) than with NC support (1.0), $t(13) = 1.8$, one-tailed $p < .05$, $\eta^2 = .20$. This effect was most obvious for the assessment of the specifics of the dura mater and of the sigmoid sinus, although these structures were directly relevant for the intervention. There was no considerable difference in respect to a correct description of the level of pneumatization.

Discussion

The current study represents the first approach to evaluate human performance consequences of automated navigation support experimentally. With respect to the effects on the surgeons' primary task performance, a mixed pattern of results emerged. On one hand, the availability of NC support contributed to an increase of patient safety. In the NC condition, even comparatively inexperienced surgeons were able to carry out the complex surgical intervention without any intraoperative injuries to important risk structures while reaching equal milling quality. This not only supports the effectiveness of NC for protection of risk structures but also suggests that this kind of support can compensate for differences in the surgical experience of surgeons. On the other hand, the efficiency, for example, time needed for the intervention, increased with NC support. This effect resembles earlier results found for pointer-based tools (Manzey, Roettger, et al., 2009; Metson, Consenza, Cunningham, & Randolph, 2000). However, in contrast to these earlier results, which primarily seemed to be related to the need for handling additional equipment, the time costs for NC appear to be mainly the result of events where the trephine stopped because of technical reasons (e.g., loss of line of sight). That is, they seem to be related not to the basic concept of the NC but to its kind of technical implementation. Specifically, the strict protection function, which leads to a complete stop of the surgical instrument and repeated interruptions of the workflow of surgeons, seems to represent a major problem in this respect, as it further raises the already high work demands of the surgery. It can be assumed

that a less "invasive" implementation of NC support, for example, an implementation that provides just an alarm signal or a slowing down of the instrument, would reduce this problem.

The repeated interruptions of workflow might also explain the observed increase of subjective workload and decrease of secondary task performance in the NC condition. Participants obviously got distracted by the repeated interruptions of workflow that captured their attention and, thus, were less able to respond to the additional task. This became particularly evident during the second part of the surgery, when the participants needed to pass the exteriorized sigmoid sinus and had to express the nervus facialis, which produced a particularly high number of interruptions.

However, despite these effects, a clear beneficial effect of NC support emerged with respect to a lower level of physiological effort during the simulated surgery. Compared to the manual condition, this was reflected in fewer increments of HR and RR as well as a recovery of initial increments of BP over the course of the surgery. This effect confirms the hypothesis that automated navigation support can contribute to lowering the level of physiological effort during a surgical intervention. It can be suggested that it directly relates to the protection function of the NC system, which effectively prevented the participants from coming too close to risk structures and, thus, lowered the perceived and actual risks for patient safety while performing the intervention. The fact that this effect was not reflected in the MF/HF index contradicts the results of Boehm et al. (2001), who found this index sensitive to differences in "mental strain" of surgeons. However, it is in line with other results suggesting that HRV sensitivity for workload variations is comparatively low for complex tasks (Jorna, 1992).

As expected, SA was affected by NC support. Specifically, participants working with NC support were less able to describe correctly the specifics of the positions of different anatomical structures. This contradicts results for the less automated pointer-based navigation systems from the survey study of Manzey, Roettger, et al. (2009), where navigation systems were reported to provide benefits with respect to SA.

The underlying reasons for this effect cannot be derived conclusively from the current data. On one hand, this finding confirms earlier results that higher automated aids may reduce SA (Endsley & Kiris, 1995). However, the effect was comparatively weak and was reflected in only one out of the four indicators used for SA assessment in the current study. In particular, doubts are cast on this interpretation by the fact that participants working with NC support also were able to describe their spatial position within the operative site as good as in the manual condition. Based on the current data, it cannot be excluded that the comparatively less surgical experience and knowledge of the participants also might have contributed to this specific effect, and more research seems to be needed to investigate this further.

In summary, the results suggest that NC support provides beneficial effects for patient safety and physiological effort of surgeons. Issues identified included increased time demands, a higher level of distraction, and at least some indications of reduced SA when using the system. Although most of these issues might be tolerable in the clinical practice, given that the higher time demands are considered in the scheduling of surgeries, it might lead to more serious effects if the systems are already used during surgical training. In this case, the higher distraction level as well as possible negative impacts on perception of specifics of anatomical structures might interfere with the acquisition of proper surgical skills and knowledge. Thus far, no study is available that investigated the effects of automation support on the acquisition of surgical skills. A second experiment was conducted that specifically addressed this question.

EXPERIMENT 2

Method

Participants. A total of 21 advanced medical students ranging in age from 21 to 29 years ($M = 23.7$) participated in the study. They were randomly assigned to the control (10) or the experimental group (11). None of the participants had any experience in performing a mastoidectomy. Participants were paid 50 as compensation for participation in the study.

Design. The study composed a 2 (group: manual vs. NC) \times 2 (training: pre- vs. posttraining assessment) design. The first factor was defined as a between-subjects factor. Although 10 of the participants practiced the mastoidectomy conventionally, the remaining 11 students were supported by an NC assistance system. Practice of the surgery comprised five different sessions lasting about 90 min each. During each session, one complete mastoidectomy had to be performed. The second factor represented a within-subjects factor. The two stages of this factor contrasted the participants' manual, that is, unassisted, performance in conducting a mastoidectomy before and after the five training sessions.

Apparatus. The same apparatus as in the first experiment was used. Again, the risk structures of the artificial petrosal bones were segmented out of the digital CT data set by one of the coauthors (M.L.).

Dependent variables. To assess practice effects in terms of surgical outcome, workload, and SA, the same measures as in the first experiment were collected.

Procedure. All participants participated in a total of nine sessions distributed across different days. The first two sessions consisted of a 1-hr lecture describing the anatomy of the petrosal bone, the general procedure of a mastoidectomy, and the functions of the NC system including a specific sensitization for the "false alarm" issue (first session) as well as a familiarization with the simulation environment, the instruments needed for the surgery, and the practical use of the trephine (second session). The third session included the pretraining assessment. During this session, all participants performed a mastoidectomy manually, that is, without automation support. The entire practice of the surgery included five sessions. Although the participants in the manual group practiced the surgery without automation support, the participants of the NC group got support by the NC system. To make the participants aware of possible differences in the anatomy of the petrosal bone and to provide new challenges for navigation during each training trial, modules with "normal" anatomy (as during pre- and posttraining assessment) were used for Sessions 5

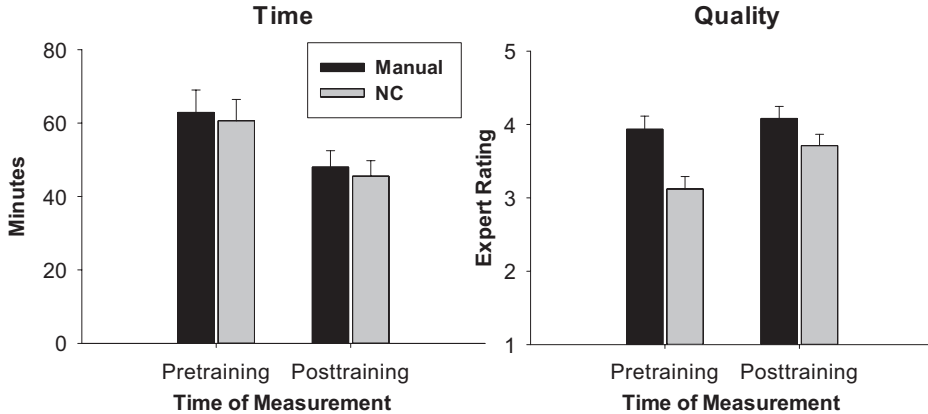


Figure 4. Effects of training on surgical performance in Experiment 2. Displayed are means and standard errors for time needed for surgery (left) and the expert ratings of different aspects of quality of surgical outcome (right) dependent on whether the simulated surgery was performed before or after the five training sessions. NC = navigated control.

and 7, modules with a deep dura mater in Sessions 4 and 8, and a module with an exteriorized sigmoid sinus in Session 6. After completion of the intervention, all participant groups got feedback about their performance. This feedback involved a comparison of the just worked module to a “standard” module, that is, a module considered as the perfect outcome of a mastoidectomy by an experienced surgeon.

Session 9 included the posttraining assessment. In this session, all participants had to perform the surgery again without NC assistance. All sessions took place in a laboratory of the ORL University Hospital in Leipzig. The setup of hardware resembled as close a possible the setup in the operating theatre used for the first experiment. Similarly, all procedural details of data collection in Sessions 3 (pretraining baseline) and 9 (posttraining assessment) corresponded to the procedure used in the first experiment.

Results

Surgical performance. Effects of practice on time needed for the surgery and the quality of surgical outcome (expert rating) are shown in Figure 4 and were analyzed by a 2 (group) \times 2 (training: pre- vs. postassessment) ANOVA. The quality of the surgical outcome was better for the manual than for the NC training group, $F(1, 19) =$

8.78, $p < .01$, $\eta^2 = .32$, and improved after training significantly for both groups, $F(1, 19) = 8.58$, $p < .01$, $\eta^2 = .31$. However, no interaction effect was found, $F(1, 19) = 3.08$, $p < .10$, $\eta^2 = .14$. For the time needed for the simulated surgery, only a main effect of training emerged, indicating an improvement in surgical skills, $F(1, 19) = 11.96$, $p = .003$, $\eta^2 = .39$. Neither the group effect nor the Group \times Training interaction became significant, all $F_s < 1$.

Even more important for the question addressed in this experiment were possible effects of manual versus NC-supported practice on the number of injuries to risk structures, as objectively registered by the simulation software, and the assessment of risk of different complications that would have arisen during the surgery as assessed by expert ratings. In general, the number of injured risk structures was very low before as well as after the practice trials. Only three incidents of this sort occurred during the pretraining assessment, one in the manual group and two in the NC group. However, it is remarkable that after the five trials of practice, participants in both groups were able to perform the mastoidectomy without any injury to risk structures. A similar effect also was reflected in the expert ratings of risks of complications. Generally, the mean risk ratings averaged across several complications were

TABLE 1: Summary of ANOVA Results for the Different Physiological Measures Sampled in Experiment 2

Effect	df1, df2	HR		HRV MF/HF		BP Syst.		Resp.	
		F	η^2	F	η^2	F	η^2	F	η^2
Group	1, 19	1.77	.09	3.31	.15	<1	.01	5.07*	.21
Training	1, 19	<1	.04	<1	.05	26.57**	.58	3.86	.17
Time of measurement (ToM)	3, 57	50.34**	.73	45.62**	.71	9.37**	.33	185.56**	.91
BP Syst.	2, 38								
Group × training	1, 19	<1	.00	<1	.00	<1	.01	<1	.01
Group × ToM	3, 57	<1	.03	<1	.03	2.63	.12	<1	.02
BP Syst.	2, 38								
Training × ToM	3, 57	2.09	.10	1.37	.07	<1	.04	<1	.01
BP Syst.	2, 38								
Group × training × ToM	3, 57	<1	.02	1.05	.05	<1	.02	<1	.02
BP Syst.	2, 38								

Note. HR = heart rate; HRV MF/HF = ratio of power in the mid- and high-frequency band; BP syst. = systolic blood pressure; Resp. = respiration rate.

* $p < .05$. ** $p < .01$.

already low for both groups in the pretest (manual = 1.03, NC = 1.10). For the worked modules of the posttraining, issues of complications were identified in neither group (all single risk ratings = 1.0). Because of the uniform ratings provided for the posttraining assessment, a formal statistical analysis could not be performed on these data.

Workload. Effects of practice on physiological effort were analyzed by a 2 (group) × 2 (training) × 4 (ToM) ANOVA. As in the first experiment, the ToM factor represented a within-subjects factor contrasting the physiological activity during the presurgery baseline, the first part of the surgery (until the SAGAT break), the second part of the surgery (after the SAGAT break), and the postsurgery baseline. The pattern of effects for the different physiological variables is shown in Table 1.

As becomes evident, the most consistent effects were found for the ToM factor. All variables indicated a significantly raised level of physiological effort during the surgical intervention, compared to the two baseline measurements. A significant effect of training was found only for systolic BP. This effect is shown

for the manual and NC training group separately in Figure 5. As becomes evident, systolic BP raises during the simulated surgery. However, for both training groups this increase was considerably reduced after training.

More evidence for training effects on workload was provided by the analyses of subjective ratings and secondary task performance. A 2 (group) × 2 (training) ANOVA of the overall mean of NASA TLX ratings revealed a significant effect of training, $F(1, 19) = 13.22, p < .01, \eta^2 = .41$, reflected in a significant decrease in experienced workload while performing the mastoidectomy after training ($M = 6.7$) as compared to the pretraining assessment (9.8). The Mental Demand (14.2 vs. 7.9), Performance (9.3 vs. 5.6), and Effort (11.7 vs. 8.2) subscales contributed the most to this effect. However, neither the main effect of group nor the Group × Training interaction became significant, all $F_s < 1.8$. Essentially the same effect was found for secondary task performance. Both groups were able to respond more quickly to the acoustic probes after the practice trials as compared to the pretraining assessment. A 2 (group) × 2 (training) × 3 (ToM: single-task RT, first part of

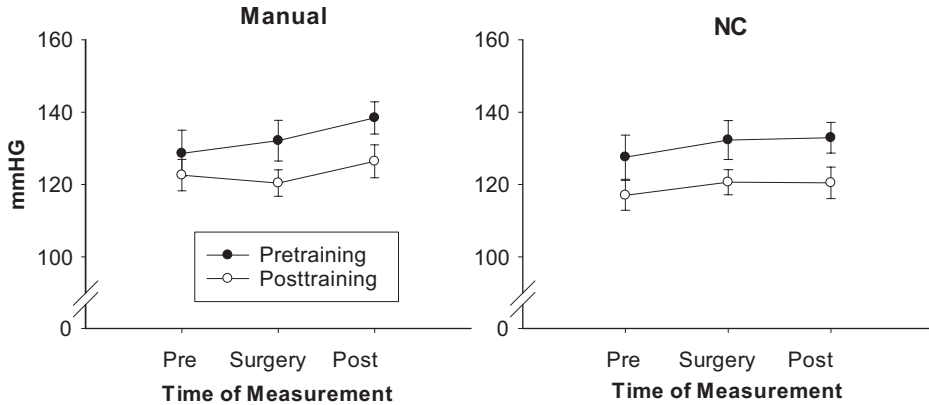


Figure 5. Training effects on systolic blood pressure (BP) for the two training groups of Experiment 2. Displayed are means and standard errors for BP measurements taken before, during, and immediately after the simulated surgery. Measurements during the surgery were taken in the break for situation awareness assessment. NC = navigated control.

the surgery, second part of the surgery) ANOVA revealed a significant main effect of training, $F(1, 19) = 62.72, p < .01, \eta^2 = .77$, and of ToM, $F(2, 38) = 135.57, p < .01, \eta^2 = .88$, as well as a significant Training \times ToM interaction, $F(2, 38) = 15.96, p < .01, \eta^2 = .46$. This effect is shown in Figure 6. As becomes evident, probe reaction times increased significantly during the surgery compared to single-task performance. Yet this effect was considerably smaller after completion of the training session. Neither a group effect nor any of the remaining interaction effects became significant, all $F_s < 1.3$.

Situation awareness. Significant effects of training were found for all measures of SA but the estimation of distances to risk structures. After completion of the practice trials, the participants were better able to correctly report which out of seven different subtasks they already had accomplished (6.7 vs. 5.6 correct answers), $F(1, 19) = 12.58, p = .002, \eta^2 = .40$, to correctly answer the three questions concerning specifics of the simulated patient's anatomy (2.0 vs. 1.1 correct answers), $F(1, 19) = 14.94, p = .001, \eta^2 = .44$, and to estimate the remaining time needed for the surgery (2.2 vs. 11.2 min deviation from real time), $F(1, 19) = 16.48, p = .001, \eta^2 = .46$. However, all of these effects turned out to be independent on whether or not the participants were provided NC support during their training sessions. Neither the main

effect of group nor the Group \times Training interaction became significant for any of these measures, all $F_s < 1$. For the estimation of relative distances to risk structures, none of the effects became significant, all $F_s < 1$.

Discussion

It was expected that the provision of automated navigation support already during surgical training might interfere with the acquisition of proper surgical skills. Specifically, it was anticipated that participants trained with NC support would be less able to protect risk structures, would experience more stress and workload, and would have a lower level of SA than would conventionally trained surgeons if the NC system were no longer available. The results of the second experiment contradict this hypothesis. After five training sessions, participants in both training groups were able to perform the simulated surgery more effectively and efficiently as compared to the pretraining baseline. The higher level of effectiveness was reflected in a reduced time needed for the surgery, a qualitatively better surgical outcome, and—most important—a higher level of patient safety. Gains in terms of efficiency were reflected in a reduced subjective workload, a higher level of spare capacity indicated by improved secondary task performance, and a reduced increase in systolic BP during the

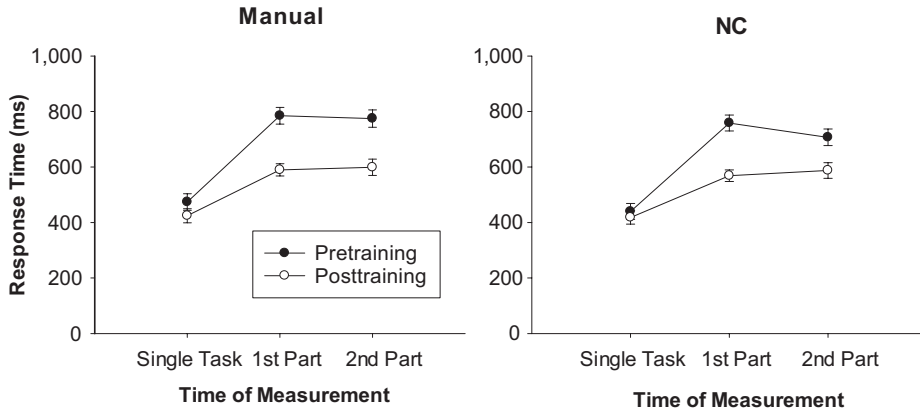


Figure 6. Training effects of secondary task performance for the two training groups of Experiment 2. Displayed are means and standard errors of probe reaction times for single-task (baseline) performance and concurrent-task performance during the two parts of the simulated surgery (before and after the break for situation awareness assessment). NC = navigated control.

simulated surgery. In addition, also the SA of the participants for different aspects was found to be improved after the training.

All of these effects emerged independent of whether the participants were trained conventionally or with the help of NC support. This suggests that providing automated navigation support already during the early phase of skill acquisition does not prevent trainees from acquiring relevant knowledge and skills needed to orientate themselves in the operative site, to protect important risk structures, and to perform the surgical intervention more quickly and with less workload and physiological effort even if the automation support is not available.

Negative findings like this are always a bit challenging to explain, and it might be argued that they just are related to the comparatively low power of the study given the limited number of participants. However, the sizes of all statistical effects involving the contrast of both experimental groups were very small. Thus, there is good reason to assume that the results of no effects of NC support on learning represent a valid finding that calls for an explanation.

The very fact that all participants were able to improve their speed and quality of surgical outcome does not seem to be surprising. These performance measures are highly dependent on the acquisition of drilling skills that were trained in both groups to a comparable degree. More remarkable is the finding that both groups also

yielded comparable improvements with respect to SA, to the protection of risk structures, and to the avoidance of complications. Because comparable studies are lacking, the underlying reasons for this result represent a matter of speculation. At least two factors might have contributed to this effect. First, the automated assistance system investigated, albeit representing the most advanced technology in the domain, still represents a comparatively low degree of automation compared to systems used in other fields of application (e.g., aviation). Instead of keeping the surgeon completely out of the loop of the navigation process, its function is limited to providing information analysis and decision support to the surgeon, who concurrently also has direct access to the “raw data.” This feature of “shared functions” might have counteracted any negative impacts on skill acquisition. More specifically, the automated blocking of inappropriate actions could have served as a “training wheel” for the participants that helped them to develop appropriate navigation and intraoperative decision-making skills. In some aspects this would correspond to approaches of error prevention training, which have been shown to improve learning efficiency even in other contexts of human-computer interaction (e.g., Carroll & Carrithers, 1984). Another related factor that might have contributed to the present findings is that the current experiments have addressed only possible consequences of NC support for an

early phase of skill acquisition. During this phase participants in the NC training group might also have been motivated not to rely completely on the automation support but to orientate themselves on the operative site. This might have been further reinforced by the fact that the model chosen for this experiment did not represent the simulation of an MIS but a surgery where the participants had a direct view in the operative site. As a consequence, they were able to directly match their perception with the automatically generated information from the NC system. It could be supposed that in the case of an MIS that provided even greater challenges for navigation, trainees might be more tempted to rely on the automated support and, thus, neglect to develop their own navigational skills. Furthermore, further research seems needed to investigate the effects of automation support on later stages of skill acquisition.

SUMMARY AND CONCLUSION

The present study provides the first comprehensive insights into the human performance consequences of advanced automated navigation support for surgeons. In summary, the results suggest that the new NC technology provides benefits in terms of increased patient safety and a release of physiological effort of surgeons. However, a clear cost effect emerged with respect to the time needed for a surgery if the system is used. If this increased time demand is not taken into consideration in the scheduling of surgeries, it can lead to a raised time pressure for surgeons, which in turn might increase the risk of errors and lower the quality of the surgical outcome.

Beyond that, only a few indications were found that the benefits of NC support might be offset by new risks or issues known from other fields of automation (i.e., increased subjective workload, increased attentional demands, and minor issues of SA). Most of these negative performance consequences do not seem to reflect principle issues related to the basic idea of NC support but seem to arise from its current features of technological implementation that lead to repeated interruptions of workflow. Improvements of the tracking technology and, in particular, a remodeling of the protection

function toward a less interruptive one might reduce these disadvantages in future systems.

ACKNOWLEDGMENTS

This research was sponsored by Research Grant MA 3759/2 / STR 696/4 provided by Deutsche Forschungsgemeinschaft.

KEY POINTS

- Navigated control represents an advanced technology of image-guided navigation support for surgeons that supports surgeons in acquisition and analysis of information and intraoperative decision making.
- The more advanced degree of automation of NC technology compared to currently used pointer-based systems might increase issues of human-automation interaction.
- The results of the first experiment reveal that NC support can reduce both the risk of intraoperative injuries and complications as well as the physiological effort of surgeons.
- Negative performance effects compared to a conventional surgery were found with respect to the time needed for the surgery, an increased subjective workload, a reduced spare capacity, and a reduced level of situation awareness.
- The results of the second experiment suggest that providing NC support already during surgical training does not affect the acquisition of surgical skills.

REFERENCES

- Boehm, B., Roetting, N., Schwenk, W., Grebe, S., & Mansmann, U. (2001). A prospective randomized trial on heart rate variability of the surgical team during laparoscopic and conventional sigmoid resection. *Archives of Surgery, 136*, 305–310.
- Carroll, J. M., & Carrithers, C. (1984). Blocking learner error states in a training-wheels system. *Human Factors, 26*, 377–389.
- Casap, N., Wexler, A., & Eliashar, R. (2008). Computerized navigation for surgery of the lower jaw: Comparison of 2 navigation systems. *Journal of Oral and Maxillofacial Surgery, 66*, 1467–1475.
- Ecke, U., Luebben, B., Maurer, J., Boor, S., & Mann, W. J. (2003). Comparison of different computer-aided surgery systems in skull base surgery. *Skull Base, 13*, 43–50.
- Endsley, M. (2000). Direct measurement of SA: Validity and use of SAGAT. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness: Analysis and measurement* (pp. 147–173). Mahwah, NJ: Lawrence Erlbaum.
- Endsley, M., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*, 381–394.

- Ferris, T., Sarter, N., & Wickens, C. D. (2010). Cockpit automation: Still struggling to catch up. In E. Salas, T. Allard, & D. Maurino (Eds.), *Human factors in aviation* (2nd ed., pp. 479–503). Amsterdam, Netherlands: Elsevier.
- Gong, J., Mohr, G., & Vézina, J. L. (2007). Endoscopic pituitary surgery with and without image guidance: An experimental comparison. *Surgical Neurology*, *67*, 572–578.
- Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, Netherlands: Elsevier.
- Jorna, P. (1992). Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological Psychology*, *34*, 237–258.
- Koele, W., Stammberger, H., Lackner, A., & Reittner, P. (2002). Image guided surgery of paranasal sinuses and anterior skull base-five year experience with the InstaTrak-System. *Rhinology*, *40*, 1–9.
- Manzey, D., Roettger, S., Bahner-Heyne, J. E., Schulze-Kissing, D., Dietz, A., Meixensberger, J., & Strauss, G. (2009). Image-guided navigation: The surgeon's perspective on performance consequences and human factors issues. *International Journal of Medical Robotics and Computer Assisted Surgery*, *5*, 297–308.
- Manzey, D., Strauss, G., Trantakis, C., Lueth, T., Roettger, S., Bahner-Heyne, J. E., Dietz, A., & Meixensberger, J. (2009). Automation in surgery: A systematic approach. *Surgical Technology International*, *18*, 37–45.
- Metson, R. B., Consenza, M. J., Cunningham, M. J., & Randolph, G. W. (2000). Physician experience with an optical image guidance system for sinus surgery. *Laryngoscope*, *110*, 972–976.
- O'Donnell, R.D., & Eggemeier, F.T. (1986). Workload assessment methodology. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and performance*. Vol. 2: Cognitive Processes and Performance (pp. 42-1–42-49). New York: Wiley.
- Parasuraman, R., Sheridan, T., & Wickens, C. D. (2000). A model of types and levels of human interaction with automation. *IEEE Transactions of Systems, Man, and Cybernetics. Part A—Systems and Humans*, *30*, 286–297.
- Reijnders, K., Coppes, M. H., van Hulzen, A. L. J., Gravendeel, J. P., van Ginkel, R. J., & Hoekstra, H. J. (2007). Image guided surgery: New technology for surgery of soft tissue and bone sarcomas. *European Journal of Surgical Oncology*, *33*, 390–398.
- Strauss, G., Koulechov, K., Hofer, M., Dittrich, E., Grunert, R., Moeckel, H., Muller, E., Korb, W., Trantakis, C., Schulz, T., Meixensberger, J., Dietz, A., & Lueth, T. (2007). The navigation-controlled drill in temporal bone surgery: A feasibility study. *Laryngoscope*, *117*, 434–441.
- Strauss, G., Koulechov, K., Richter, R., Trantakis, C., Dietz, A., & Lueth, T. (2005). Navigated control in functional endoscopic sinus surgery. *International Journal of Medical Robotics and Computer Assisted Surgery*, *1*, 31–41.
- Wickens, C. D., Huiyang, L., Santamaria, A., Sebok, A., & Sarter, N. B. (2010). Stages and levels of automation: An integrated meta-analysis. In *Proceedings of the Human Factors and Ergonomics Society 54th annual meeting* (pp. 389–393). Santa Monica, CA: Human Factors and Ergonomics Society.

Dietrich Manzey is a university professor of work, engineering and organizational psychology at the Institute of Psychology and Ergonomics, Berlin Institute of Technology, Germany. He received his

PhD in experimental psychology from the University of Kiel, Germany, in 1988 and his habilitation in psychology from the University of Marburg, Germany, in 1999.

Maria Luz is a PhD candidate at the Department of Psychology and Ergonomics of the Berlin Institute of Technology. She obtained a diploma in psychology from the Berlin Institute of Technology in 2008. As project scientist, she has been responsible for conducting the experiments presented in the current publication.

Stefan Mueller is a third-year trainee in ENT at University Hospital Leipzig, Germany. He finished his basic medical training in 2008 at the University of Leipzig. As project scientist, he was involved in parts of the research presented in the current publication.

Andreas Dietz is university professor for ENT and director of the ENT department of the University Hospital Leipzig, Germany. He is executive board member of the Innovation Center for Computer Assisted Surgery at the University of Leipzig and director of the IRDC-Academy. He received his MD from the University of Heidelberg, Germany, in 1985 and his habilitation in neurosurgery from the same university in 2000.

Juergen Meixensberger is professor of neurosurgery at the medical faculty of the University of Leipzig, Germany, and head of the department of neurosurgery at the University Hospital Leipzig. He has been director of the board of the Innovation Center for Computer Assisted Surgery at the University of Leipzig since 2005. He received his MD from the University of Wuerzburg, Germany, in 1985 and his habilitation in neurosurgery from the same university in 1993.

Gero Strauss is a professor of ENT surgery at the medical faculty of the University of Leipzig, Germany. Since 2009 he has been head of the International Development Centre for Surgical Technology in Leipzig. He received his MD from the University of Leipzig, Germany, in 1999 and his habilitation in ENT from the same university in 2006.

Date received: May 31, 2011

Date accepted: September 2, 2011