THE IMPACT OF TECHNOLOGY ON PHYSICIAN COGNITION AND PERFORMANCE

David D. Woods, PhD, Richard I. Cook, MD, and Charles E. Billings, MD

Woods DD, Cook RI, Billings CE. The impact of technology on physician cognition and performance.
J Clin Monit 1995;11:5–8

PHYSICIAN PERFORMANCE AND NEW TECHNOLOGY

Anesthesiologists and intensive care specialists are struggling with the question of how to evaluate the impact of new information technology, especially when devices promise to benefit patient safety [1]. The value of many innovations in technology for anesthesia and critical care medicine, like automated record keeping, pulse oximetry, or integrated patient monitors, lies in their impact on the cognition of medical personnel—processes such as attention, workload, and reasoning. In evaluating the impact of new technology, the unit of analysis cannot be the new box or capability in isolation; rather, it must be the larger information processing system that includes the physician's cognitive processes, the set of interacting physicians and medical personnel for that setting, and the set of tools that influence the information processing of these people [2]. It is the person–machine system that needs to be evaluated, not simply an electronic box. The work by Loeb investigating the impact of record-keeping technology on anesthesiologist workload and attention is one example of the direction in which technology assessment needs to move. He has taken one approach for investigating human performance, the dual task class of methods, and tried to adapt it to investigate anesthesiologist performance during actual cases. In this class of methods, the investigators introduce an additional, or secondary, task that can be related or unrelated to the primary task. Variations in performance on the secondary task provide one indication of the workload or mental resources devoted to the primary task.

HUMAN-CENTERED AUTOMATION

While the goal of the Loeb study is to assess the impact of automatic record keeping, he has actually manipulated whether a human scribe is present or not to assist the anesthesiologist responsible for the case. The scribe functioned as a team player: responsive, directable, intelligent, nonintrusive. However, research on human interaction with automation in many domains, including aviation and anesthesiology [3–5], has shown that automated systems often fail to function as team players. To summarize the data, automated systems that are strong, silent, clumsy, and difficult to direct are not team players. Automated systems are

From the Cognitive Systems Engineering Laboratory and the Department of Anesthesiology, The Ohio State University, Columbus, OH.

Received Jul 19, 1994, and in revised form Aug 29, 1994 and Sep 15, 1994. Accepted for publication Sep 15, 1994.

Address correspondence to Dr Woods, Cognitive Systems Engineering Laboratory, The Ohio State University, 1971 Neil Avenue, Columbus, OH 43210.

1. *strong* when they can act autonomously;
2. *silent* when they provide poor feedback about their activities and intentions;
3. *clumsy* when they interrupt their human partners during high workload and high-criticality periods, or add new mental burdens during these high-tempo periods; and
4. *difficult to direct* when it is costly for the human supervisor to instruct the automation about how to change as circumstances change.

Systems with these characteristics create new problems for their human partners and new forms of system failure.

There is a great deal of work going on, especially in the aerospace arena, about how to develop automated systems that do function as team players—what is now called human-centered automation [3].

VARIETIES OF ATTENTION: SUSTAINED ATTENTION

The Loeb study attempts to examine the impact of record-keeping technology on anesthesiologist "vigilance." Anesthesia practitioners understand "vigilance" as part of the commonsense notion that it is sometimes difficult to track or attend to patient state closely. Anesthesiologists are aware of this threat in situations where they are very busy with multiple tasks, such as induction and emergence, and in situations where there is very little for them to do for long periods of time. However, for cognitive psychologists, these two situations involve very different kinds of attentional and other mental processes [6].

It is important to keep the different varieties of attention in mind if we are to make sense of the impact of new technology on physician performance. *Sustained attention* (what cognitive psychologists call vigilance) refers to "a state of readiness to detect and respond to certain small changes occurring at random time intervals in the environment" [6: p244]. The basic empirical finding is called the vigilance decrement—the rate of correct detection of infrequently occurring signals declines over time, although a number of factors influence this phenomenon [6,7]. In anesthesia, such low-tempo, low-signal rate periods generally occur only during quiescent periods of the maintenance phase.

During high-tempo phases, such as induction and emergence, different attentional processes are at work. Anesthesiologists switch attention between different streams of activity and divide their attention across different data channels. The processes of *divided attention* and *attention switching* in busy multitask situations are fundamentally different from the processes involved in sustained attention during quiescent, low-signal rate periods [8]. In the former, we are concerned with *overload* and workload management (timesharing, shedding loads, task priorities, shifting workload over time); in the latter, we are concerned with *underload* and the early detection of the onset of a problem.

Loeb found no effects of record keeping during the maintenance phase. Performance on the secondary task was good (only 1% errors and relatively short reaction times on the secondary task) and was unaffected by the scribe. In the logic of dual-task methods, when the secondary task and the primary task can be performed successfully in parallel, it could indicate that the secondary task was too easy (a ceiling effect), that the two tasks do not compete for the same limited resources, or that the primary task is not very demanding relative to the visual, cognitive, or manual resources needed to perform the secondary task.

What could these results tell us about the possible effects of record keeping on sustained attention? Sustained attention during the maintenance phase is a potential benefit, according to some advocates for automatic record keeping. Critics, on the other hand, claim that automated record keeping harms sustained attention during low-workload, uneventful periods. For example, when there are few other tasks (underload), record keeping directs attention to patient vital signs, which are themselves indicators of developing problems, and keeps physicians mentally involved in the case. This activity may help maintain what human factors specialists call *situation awareness* [9]. The signals one wishes to detect are related to the very vital signs being checked and recorded. Delegating vital sign recording can affect one's ability to pick up early signs of problems. Measuring the effect of record keeping on sustained attention through an unrelated secondary task ignores the critical role of situation awareness in anesthetic practice [10].

Arousal level, a physiologic state related to the readiness to respond to new stimuli, is a key factor that affects performance at sustained attention tasks. Events that occur during maintenance are important, in part, because they affect arousal level. If other activities maintain arousal levels, one would not expect to see a vigilance decrement on the detection task. Data about the context of other activities and their mental demands is critical if one intends to study sustained attention in naturally occurring situations. Unfortunately, the dual-task method used here did not include this finer grain data.

Thus, on methodological and conceptual grounds, the Loeb study leaves questions surrounding the impact of automatic record keeping on *sustained attention* unexamined. While questions about how automatic record keeping affects sustained attention remain open, there are many other factors present in the current practice of anesthesia that are known to have large effects on arousal and, therefore, on sustained attention (e.g., sleep loss, shift schedule [11]). Research on sustained attention, if applied to anesthesiology, could identify changes with large potential effects on performance and safety.

HIGH-TEMPO, MULTI-TASK SITUATIONS AND WORKLOAD

The dual-task technique used in this study typically is used for studying workload, timesharing, and task switching,* rather than sustained attention. The characteristic that defines a sustained attention task is that it is uneventful, i.e., *all* of the activities carried out by the participant are relatively uneventful and nonarousing, not just the signal rate on the secondary task. This criterion is clearly not met for induction and emergence.

Loeb finds a difference between the scribe and no scribe conditions during emergence. During induction, performance on the secondary task dropped, even with a scribe to perform record-keeping tasks (compare missed signal rate and reaction time in the Induction/Scribe condition to performance during the maintenance phase and to performance in the Emergence/Scribe condition; Loeb, Table 4).† In other words, the secondary task method shows that the physical, visual, and cognitive activities of anesthesiologists are many and intense during induction.

One argument advanced in favor of introducing automated record keeping is that it off loads the anesthesiologist during high-tempo periods. The Loeb data indicate that these phases of practice are, indeed, intense and active (consistent with practitioner commonsense and other studies of anesthesiologist activities [12]). One, then, could interpret this study to imply that automatic record keeping, if it *lowers* physician workload during these busy periods, will be beneficial.

However, the Loeb study does not delve deeply enough into the workload issue or the impact of the technology of record keeping on workload to provide the kind of results needed to assess the impact of new technology on practitioner performance. For example, in multitask situations, especially when they are as complex as anesthetic practice, workload competition and workload management strategies are critical factors to measure and trace throughout each case. The coarse grain data reported indicate that as workload associated with primary task increases, the priority of the secondary task and the resources devoted to it drop. In other words, people actively allocate their limited visual, motor, and cognitive resources to higher-priority tasks when multiple tasks impose competition for these resources.

The Loeb study does not help us isolate the source of workload bottlenecks. Is it competition for limited motor channels, e.g., they know that a signal occurred, but are unable to respond because of ongoing tasks that occupy their hands? Is the bottleneck limited visual resources, e.g., the visual channel is devoted entirely to ongoing or high-priority tasks, leaving no spare visual capacity to monitor the secondary data channels (a tunnel vision effect)? Does the bottleneck occur at the level of central cognitive processing, where limited channel capacity requires some kind of attention switching or divided attention to process all of the relevant stimuli across all of the relevant tasks? Or, is it all of these in different combinations and at different times? These questions illustrate that workload research is concerned primarily with how people actively manage their internal and external resources to cope with different workload bottlenecks.

Measuring strategies for workload management has been recognized as particularly important because the primary task—in this case, anesthetic practice—is actually a complex of multiple coordinated tasks that provides practitioners with multiple degrees of freedom. Practitioners shift resources to higher-priority tasks, shed lower-priority tasks, seek additional resources (e.g., other people), and integrate these additional resources smoothly into a team effort, as workload and tempo increase in an evolving situation [13].

CONCLUSIONS

We have mentioned a few methodological points to illustrate that settling issues about the impact of technology on human performance requires a finer grain of analysis of human activity than is provided by the Loeb study. The methods for assessing the impact of technology on human performance are rich in possibilities and subtle in execution. There are no standard methods that can be applied directly; rather, each class of technique requires the investigator actively to design and adapt the technique to the conditions of the field of practice and the role of technology. As the anesthesia commu-

*See Wickens, 1992, chapter 9, for one standard introduction to workload and methods for studying workload.

†The difference between missed signal rate in the Induction/Scribe condition and the Emergence/Scribe condition did not reach statistical significance at the 0.05 level, although the statistical power of comparisons of the emergence and induction data is impaired by low N relative to high variability in the data.

nity struggles with questions about the impact of new technology, it will find, as have others (e.g., aviation), that a fine grain and cognitively based analysis will be one necessary tool.

Loeb's study is important because it addresses the need to resolve the ongoing controversies about the impact of automatic record keeping and similar debates. Very often, advocates of new technology claim that, because practitioners are busy during some phase of activity, automating a subtask will automatically be useful. Loeb attempts to test the claim: Does additional help in the form of a scribe to perform one subtask during high-workload phases improve performance, as measured by a nonrelated secondary task? Advocates of automatic record keeping will argue that if a scribe is useful, then an automatic machine as scribe will be useful. However, studies of the impact of automation on human performance show that following this logic frequently produces automated devices that create new cognitive burdens *especially during high-tempo periods and nonroutine cases* [5].

It is not controversial that having a good team member would be useful during busy times, such as induction and emergence. It is also clear from the growing database about people interacting with automated systems that automated devices frequently fail to function as team players and, as a result, create new problems and new errors. The important question is to determine what are the characteristics of good team players, either human or machine. Perhaps we should study how people manage evolving high-tempo, high-workload situations; how additional people are integrated smoothly into the team to handle busy or critical periods; how people shift resources to high-priority tasks and away from low-priority tasks as a case escalates in difficulty; how work is divided up usefully among different team members in different situations. One skill in domains like anesthesia is the ability to prioritize resources and focus on the critical issues when there are multiple competing demands. Another skill is the ability to coordinate team resources, especially in crisis situations [14]. Studying and modeling how people cooperate and coordinate their activity—how they generally succeed, but sometimes fail to do this well—is a strategy that can point the way to the development of automated systems that *will* function as team players.

The success or failure of new technology that is not directly therapeutic depends on how using a device in situ influences physician cognitive processes and cooperative activity. Loeb has recognized this and has tried to examine some of the relevant cognitive factors, workload and attention. It is important to continue to move down this path in anesthesia and other medical specialties, using research techniques that provide a finer grain view of physician–technology interaction and using concepts about human-centered automation as a guide.

REFERENCES

1. Orkin FK, Cohen MM, Duncan PG. The quest for meaningful outcomes. Anesthesiology 1993;78:423–426
2. Norman DA. The psychology of everyday things. New York: Basic Books, 1988
3. Billings CE. Human-centered aircraft automation philosophy. (NASA technical memorandum 103885). Moffett Field, CA: NASA-Ames Research Center, 1991
4. Cook RI, Woods DD, Howie MB. Unintentional delivery of vasoactive drugs with an electromechanical infusion device. J Cardiothorac Vasc Anesth 1992;6:1–7
5. Woods DD, Johannesen L, Cook RI, Sarter N. Behind human error: Cognitive systems, computers and hindsight. Dayton, OH: Crew Systems Ergonomic Information and Analysis Center, 1994
6. Parasuraman R, Davies DR. Varieties of attention. London: Academic Press, 1984
7. Wickens C. Engineering psychology and human performance, ed 2. New York: Harper Collins, 1992
8. Moray N. Monitoring behavior and supervisory control. In: Boff K, Kaufman L, Beattie J, eds. Handbook of human perception and performance. New York: Wiley, 1986
9. Gilson R, ed. Special issue on situation awareness. Human factors. (in press)
10. Gaba DM, Howard SK, Smith SD. Situation awareness in anesthesiology. Hum Factors. (in press)
11. Weinger MB, Englund CE. Ergonomic and human factors affecting anesthetic vigilance and monitoring performance in the operating room environment. Anesthesiology 1990;73:995–1002
12. McDonald JS, Dzwonczyk RR. A time and motion study of the anesthetist's intraoperative time. Br J Anesthes 1988;61:738–742
13. Cook RI, Woods DD. Operating at the "Sharp end:" The complexity of human error. In: Bogner MS, ed. Human error in medicine. Hillsdale, NJ: Lawrence Erlbaum, 1994
14. Gaba DM, Howard SK, Fish KJ. Crisis management in anesthesiology. New York: Churchill-Livingstone, 1994