Improving Clinical Communication:
A View from Psychology

JULIE PARKER, MSc, ENRICO COIERA, MB BS, PhD

Abstract Recent research has studied the communication behaviors of clinical hospital workers and observed a tendency for these workers to use communication behaviors that were often inefficient. Workers were observed to favor synchronous forms of communication, such as telephone calls and chance face-to-face meetings with colleagues, even when these channels were not effective. Synchronous communication also contributes to a highly interruptive working environment, increasing the potential for clinical errors to be made. This paper reviews these findings from a cognitive psychological perspective, focusing on current understandings of how human memory functions and on the potential consequences of interruptions on the ability to work effectively. It concludes by discussing possible communication technology interventions that could be introduced to improve the clinical communication environment and suggests directions for future research.

Communication between health care workers accounts for the major part of the information flow in health care, and growing evidence indicates that errors in communication give rise to substantial clinical morbidity and mortality. Covell has reported that about 50 percent of information requests by clinicians in clinic were met by colleagues rather than by documented sources. Safran et al. reviewed the information transactions in a hospital with a mature computer-based record system and still found that about 50 percent of information transactions occurred face-to-face between colleagues, with e-mail and voicemail accounting for about another quarter of the total.

Given the importance of interpersonal communication as a means of information exchange, it is not surprising that communication failures are a large contributor to adverse clinical events and outcomes. In a retrospective review of 16,000 in-hospital deaths, communication errors were found to be the leading cause, twice as frequent as errors due to inadequate clinical skill. Furthermore, in a study of primary care physicians, about 50 percent of all detected adverse events were associated with communication difficulties.

The causes and remedies of poor communication in the health care system are consequently of critical interest to the study of informatics. Much work has examined the dynamics of communication between individual health care providers and patients, and this body of work can help optimize that interaction. Similarly, there are bodies of work on nurse and physician perceptions of, and satisfaction with, their communication and on physician teaching behaviors.

However, there is a paucity of detailed data about the effect of communication behaviors on overall organizational efficiency and effectiveness in health care. Some studies have investigated the use of communication technologies in health care settings. One study in a nursing home found that the number of telephone calls nurses received constituted a significant communication burden and that most calls were routine or informative only. Introduction of a voicemail system allowed the vast majority of calls to be transferred to that medium, saving both time and unnecessary extra communication.

Spurck et al. found that nurses and physicians felt that their hospital’s telecommunication system was more effective when nurses were given portable...
A large number of factors might influence communication behavior in organizations, including the nature of the available communication infrastructure, the nature of the work undertaken, and the practices that are routinely applied in the organization by individuals. A limiting factor in any communication analysis is the cognitive capacity of individuals to undertake their work, and in studies of high cognitive workload, it has often been shown that error or inefficiency results when cognitive limits are exceeded.23

In this paper, we focus on the cognitive limits of individuals and how these might explain the inefficient and interruptive clinical environment described in the Coiera and Tombs study. We do so by drawing together the available clinical findings with empirical research from cognitive psychology. In particular, we are interested in human memory, and we speculate that the burden on memory that results from a highly pressured working environment is a significant contributor to the interruptive behaviors observed.

By adopting this approach, we hope to achieve several goals. Our first goal is to provide a principled framework in which existing findings about communication patterns in health care organizations can be understood and can direct further research. Our second goal is to promote an understanding, based on empirical psychological research, of the sorts of problems in real working environments that can, potentially, be addressed by the appropriate application of new communication technologies.

In addition, in discussing the social environment into which new communication technologies are introduced, we aim to caution against an approach that does not make the social environment the object of study. Enthusiastic applications of new technologies do not always have the consequences expected of them. This is not least because they are always introduced into a social environment, and this often acts strongly to modify the ways in which their capabilities are harnessed.

Finally, we hope to stimulate psychologically and sociologically informed research into the effective application of new communication technologies to hospitals.

In the next sections, we introduce relevant concepts from cognitive psychology and then relate these to the existing clinical communication findings. In particular, we discuss specific types of memory error that are likely to be seen in interrupt-driven environments. We then use this model to discuss the likely consequences of introducing different communication technologies into this environment.

**Clinical Communication Patterns**

Coiera and Tombs22 observed the communication patterns of eight physicians and two nurses in an English district general hospital. The available channels of communication for these highly mobile professionals consisted of face-to-face meetings, both impromptu and planned; desktop telephones; paging; written notes for colleagues in patient notes; notes at ward desks; notice boards; and pigeon holes for personal memos. Voicemail and e-mail were not supported, and mobile telephones were not used. The subjects in this study, like those in others, made little or no use of more formal sources of information, with the exception of data from the medical record.

One of the communication behaviors observed was a bias toward interruptive communication. Interruptive, or synchronous, communication methods require the simultaneous interaction of the two parties to the communication: the telephone and face-to-face discussions are two such methods. In contrast, an asynchronous method, such as writing a note or leaving a voicemail or answerphone message, allows the recipient to deal with the communication at a time of his or her choosing.

The authors reported that staff showed strong preferences for making telephone calls and for taking advantage of chance face-to-face meetings with colleagues. There was little evidence that staff’s own experience of interruptions encouraged them to adopt more “considerate” communication methods when contacting their colleagues. The authors also observed that the reliance on synchronous methods can be a source of inefficiency for the person attempting to communicate. Recipients may be unavailable or occupied, or the communication channel may be busy, and tasks remain undone until these conditions change.

Coiera and Tombs’ observations are of interest from a psychological perspective, for two reasons. First, cog-
Human Memory

A knowledge of the way in which human memory is believed to function is key to understanding the probable effects of working in an interrupt-driven environment. It is also key to understanding the requirements of technologies that might be introduced to support those in such environments.

The functioning of human memory has been the subject of empirical study in cognitive psychology for many decades. A basic division of memory into short- and long-term functional components provides the cornerstone on which the working of memory is understood.

Our knowledge is believed to be stored in notional "repositories," known as long-term memory. Remembering medical facts, significant dates, events from childhood, and how to drive a car all draw on long-term memory. Much of the time, most of our fund of knowledge and meaning in long-term memory is inactive; that is, it is not the current focus of attention.

Working memory is believed to be the activated state of information held in memory. It may be equated with the component of memory we associate with attention. Working memory actively processes information, whether the information is sensory input (for example, sounds, sensations, or sights currently being experienced) or items from long-term memory. When carrying out a mental calculation, making a plan to do something, recalling a phone number, or writing a note, it is working memory that allows the various "pieces" of information to be attended to, integrated, and manipulated.

Working memory has some interesting characteristics. In particular, it is extremely limited in its capabilities. The number of items—such as thoughts, sensory impressions, and plans—that can be held in working memory is very small. Furthermore, items in working memory are easily disturbed by each other. This is particularly the case when someone is distracted from thinking about one task by a new one that supervenes. An intention to carry out an act can be forgotten by the intrusion of another plan, even when only ten seconds separates the intention from the intrusion.

Working memory is also severely limited in duration. Without conscious attention to plans or other items in working memory, the accurate memory persists no longer than about 20 seconds. This decay can be overcome by acts of conscious self-reminder: these serve to refresh and re-prioritize the items in working memory. If there are competing demands on working memory, however, such as executing another task or communicating with a co-worker, then such rehearsal of intention becomes impossible, with the same effect—a plan may be forgotten. Considerable empirical evidence shows the powerful negative effects of both interference and diversion of attention on working memory.

The characteristics of the process of forgetting are not random; in particular, two serial position effects, known as the primacy and the recency effects, are known to affect retention in systematic ways. The primacy effect describes the tendency toward superior recall of items that have resided longest in working memory; the recency effect describes a similar superiority of items most recently added to working memory. The combined result of these two effects is poor retention of items in the middle of the "mental list." In addition, a distracting task before recall can obliterate the recency effect but not affect the primacy effect.

A further distinction may be drawn in long-term memory between retrospective memory and prospective memory. Retrospective memory refers to the factual, autobiographical, and instructional ("how-to") knowledge we possess. Prospective memory, in contrast, is the memory for a future act, or the memory to remember to do something. It necessarily draws on retrospective memory and entails complex planning and coordination. Like retrospective memory, prospective memory relies on working memory for its processing work. For example, in remembering that you need to contact someone later in the day, you draw on your retrospective memory in deciding how to make the contact.

Failures of Working Memory in an Interrupt-driven Environment

Those who work in an interrupt-driven environment are likely to suffer failures of working memory. As
interruptions occur, interfering with the active cognitive rehearsal of what is to be done and generating new tasks for their recipients, prospective plans may be partly or fully forgotten. They may or may not be recalled subsequently, depending on appropriate cues for recall. The greater the number of such plans, the more the effect will be exacerbated. Since planning for prospective activities is an activity of working memory, which is limited in the number of distinct items it can retain, then forcing more items into working memory will cause some plans to be displaced from it, perhaps to be forgotten. The tasks that should suffer most from interruptions are all but the oldest tasks in a nurse’s or doctor’s mental “to do” list.

Coiera and Tombs proposed that immediate acknowledgement of a message seemed to be needed in such an interruptive environment to permit workers to complete a task. Good psychological reasons may be advanced for this and for the “selfish” behavior observed. When a worker’s working memory is operating to capacity, the highest priority is likely to become the reduction of this mental burden by completing the tasks that are consuming memory resources. Reinforcing this, it is probable that when the consequences of errors can be so serious, it is difficult for a doctor or nurse to feel that he or she has truly “handed over” responsibility for a task without an explicit acknowledgement from the recipient. Existing asynchronous methods of communication, such as hand-written notes, voicemail, and e-mail do not easily or routinely offer this feature. The consequence may be that the task cannot be removed from working memory.

Errors of Reality Monitoring and Temporal Association

Two further types of memory error may be made more likely by interruptive working environments.

“Reality monitoring” is the ability to discriminate between “true” and “false” memories. True memories are memories of events, objects, and actions that really occurred or were experienced. False memories derive from the imagination that something occurred or was experienced. Confusion between the two is a quite ordinary occurrence. It may be seen as a consequence of the way in which memory functions as creatively constructed representations rather than as simple records of sensory information. Childhood memories, for example, are often a composite of the original experience interwoven with what other people said about it, as well as the embroideries that are overlaid on the original fabric as the experience is mentally revisited over time.

Reality monitoring errors may be either omissive or repetitive. If the memory of an intention to act is confused as a memory of having acted, an error of omission will occur. If the memory of the performed act is mistaken as the memory of a plan to act, the error will be one of repetition. For example, if an intention to take a dose of medicine is confused as the action of having done so, a dose will be missed. If the action is mistaken as the intention, an extra dose will be taken.

The second type of errors are those of “temporal association.” These errors are thought to be strongly associated with routine and frequently performed actions. The more repetitive and routine an action is, the more difficult it can be to decide whether a memory of the action is today’s or yesterday’s memory. For example, did I really clean my teeth this morning or am I remembering cleaning my teeth yesterday morning?

Both reality monitoring and temporal association errors can, therefore, result in the omission or repetition of tasks. Given that they are fundamental tendencies of the memory system, we may expect them to be additional sources of memory failure in an interruptive working environment.

These errors will be especially likely under particular circumstances. Outstanding tasks that are simple, routine, and repetitive are particularly vulnerable. Under conditions of high work pressure, when there is insufficient time to perform a reality check, errors may be more readily accepted. Furthermore, since junior doctors, like nurses, are required to undertake many more routine and repetitive tasks than are senior doctors (for example, ordering laboratory tests and securing ward beds for new admissions), it is probable that they are more vulnerable to failures of reality monitoring and temporal association than are senior doctors.

The Effects of Expertise on Memory

There is a large body of research on the effects of skill acquisition on problem solving and memory, which indicates that with experience, some components of tasks can be performed automatically. They are sufficiently well learned that once set in train they do not rely on working memory to be enacted, thereby freeing components of working memory for alternative use.

This means that the probability of a memory error is greater for less experienced members of staff. Junior medical officers, for example, are novices in the practice of any given speciality of hospital medicine. As
the tasks associated with each level of seniority differ, this will also be true of newly appointed specialist medical staff.

Since experts need to rely less on general attentional resources than do novices, it is probable that more experienced doctors at every level of the hierarchy, as well as nurses and other health care professionals, will suffer less from the effects of interruptions in the performance of specific tasks than will their less experienced colleagues. In the Coiera and Tombs study, the greatest communication burden actually fell on the most junior staff, whom one would expect to be the group most likely to make errors in such circumstances.

Discussion

New Communication Technologies

Coiera and Tombs observed two main contributors to the interrupt-driven nature of the hospital environment—the behaviors of hospital workers and the characteristics of the work itself. The work was highly mobile, conducted in multidisciplinary teams, and involved many simultaneous tasks and responsibilities. The existing communication environment in the hospital in which the study took place relied largely on synchronous communications and did not support mobility. Individuals responded to these factors by favoring synchronous communications even when they were not necessary or even productive, thus increasing the interruptiveness of the working environment.

The previous exploration of memory functioning demonstrates that working in a busy and interrupt-driven environment can over-extend the capabilities of the human cognitive system. In such an environment, there is a premium on immediate task completion and reliance on synchronous communication. Such behaviors permit ambiguities and uncertainties to be dealt with on the spot and can thus be construed as reasonable adaptations to working in such an environment.

Coiera and Tombs suggested technologies that could reduce the interrupt-driven nature of hospital work. Portable telephones could support mobility, and asynchronous communications technologies such as voicemail and e-mail, with acknowledgements, could fulfill the initiator’s need for immediate task completion without generating an interrupt for the recipient.

The introduction of new technologies seldom, however, permits such straightforward predictions to be made. Social influence approaches to studying information technology have shown that the use of new technologies is not predicted solely by the characteristics of the technologies themselves. Instead, the human environment into which they are introduced is critical in shaping how their capabilities are actually used. Attitudes of key individuals and organizational norms are among the important factors that shape adoption and use.29

There are thus many uncertainties about the actual use of technologies. Correspondingly, this is a fertile area for research, no less in the medical field than in the world of the office. In the following sections, some of the difficulties with these simple predictions about introducing new technologies are explored, to underline the difficulties that unexpectedly arise when apparently simple solutions are introduced into complex human work environments.

Increasing Interruption and the Bias to Synchronous Communication

What might be achieved by providing staff with mobile phones? The general effect of introducing mobile phones is to make individuals more available.20 Since failure to reach individuals results in further attempts to make contact, mobile phones would be expected to reduce the call failure rate and thus the overall call traffic for the organization. For the individual caller, it could mean a reduction in the number of outstanding tasks, as call recipients are easier to reach. This means that the corresponding number of items in working memory that are associated with a task no longer compete for attention.

For the recipient, however, the picture is less clear. If providing mobile phones only reduces the number of call re-tries, we would expect interruption levels to be unchanged. However, if the ease of contacting individuals has the effect of creating additional calls for conversations that would not have occurred previously, then the overall interruption level would increase. We would predict in these circumstances that each new call generates an interruption, and an addition to working memory.

The synchronous bias hypothesis22 predicts that individuals preferentially use synchronous communication channels. Mobile telephones, by making synchronous communication easier, would thus be predicted to result in new calls being made and consequently would result in an increase in the overall interruption rate for individuals.

In such circumstances, new synchronous technologies therefore would not, on their own, resolve the practical or the cognitive difficulties faced by those in interruptive working environments.
Asynchronous Messaging

What might the provision of asynchronous technologies mean? The technology would permit the message sender to achieve task completion independently of the recipient’s location and current activity. The recipient may choose a convenient time to consult and act on his or her messages.

The cognitive benefits of voicemail and e-mail for both message senders and recipients could be substantial. For callers, independent completion of communication tasks reduces the number of pending tasks in working memory. For call recipients, the receipt of fewer interruptive calls would be likely to contribute to greater chunks of uninterrupted time and greater ability to rehearse and recall existing outstanding tasks. It would allow completion of more tasks, fewer errors in task completion, and fewer forgotten tasks. These probable benefits from moving some communication tasks from synchronous to asynchronous channels would result from a decreased incidence in the factors contributing to memory errors, such as distraction, interference, and new involuntary additions to prospective memory.

There is much to be learned about the degree to which and circumstances under which callers might choose to employ asynchronous channels. While call recipients would probably choose to deal with their calls at one time, the same is not likely of callers. Since each undone communication task remains an item in prospective memory, it is instead desirable to carry out communication tasks as the need for them occurs. Tasks that do not require immediate acknowledgement or completion lend themselves to asynchronous methods of communication.

Under what circumstances might the employment of an asynchronous method form the sender’s first preference, rather than the last resort, when attempts at synchronous communication have failed? Might callers choose their communication methods on the basis of the demands of the task, and could careful design of the technologies encourage a shift to task-based use of communications?

The Effects of Social Influence on Technology Use

There are a cluster of very interesting questions relating to the use of communication technologies and status. How might the choice and use of synchronous and asynchronous communication be affected (if at all) by an individual’s status in the hospital hierarchy? Might the greater certainty of connection with junior staff encourage more “selfish” behavior by those higher in the hierarchy? That is, for routine communication, might more senior members of a team feel more free to interrupt junior members with synchronous communication and, conversely, would more junior members, more reluctant to interrupt, tend to use asynchronous voice messaging or e-mail to communicate with senior staff?

As research with office workers indicates, it is probable that staff would influence one another’s adoption and patterns of use and that local norms might evolve. Since communication technologies are essentially shared tools, the degree to which they may be fully exploited is contingent on all parties to a communication being prepared to use, and feeling satisfied with their ability to use, the capabilities provided. If, for example, one member of a team were reluctant to access text messages, the behavior of other members of the team would either have to evolve to accommodate that antipathy or to exert influence to convince that team member to behave differently. Thus, it is probable that differences in use of available technologies might be seen from team to team, shaped by key individuals.

Effects on the Nature of Conversations

Coiera and Tombs indicated that chance face-to-face meetings were an important medium of communication. One of the reasons that such meetings were so eagerly seized on was precisely the difficulties the study participants had in either locating colleagues or setting up synchronous conversations. These meetings often provided the first opportunity one colleague had to confer with another on a particular matter and sometimes substituted for the failed attempts at earlier conversations.

We might expect that the use of voicemail and e-mail would cause some changes to the content of opportunistic exchanges between colleagues. Instead of communication of the nature “I need to speak to you about X,” we could expect that opportunistic exchanges might instead be more of the nature “Did you read/receive/act on my voice/text message to you about Y?” or “Thanks for your voice message about Z; I’ll see it this afternoon.” In other words, instead of representing the first opportunity one colleague has to communicate with another, opportunistic meetings might become opportunities to confirm earlier communications and, perhaps, elaborate on them.

Message Acknowledgement

Coiera and Tombs suggested that the need for acknowledgement of receipt of a message was one of
the drivers behind the preference for synchronous communication. The cognitive reasons for this preference were discussed above: without confidence that the receiver has taken over the task, it remains an unfinished task in the caller’s working memory. Acknowledgments could, perhaps, be required for different purposes. Has the message arrived safely in the recipient’s “in” tray? Has the recipient listened to or seen the message yet? The message sender might require different types of acknowledgment, to be able to feel that the communicated task has truly been delegated and that he or she is thus able to “remove” that task from working memory.

Additional interesting questions are raised by research in the office on acknowledgments mediated by e-mail. In some circumstances, agreements mediated by e-mail were not viewed by recipients as equivalent in strength or reliability to face-to-face agreements on a course of action. Co-workers wished to look each other “in the eye” when negotiating and agreeing on commitments. Might this also be true in the hospital environment?

**Future Research**

We have hypothesized that in some cases, over-dependence on synchronous channels of communication may come about because these conditions result in excessive burdens on memory. At face value, synchronous communication may seem to provide the best way to complete some tasks and reduce unfinished tasks held in working memory; in fact, difficulties with this channel, ranging from unanswered calls to not knowing the location of a colleague, often mean that many inefficiencies are introduced into the process.

These are hypotheses that require further examination in the field. Studying cognitive phenomena in a naturalistic setting is a method of enquiry that is increasing in popularity and is increasingly seen to complement the older laboratory-based method of enquiry, to the benefit of both. Laboratory studies, with their strict experimental controls, enable cause-and-effect relationships to be postulated and examined and theory to be developed. However, laboratory experiments are often criticized as unrepresentative of real life. Field studies, on the other hand, lack this control and therefore cannot demonstrate cause-and-effect relationships. This does not mean, however, that laboratory-based theory cannot be applied to the field and used to shape and guide enquiry.

Given the complexity of this field of study, we propose a program of research to address its many facets. At least three types of approach are required. First, future studies should be designed specifically to examine memory errors in the hospital setting. Structured observation techniques could be used to focus on the occurrence of memory errors, and interviews with study participants could draw on their own accounts of observed data. While self-accounts do not offer proofs of theory, they enrich and inform the observer’s understanding; combined observation and interviews are the basis of ethnographic approaches.

The same approach is required to further understand communication behaviors. Furthermore, semi-experimental studies may bridge the need to understand behaviors in a working environment while executing controlled experiments. For example, artificially structured memory tasks could be given to staff to carry out during their routine work.

Second, and informed by the ethnographic approach, quasi-experimental studies are needed to assess the consequences of introducing new communication technologies. For example, asynchronous technologies may be introduced to a mobile and distributed team. Studies of relevant behaviors using before and after comparisons—as well as comparisons over the same period with a different team that has not used the new technologies—should also be carried out. Because of the complexity of the phenomena under study, separate studies investigating communication behaviors and memory errors will be required.

Third, the introduction of new technologies also requires investigation of the social environment into which they are introduced. The ethnographic approach, a method of observation focused on learning the meaning for participants of particular behaviors, will help characterize the social aspects of the communication environment, and drive hypothesis construction and further experiments.

**Conclusions**

This report has taken a cognitive psychological approach to set the scene for future investigations of clinical communication behavior. It has outlined some of the negative consequences that interruptive communication patterns could have on the ability of hospital staff to manage successfully their current and impending task loads. It has detailed several characteristics of human memory that might contribute to greater cognitive burdens. It has indicated some of the characteristics of tasks that contribute to the probability that they will be forgotten in interrupt-driven circumstances. It has suggested that the degree of an individual’s expertise is one characteristic that could be used to discriminate between the differing vulner-
abilities of individuals to interruptions. Finally, it has combined this cognitive approach with a social information approach to consider the effects of new communication technologies in the interrupt-driven environment of the hospital.

This approach reveals great scope for studying existing communication patterns and work practices in hospitals to understand their effects on memory and task performance. This is a necessary precursor to developing an understanding of the consequences of introducing new communication technologies.

References

25. Miller GA. The magic number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev. 1956;63:81–97.
34. Shiffrin RM, Schneider W. Controlled and automatic human information processing, part II: perceptual learning, automatic attending, and a general theory. Psychol Rev. 1977;84:127–90.
40. Soe LL, Markus ML. Technological or social utility? Unrav-
eling explanations of e-mail, v-mail, and fax use. Inf Soc. 1993;9:213–36.


