

Why is it so difficult to measure the effects of interruptions in healthcare?

Farah Magrabi, Simon Y. W. Li, Adam G. Dunn, Enrico Coiera

Centre for Health Informatics, Australian Institute for Health Innovation, University of New South Wales, Australia

Abstract

Interruptions are a complex phenomenon where multiple variables including the characteristics of primary tasks, the interruptions themselves, and the environment may influence patient safety and workflow outcomes. Observational studies present significant challenges for recording many of the process variables that influence the effects of interruptions. Controlled experiments provide an opportunity to examine the specific effects of variables on errors and efficiency. Computational models can be used to identify the situations in which interruptions to clinical tasks could be disruptive and to investigate the aggregate effects of interruptions.

Keywords:

Interruption, Medical error, Safety, Efficiency, Evaluation studies, Observation, Computer simulation.

Introduction

Interruptions seem to be inherent in the way work is undertaken in many clinical settings [1, 2]. Numerous studies have characterised this phenomenon over the last decade, and examined the extent to which doctors and nurses are interrupted in undertaking routine clinical tasks. Hospital doctors and nurses are interrupted anywhere between once every two hours, up to 23 times every hour in emergency, intensive care and surgery [3]. Interruptions may be a risk to patient safety in certain types of clinical tasks [4].

While there is solid evidence from psychology on the disruptive effects of interruption on human cognition [5], there is little evidence to date about the domain specific consequences of interruptions on clinical work. Interruptions have been reported as a contributing factor in IT-related medication incidents. Examination of 7029 CPOE-related medication incidents in the United States Pharmacopeia Medmarx database found that distractions were a significant contributing factor in eight out of ten errors [6]. In a UK study, interruptions and distractions were reported as a factor in up to 11% of errors in dispensing medications [7]. Interruptions also have a time cost. In one study clinical staff in an emergency department spent 24% of their time in dealing with interruptions [8].

A recent review of clinical interruption studies found that most studies were observational, describing the frequency, duration and types of interruptions [3]. Only one controlled experimental study examined the effect of interruptions on programming

an infusion pump [9]. The review concluded that it is currently not possible to be certain about the causal links between interruptions and errors in healthcare. Yet interruptions have been well studied in other domains such as psychology, human-computer interaction [10] and aviation [11]. In most instances interruptions have been found to be disruptive, increasing time spent on a task and leading to errors in completing tasks. Studies also identify some of the characteristics that make interruptions disruptive such as the interruption complexity [12], similarity of an interrupting task to the primary task [13] and the availability of retrieval cues in the primary task [14]. While this literature provides a useful starting point for examining interruptions in healthcare, the disparity in tasks and environmental context make it necessary to investigate the specific consequences of interruptions to clinical work and their impact on patient safety.

In this paper we seek to examine the problem of studying interruptions in healthcare. We argue that the complexity in studying interruptions and measuring their impact on clinical work is one reason why little is still known about the clinical consequences of interruptions and this may require different research methods to those currently used in interruption research.

Methods

This study draws upon general literature about interruptions in non-health domains; experimental studies of the impact of interruption on electronic prescribing [15] and error behaviour [16]; and a series of observational studies in emergency and intensive care [8, 17, 18]. We firstly bring together a range of process and outcome measures to describe the complexity and consequences of interruptions. This is followed by a discussion of how observational studies, controlled experiments and computer simulations can be used to study interruptions and their impact on the safety and efficiency of clinical tasks.

Results

Table 1 lists primary task and interruption variables that may contribute to the outcomes of interruptions.

Task and interruption variables

1) Task type: Primary tasks and interruptions can be described in terms of their phenotype or their genotype. The *phenotype* is used to refer to the surface characteristics of a clinical task. In an observational study we found that doctors attended to a

range of task phenotypes that included prescribing medications, answering their pagers, phone calls, conversations with patients and colleagues on the ward, consulting medical records and reviewing investigation results [19].

The study of interruptions also requires description of underlying task characteristics or *genotype* such as task complexity and working memory load requirements, which have been shown to influence the disruptiveness of an interruption [20, 21]. Few clinical tasks fit neatly under the broad heuristic classes of procedural or problem solving traditionally used in interruption studies to describe task phenotype. A more sophisticated approach may be Wood's definition based on acts and information cues allowing task complexity to be quantified and compared [22, 23]. Models such as GOMS (Goals, Operators, Methods, and Selection rules), a well-recognised cognitive engineering method, also allow differentiation of the relative complexity of computer-based tasks by working memory load requirements. We have applied this method to analyse electronic prescribing tasks [19].

Table 1- Variables associated with effects of interruptions.

| |
|----------------------------------------------------------------|
| Task and interruption variables (process measures) |
| 1. Task type (primary and interrupting task) |
| 2. Point of interruption |
| 3. Duration of interruption |
| 4. Similarity of interruptive task to primary task |
| 5. Modality of interruption |
| 6. Environmental cues |
| 7. Interruption handling strategy |
| Impact on task performance (outcome measures) |
| 8. Safety: task errors |
| 9. Efficiency: time on task, interruption lag, re-sumption lag |

2) Point of interruption: Interruptions at the beginning of a task appear less disruptive than those in the middle of a task [24]. Interruptions at the end of a task may also be particularly disruptive, especially when the main goal is accomplished and a small sub-task remains. For example, a nurse who administers a medication may fail to complete documentation relating to that medication when interrupted. Such errors are known as post-completion errors [16, 25].

3) Duration of interruption: The longer an interruption, the longer it takes to re-orient and restart the primary task afterwards, indicating that the disruptiveness of an interruption is directly related to its duration [26].

4) Similarity of interruptive task to primary task: In computer-based tasks, interruptions have been found to be less disruptive when they are dissimilar to the primary task [27].

5) Modality of interruption: A range of modalities like heat, smell, sound, vibration, and light can influence interruption effect [28]. For example, a doctor interrupting a nurse who is administering medications, to discuss the care of a patient, is a different experience to a device mediated interruption such as a phone, pager or an alarm, which could be turned off.

6) Environmental cues: The availability of information cues in the primary task, such as an x-ray, should help return to it after

interruption. Studies have shown blatant cues are highly effective, but the availability of subtle cues is the same as having no cues at all [29]. While the clinical environment is full of cues, some may be less obvious than others. A doctor who is interrupted when reviewing an x-ray is cued to resume and complete her primary task by the image display. The absence of cues when an interruption dislocates a clinician from her primary task may be particularly relevant in hospital settings where some roles require staff to be highly mobile. For example, when a pharmacist reviewing a chart is called away to a different unit, is not cued by the new environment to return to the primary task.

7) Interruption handling strategy: Clinical workers may use different strategies to handle interruptions, which may influence the extent of disruption to primary tasks. Consider the case of a doctor prescribing a number of medications who is interrupted by a mobile phone call:

- Attend to interruption:** The doctor may choose to take the call immediately or, with a momentary delay to rehearse the name of the next medication to be prescribed. The availability of an *interruption lag*, the time taken to attend to an interruption, may allow encoding of the next action. This can be effective in reducing interruption disruptiveness. After attending to the interruption the doctor may either *switch* to the interrupting task (i.e. suspend prescribing) or *multi-task* (i.e. prescribe while on the phone).
- Delay interruption:** The doctor may choose to switch off the phone and check for a message after finishing the primary prescribing task.

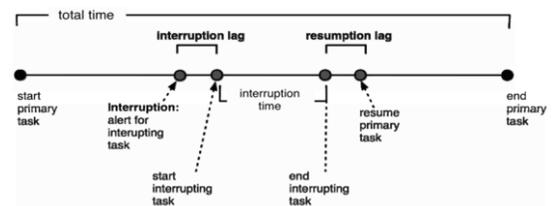


Figure 1- Time-course of an interruption (adapted from [30]).

Effects of interruptions on task completion

Interruption effects are typically quantified by examining errors [11, 16] and time penalties [5, 10].

Errors

An error is a "failure to carry out a planned action as intended, or application of an incorrect plan" [31]. Similar to task type, errors can be described in terms of their phenotype or genotype. We will use the example of electronic prescribing using a computerised physician order entry (CPOE) system to distinguish error phenotypes and genotypes.

Error phenotype: This is the observable property of an error. Or, the clinical manifestation of the error depending on the type of task. For example, a range of prescribing error types can be identified, such as wrong patient; missed allergy status; incorrect medication name dose, route or formulation; admini-

stration time; frequency; missing instructions to administer medications; and failure to prescribe or cease medications [15].

Error genotype: Is used to describe the underlying causes of an error or the ways of making the error.

- a. *Errors of commission:* are manifest by doing the wrong thing" [31].
- b. *Errors of omission:* are failures to do the right thing [31].

Errors can also be distinguished into slips and mistakes based upon knowledge about the intentions of an actor. A *mistake* is an "action that may conform exactly to the plan, but the plan is inadequate to achieve its intended outcome" [31]. Mistakes are associated with lack of knowledge about a task. A *slip* is an error of action, defined as "an unintended error or execution of a correctly intended action" [31]. Slips are associated with human fallibility in executing routine procedures despite having the correct how-to knowledge.

As with any other risk to patient safety measuring the frequency of errors due to interruption along with identification of potential causes is needed to facilitate investigation of corrective and preventative strategies. For example, in electronic prescribing identifying user interface features (i.e. dropdown menus, text entry) and actions associated with specific errors provides a basis to improve design of the CPOE system.

Measuring task errors: As clinical tasks often provide multiple opportunities to make an error, a normalised error rate, which takes into account the number of opportunities, can be calculated to compare the impact of interruptions across tasks (Equation 1) [16, 25].

$$\text{Error_rate} = \frac{E_j}{E_n} \quad (1)$$

where, E_j =no. of errors and E_n =no. of error opportunities.

Error rate can be aggregated by phenotype. For example, error rates can be calculated for each prescribing error type or an overall error rate can be computed by aggregating all error types. As some errors may pose a greater risk to patient safety than others, another way is to calculate a weighted score. For example, errors can be weighted with a value of 1 (low risk), 2 (moderate risk), 3 (serious risk) to calculate an error score.

Task efficiency

The impact of interruptions on efficiency is examined in terms of the resumption lag and time-on-task (TOT).

Resumption lag: The time taken to re-orient and then restart the primary task after an interruption is generally regarded as a measure to examine the time cost of interruption (Figure 1) [26, 30].

TOT: The time taken to complete a primary task is used to examine any residual effects of an interruption (Equation 2). Studies have shown that for simple office tasks participants tended compensate by speeding up post-interruption [20, 21]. This measure may be highly relevant in a hospital context where clinicians are time pressed with a finite amount of time to complete tasks.

$$\text{TOT} = T_{\text{total}} - T_i - t_{\text{inlag}} - t_{\text{rlag}} \quad (2)$$

where, T_{total} = Time taken to complete the primary task including any interrupting tasks;

T_i = Time taken to complete the interrupting task;

t_{inlag} = interruption lag; and t_{rlag} = resumption lag.

Discussion: Challenges in measuring the effects of interruptions

As we have shown, interruptions are a complex phenomenon where multiple variables including characteristics of primary tasks, different dimensions of the interruptions and the environment may influence outcomes. In this section we discuss some of the challenges in using observational studies, controlled experiments and computer simulations to measure the effects of interruptions in healthcare.

Identifying situations where interruptions are problematic

Observational studies have been successfully used to identify *types and frequency of primary and interrupting tasks; sources of interruption; and time on tasks*. A range of highly resource intensive methods and tools have been used to examine interruptions including audio (e.g. COM [8]), video recordings (e.g. computer screen-capture [32], eye-trackers) and direct observations of work patterns (e.g. PDAs) in clinical settings.

The time cost of interruptions is not easily measured in observational studies. Whilst it is possible to examine time on task, observational studies may only provide rough estimates of resumption lag, which is typically measured in the order of a few seconds. For example, it is possible to only broadly compare resumption times for a medication administration task with a documentation task that may require a nurse to spend some time gathering his thoughts prior to resuming.

The effect of interruptions on errors is less clear-cut. Errors often result from the interplay of multiple events over time, and the error contributed to by an interruption may only occur after a researcher has stopped observation. Whilst slips may be observable in tasks with well-defined structure (for example, a nurse does not check the name of a patient before administering medications) mistakes are less identifiable because the intention of the actor is not always clear to the observer. Further, error identification is complicated in tasks with less rigid task structures, i.e. when there are many ways of achieving a goal.

Given that multiple variables affect the outcome of an interruption, interruptions are likely to be problematic only in certain situations e.g. poor cue availability. While observational studies may be useful to identify particular situations in which interruptions are disruptive, they present significant challenges for recording many of the process variables that influence the effects of interruptions. These variables include the *point at which primary tasks are interrupted, the modality, length and similarity of the interruption to the primary task, environmental cues and interruption handling strategies* (Table 1).

Quantifying effects in the laboratory

Controlled experiments provide an opportunity to examine the effects of specific *task and interruption variables* on task completion (*errors and efficiency*) and have been effectively used to examine the effects of interruptions in other disciplines. In healthcare, controlled studies provide an opportunity to exam-

ine the impact of interruptions in a range of task environments from low fidelity laboratory tasks (e.g. use of a CPOE [15], programming of an infusion pump [9]) to high fidelity simulation environments (e.g. surgery [33], medication administration on a ward). However, within a laboratory setting there is a trade-off between obtaining causal relationships among controlled variables and ensuring ecological validity.

One of the challenges is to design tasks and interruptions that are representative of typical tasks undertaken by the intended participants. For example, to examine the impact of interruptions to electronic prescribing using a CPOE by junior doctors, tasks could be based on hypothetical clinical scenarios representing typical prescribing tasks undertaken by doctors working on a medical ward. Interruptions in such a setting will often be initiated by a phone or pager (*modality of interruption*) and require the doctor to walk away from her primary task.

Another challenge is to effectively control for all independent variables. For example, if several different clinical scenarios are used to examine interruptions in electronic prescribing all scenarios must be of similar task complexity. Interruption variables such as the *length of the interruption*, *similarity of the interruption to the primary task* and *cues available in the task environment* require meticulous consideration when designing the experiment. The experimental procedure for such interruptions requires careful coordination by investigators, especially to control for variables such as the *point at which primary tasks are interrupted*, and *interruption handling strategies*.

A third challenge relates to the measurement of outcomes. In particular, error rates as a measurement may not be sensitive enough to detect the disruptive effect of interruptions compared to measures such as resumption lag, which can be reliably measured in laboratory studies. Baseline error rates in clinical tasks are fairly low – for example, baseline error rates in electronic prescribing using a CPOE ranged from 0.5% to 15.6% [15]. Thus, designing an experiment that is adequately powered to detect a small effect size is a challenge (Table 2). While the sample size may be adequate to detect a difference in resumption lag, the study may not be adequately powered to examine the smaller differences in prescribing errors.

Table 2 - Sample size required to detect error rate with interruptions using a paired samples t-test [34]

| Effect size (d) | Small | Medium | Large |
|------------------|-------|--------|-------|
| | 0.2 | 0.3 | 0.5 |
| Sample size* (n) | 156 | 71 | 27 |

$\alpha=0.05$, $1-\beta=0.8$

In comparison, time on task and resumption lag are highly sensitive measures, which can be accurately captured. For example, screen capture software (e.g. TechSmith Morae®) can be used to record and analyse computer-based tasks and eye-tracking systems can be used for mobile task environments.

Examining impact on patient safety and task efficiency

The impact of interruptions on specific clinical tasks ought to ultimately be examined within the wider clinical workflow within which tasks are typically undertaken. Inbuilt error checking and redundancy within workflow plays a significant role in minimising the effects of errors. By investigating the

aggregate effects of interruptions over a large group of clinicians, we may better establish links between errors caused by interruptions and the adverse events that occur relatively infrequently. In addition, we can describe situations in which an individual's interruption time-cost does not necessarily correspond to an overall negative cost in efficiency for the group.

We know that very few errors lead to an adverse event (in one study, 1.4% of the total led to adverse events compared to a 60% error rate [35]). However, when we aggregate the effect of all errors across a hospital, the result is that 10% of admissions are associated with an adverse event [36]. This implies that the aggregate effect of errors on patient safety, whilst low in probability, is considerable and measurable. Since we also know that a considerable proportion of errors are associated interruptions, it follows that large numbers of interruptions contribute to errors that are associated with small but significant numbers of adverse events.

When extending beyond the individual to larger groups of cooperating clinicians, there may be a nonlinear trade-off between efficiency and interruptions. Intuitively, some interruptions are necessary and can improve the net task efficiency of the group (e.g. less need for idle waiting and asynchronous communication). However, given the relationship between interruptions and the potential for medical error, it follows that an increase in some types of interruptions may increase the overall potential for adverse events. This suggests an optimal level of interruptions that minimises error and maximises efficiency. Another limit occurs when the net time-cost of interruptions is negative – the aggregate time-cost to individuals caused by interruption outweighs the aggregate time-savings associated with reductions in idle time and asynchronous communication. There may be a non-intuitive optimal level of interruptions unique to each group and environment.

Computational modelling of interruptions

There is a strong precedent for using computational models in healthcare [37], and computational models have a role in the study of interruptions. They permit the explicit modelling of interactions between individuals and groups beyond what is typically done in a laboratory, assist in the development of hypotheses for controlled studies, and present one way to conduct first-phase validation of interventions prior to controlled studies [38]. Computational models provide an efficient way of examining the effects of different combinations of variables and can be used to identify the situations in which interruptions have the potential to be disruptive. Models of this type may help describe patterns of variability in interruption effects on groups and therefore locate problematic interruptions (problematic to overall efficiency and safety rather than just to the individual's efficiency and error rate).

Conclusion

Interruptions are a complex phenomenon with multiple variables that affect task performance. Disruption to clinical tasks can be understood by firstly using observational studies and computational models to identify the situations in which interruptions are particularly problematic. Secondly, using controlled experiments to measure the extent to which interruptions generate errors and impose a time cost. Computational

models can be used to examine overall impact on patient safety and task efficiency.

Acknowledgments

This research is supported by Australian Research Council (ARC) grants DP0772487 and LP0775532; NHMRC program grant 568612; and FM is supported by an ARC APDI Fellowship and the UNSW Faculty of Medicine.

References

- [1] Coiera E. Clinical communication: a new informatics paradigm. *Proc AMIA Annu Fall Symp.* 1996;17-21.
- [2] Coiera E, et al. Communication behaviours in a hospital setting: an observational study. *BMJ.* 1998 Feb 28;316(7132):673-6.
- [3] Grundgeiger T, et al. Interruptions in healthcare: theoretical views. *Int J Med Inform.* 2009 May;78(5):293-307.
- [4] Parker J, et al. Improving clinical communication: a view from psychology. *JAMIA.* 2000 Sep-Oct;7(5):453.
- [5] Altmann EM, et al. Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin & Review.* 2007;14(6):1079-84.
- [6] Zhan C, et al. Potential benefits and problems with computerized prescriber order entry: analysis of a voluntary medication error-reporting database. *Am J Health Syst Pharm.* 2006;63(4):353-8.
- [7] Ashcroft DM, et al. Prospective study of the incidence, nature and causes of dispensing errors in community pharmacies. *Pharmacoepidemiol Drug Saf.* 2005 May;14(5):327-32.
- [8] Spencer R, et al. Variation in communication loads on clinical staff in the emergency department. *Ann Emerg Med.* 2004 Sep;44(3):268-73.
- [9] Ginsburg G. *Interruptions: A Criterion in the Design and Evaluation of Human-Computer Interfaces*: University of Toronto; 2004.
- [10] Bailey BP, et al., editors. The effects of interruptions on task performance, annoyance, and anxiety in the user interface. *Human-Computer Interaction - INTERACT 2001 Conf Proc.*
- [11] Latorella KA. Investigating interruptions: An example from the flightdeck. *Proc of the Human Factors and Ergonomics Society 40th Annual Meeting* 1996. p. 249-53.
- [12] Cades DM, et al. Dealing with Interruptions can be Complex, but does Interruption Complexity Matter: a Mental Resources Approach to Quantifying Disruptions. *Proc of the Human Factors and Ergonomics Society 52nd Annual Meeting* 2008.
- [13] Oulasvirta A, et al. Long-term working memory and interrupting messages in human-computer interaction. *Behav Inf Technol.* 2004 Jan-Feb;23(1):53-64.
- [14] Altmann EM, et al. Task interruption: Resumption lag and the role of cues. *Proc of the 26th annual conference of the Cognitive Science Society* 2004.
- [15] Magrabi F, et al. Errors and electronic prescribing: a controlled laboratory study to examine task complexity and interruptions: CHI Technical Report UNSW, Oct 2009.
- [16] Li SYW, et al. The Effect of Interruptions on Postcompletion and Other Procedural Errors: An Account Based on the Activation-Based Goal Memory Model. *J of Exp Psychol: Appl.* 2008;14(4):314 - 28.
- [17] Alvarez G, et al. Interruptive communication patterns in the intensive care unit ward round. *Int J Med Inform.* 2005 Oct;74(10):791-6.
- [18] Coiera EW, et al. Communication loads on clinical staff in the emergency department. *Med J Aust.* 2002 May 6;176(9):415-8.
- [19] Magrabi F. Using cognitive models to evaluate safety-critical interfaces in healthcare. *CHI '08 extended abstracts on Human factors in computing systems*; Florence, Italy: ACM; 2008.
- [20] Speier C, et al. The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences.* 1999;30(2):24.
- [21] Speier C, et al. The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences.* 2003;34(4):771-97.
- [22] Wood RE. Task complexity: Definition of the construct. *Org Behav Human Dec Process.* 1986;37:60-82.
- [23] Sintchenko V, et al. Decision complexity affects the extent and type of decision support use. *AMIA Annu Symp Proc.* 2006:724.
- [24] Adamczyk PD, et al. If not now, when?: the effects of interruption at different moments within task execution. *Proc of the SIGCHI conference on Human factors in computing systems*; Vienna, Austria: ACM; 2004.
- [25] Byrne MD, et al. A working memory model of a common procedural error. *Cognitive Science.* 1997 1997/3/;21(1):31-61.
- [26] Monk CA, et al. The effect of interruption duration and demand on resuming suspended goals. *J of Exp Psychol: Appl.* 2008;14(4):299-313.
- [27] Edwards MB, et al. Task interruption and its effects on memory. *Memory.* 1998;6(6):665-87.
- [28] Arroyo E, et al., editors. *Interruptions as multimodal outputs: which are the less disruptive? Multimodal Interfaces, 2002 Proc Fourth IEEE International Conference on*; 2002.
- [29] Trafton JG, et al. Huh, What Was I Doing? How People Use Environmental Cues after an Interruption. *Proc of the Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, Florida, September, 2005* 2005:468 - 72.
- [30] Trafton JG, et al. Preparing to Resume an Interrupted Task: Effects of Prospective Goal Encoding and Retrospective Rehearsal. *Int J of Hum-Comput Studies.* 2003;58(5):583 - 603.
- [31] Runciman W, et al. Towards an International Classification for Patient Safety: key concepts and terms. *Int J Qual Health Care.* 2009 Feb;21(1):18-26.
- [32] Collins S, et al. Multitasking by clinicians in the context of CPOE and CIS use. *Stud Health Technol Inform.* 2007;129(Pt 2):958-62.
- [33] Liu D, et al. Interruptions and blood transfusion checks: lessons from the simulated operating room. *Anesth Analg.* 2009 Jan;108(1):219-22.
- [34] Faul F, et al. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods.* 2007 May;39(2):175-91.
- [35] van Doormaal JE, et al. Medication errors: the impact of prescribing and transcribing errors on preventable harm in hospitalised patients. *Qual Saf Health Care.* 2009 Feb;18(1):22-7.
- [36] Runciman WB, et al. *Safety and Ethics in Healthcare: a guide to getting it right*: Ashgate Publishing; 2007.
- [37] Fone D, et al. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *J Public Health Med.* 2003 Dec;25(4):325-35.
- [38] Epstein JM. Why Model? *J Artif Societies and Soc Simulation.* 2008;11.

Address for correspondence

Centre for Health Informatics, University of New South Wales, Sydney 2052, Australia. email: f.magrabi@unsw.edu.au