Violations and migrations in health care: a framework for understanding and management

R Amalberti, C Vincent, Y Auroy, G de Saint Maurice

Violations are deliberate deviations from standard procedure. The usual reaction is to attempt to eliminate them and reprimand those concerned. However, the situation is not that simple. Firstly, violations paradoxically may be markers of high levels of safety because they need constraints and defences to exist. They may even become more frequent than errors in ultrasafe systems. Secondly, violations have both positive and negative aspects. On the one hand they occur frequently, increase system performance and individual satisfaction, are mostly limited to practices with limited safety consequences, and therefore are often tolerated or even encouraged by the hierarchy. On the other hand, extreme violations can lead to real danger or actual harm. This paper proposes a three phase model derived from Rasmussen’s theory of migration to boundaries to explain the mechanism by which the deviance occurs, stabilizes, regresses, or progresses to harm. The model suggests that violations are unavoidable because system dynamics and deviances are markers of adaptation to this dynamicity. Violations cannot be eliminated but they can be managed. Solutions are specific to each step of the model, with a mix of relaxing constraints, increasing peer control (staff), and constraining dangerous individuals.

In 1987 the ferry Herald of Free Enterprise left Zeebrugge’s inner harbour, took on water, and sank. The immediate cause of this loss was that, to save time, the back ramp had not been fully closed before the ship left harbour. In 1999 an accident occurred at an uranium processing plant operated by JCO in Tokai-Mura, Japan. The workers were anxious to finish their job at the conversion building and decided to use the precipitation tank instead of a buffer column (a much smaller device) to increase their performance when purifying and homogenizing uranium. The concentration of product became critical and the system exploded, causing the most severe nuclear accident since Chernobyl. Both these accidents were caused not by unintended errors but by deliberate deviations from rules and standards. In both cases, accident analysis showed that the workers’ deviations from normal operating procedures resulted from a long progressive drift in practice. These deliberate deviations—known as violations—have been an important topic in safety analyses in industry but have been little studied in health care.

Violations are deliberate deviations from standard instructions. However, defining non-compliance is not straightforward. The expected level of compliance—and therefore the interpretation of non-compliance and violation—varies according to the type of instruction, the nature of the work, and the social and organizational context. In some cases strict observance of rules is expected, whereas in other cases a certain degree of flexibility is tolerated or even expected. In health care, for instance, evidence based medicine is seen as a guide to practice, not a mandatory set of instructions. Serious violations of strict rules are likely to be severely penalised, if discovered, in any environment. This makes the study of violations particularly challenging in that studies are addressing a topic that is both sensitive and partially hidden from the usual lines of enquiry. The difficulty of studying violations partially explains the lack of knowledge and understanding of the topic.

Violations have been the cause of some major healthcare incidents. For instance, five deaths were recorded in Florida after liposuction was given in doctors’ offices. The death rate for this procedure was 1 per 5000 in offices compared with 1 per 200 000 in hospital. The primary cause of death in the doctors’ offices was that patients were given very high doses of lidocaine that “pushed the envelope” of use of that drug. Following these deaths the state of Florida imposed a 90 day moratorium on all ambulatory surgery to allow further investigation; there were very few regulations covering procedures in doctor’s offices and there were clear and frequent violations of the recommended protocol for the use of local anesthesia. However, this last example also shows how difficult it is to define a violation. Although such high doses of lidocaine were clearly unacceptable violations of accepted anesthesia practice, they were nevertheless tolerated if not completely accepted in the professions of dermatology and plastic surgery.

Violations occur frequently in all industries, even those with very good safety records. In aviation, an extensive study of crews’ deviations from procedures (noted by trained observers sitting on the jump seat for some 3500 flight segments) showed that “intentional non-compliance” represented 55% of all errors and violations, but only 3% of these affected the flight in any adverse way. In another study Degani and Wiener showed that about half the checklists on airplanes were not correctly completed mainly because of interruptions and
distractions and poor design of check lists. In medicine a total of 67 violations of procedure over 59 operations were noted during a large study of alarm systems in the operating theatre.\textsuperscript{7} The most significant event occurred in a private hospital that was close to bankruptcy. One of the anaesthesiologists, apparently from a desire to save money for the clinic, decided to reuse the same syringe (of diprivan) for consecutive patients. He loudly justified this deviant behavior to the nurses, telling them that the surgical list in the theatre that morning consisted only of a series of short, rapid, turnover cases. If he were to use a fresh ampoule of diprivan for each patient, “then there would be a considerable cost associated with this practice, without any safety benefits”. Note that this doctor did not consider the risks of this behavior, justifying it solely on the grounds of economy, and also that the behavior had no immediate consequences for patients. This behavior was apparently accepted by the operating theatre medical staff, nurses, and surgeons. This illustrates the fact that violations may need to be explained both by individual motivation and wider social and organizational processes.

Violations, therefore, are a complex multifaceted phenomenon. They occur frequently and may save time and bring benefits to both individuals and systems. They may be tolerated by the wider clinical team and even actively encouraged if there is pressure to increase workload and throughput of patients. However, extreme violations may put both people and systems at risk. The existence of violations poses a number of problems for health care:

- Are some violations acceptable if they do not lead to danger or harm?
- Are acceptable and unacceptable violations part of the same continuum?
- What are the criteria for tolerance?
- Should they lead to different safety approaches?

This paper addresses these questions by providing a framework for understanding violations and system migration.

**ABSENCE OF DATA SHOULD NOT MASK THE IMPORTANCE OF THE PROBLEM**

Data on violations are sparse in health care in that they are seldom explicitly assessed. For instance, in the major epidemiological studies of adverse events the term “violation of a protocol or a rule” is explicitly used only in the Quality in Australian Health Care Study of 1999 where “violations were a cause of 4.8% of adverse events. There are no data on violations in the other national studies of adverse events.\textsuperscript{7–10} Three reasons can be put forward to explain the relative absence of information about violations in health care.

Firstly, health care has fewer explicit rules than other high hazard industries. While there are a large number of protocols and guidelines, there is usually sufficient flexibility and room for clinical judgement so it can be difficult to state unequivocally that a violation has occurred. The multiplicity of guidelines and recommendations at both national and local levels means that there can never be strict compliance to any particular set of rules or guidelines; inevitably, national guidelines in particular are never seen as more than recommendations. Secondly, much of the information about safety problems in health care has been derived from incident reporting systems and it is difficult to assess precisely the presence or absence of errors and violations from summary narrative data. Last, but not least, violations are the most difficult unsafe acts to gather in any voluntary reporting system. Reporting violations may lay the individual open to accusations of negligence or professional misconduct even if there was no intention to harm.\textsuperscript{11–13} For all these reasons, violations are insufficiently studied in health care, despite the fact that they probably represent a serious source of danger and because they tend to become more frequent than errors as systems become safer.\textsuperscript{14}

**PERSPECTIVES ON VIOLATIONS**

Violations can be understood from a number of different perspectives, which vary in the nature of the explanation advanced and the discipline from which they are derived. We briefly outline some of the principal theories and then discuss an overall framework which endeavors to integrate the cardinal elements of these various perspectives.

**Motivation and attitude**

The first category of theory emphasises the local and contextual motivation of individuals as the main source of the violation. The best example of this category is the theory of planned behavior,\textsuperscript{15} where the likelihood of violation is determined by the individual’s assessment of the consequences, the social influences on them, beliefs about control, and personal moral codes and beliefs. This theory was used, for example, to measure the attitudes and intentions of drivers to commit violations such as drunk driving, speeding, close following, and dangerous overtaking.\textsuperscript{16} The theory relates the occurrence of violations to individuals’ willingness to break rules, the likelihood of detection and of consequences.

**Organizational and cultural approaches**

The second category of theories points to organizational and cultural factors as the principal causes of violations. We can distinguish two broad theoretical frameworks in this category.

The first framework relates both to the “vulnerable system syndrome” (VSS) proposed by Reason\textsuperscript{17} and Tucker and Edmondson’s concept of first and second order problem solving.\textsuperscript{18} Reason et al proposed that a cluster of organizational pathologies—the VSS—renders some systems more liable to errors and violations and, because of this, to accidents and adverse events. VSS has three interacting and self-perpetuating elements: blaming frontline workers, denying the existence of systemic error provoking weaknesses, and the blinkered pursuit of productivity and financial indicators. Such organizations fail to learn and so perpetuate problems. This parallels Tucker and Edmondson’s description of hospital workers responding to problems with a quick fix—the immediate problem is solved but the underlying problems are ignored and allowed to continue.

A second well known framework, proposed by Diane Vaughan, also considers violations as a result of a pathological culture but highlights the role of social routines that progressively mask the problem. Vaughan considers that violations set in gradually over time, with operators becoming gradually more lax in their performance. This “normalisation of deviance” is maintained by structural secrecy, the absence of incidents, and the tolerance and absence of reaction of senior management. The best example of this approach is probably the Challenger space shuttle explosion.\textsuperscript{19}

**Adaptation and flexibility**

The third category of theories was developed within the Russian and French occupational psychology tradition of adaptation, health and a consideration of violations as a necessary adaptation of professionals in coping with the conflicting demands of complex work situations.\textsuperscript{20–24} From this perspective, violations are not a hazard; rather, they reflect the intelligence and flexibility of frontline workers. Extensive studies, mostly reported in French, have been devoted to understanding the adaptation of procedures by workers to deal with the demands of the work. The catachreses (creation

www.qphc.com
of new usage for a tool) have been specifically studied within this framework. Solutions to the problems of violations rely less on asking workers to increase their adherence to procedures than on changing the design of the work to tolerate greater flexibility of practice.

Another well known contribution in this same field of adaptation is the law of requisite variety expressed by Ashby in 1956. This law considers that a large procedural variation is the only guarantee of an effective learning process and should not be assimilated to error.

All of these approaches on adaptation have largely inspired the basis of the high reliability organization’s theoretical framework and the need for organizational learning in health care.

The theories outlined in this section address different aspects of the same problem from different perspectives at different stages of the evolution of the social system. Some address individual motivation while others address organizational and cultural influences. Most tend to suggest that violations are behaviors that need to be eradicated in the pursuit of safety, although it is clear that violations are often unavoidable even if not desirable.

A GENERAL FRAMEWORK FOR UNDERSTANDING THE POTENTIAL FOR VIOLATIONS

The final theory to be described derives originally from Rasmussen’s framework model of system migration which has been subsequently extended by Amalberti. The Rasmussen/Amalberti framework attempts to integrate the various perspectives on violations and to resolve some of the apparent contradictions between them. Rasmussen emphasized that front line workers do not follow procedures in a strict and logical manner, but try to follow the path that seems most useful and productive at the time. Workers operate within an envelope of possible actions which is influenced all the time by wider organizational and social forces. Rasmussen also described the pressures on individuals and systems to move towards the boundaries of safe operation, as workers are constantly having to adapt and react to pressures for increased performance and productivity which erode the margins of safety. Furthermore, these violations can become more frequent and more severe over time so that the whole system “migrates” to the boundaries of safety until an accident or recalibration occurs which forces a realignment. These external pressures, coupled with individual rewards and benefits, may over time modify the work being carried out, lead to rules and recommendations being progressively ignored, and eventually greatly increase the possibility of disaster as the organization becomes accustomed to operating at the margins of safety.

Amalberti has extended Rasmussen’s model to study violations in aviation, train drivers, and rotary press. In particular, Amalberti has drawn attention to the time course of system development and the manner in which a system migrates to the boundaries of safety. Three phases in this process can be distinguished:

- Initial safe space of action.
- Creation of borderline tolerated conditions of use (BTCUs).
- Normalization of deviance and reckless individuals.

Initial safe space of action

The first phase corresponds to the initial design of the work process. At that stage the process is, ideally, designed to operate according to a set of rules and procedures with some regard for the likely pressures of production. In production line or automated systems, many constraints and failsafe procedures will be introduced to act as defences against error and violation and constrain the limits of human action (fig 1). In health care, processes often evolve rather than being...
designed and constraints may not be strong, but there will at least be an informal understanding of the proper procedures.

Creation of borderline tolerated conditions of use

The second phase occurs when the process commences operation and must continuously adapt to new social and technical demands. The system migrates towards the boundaries of safe operation through the combination of pressure for greater performance (horizontal axis) and the secondary advantages for the individual (vertical axis). Barriers are quickly bypassed under the pressures of real life. Senior management are frequently under pressure to increase output and constrain the use of resources (do the same amount of work with less staff, with equipment missing or out of order, etc.). These demands transmit at the front line to a pressure to act more quickly and, ultimately, to violate basic procedures. Once this has occurred there may be a second migration in the sense of a further move towards unsafe practice. Individuals who have been pressured to cut corners in order to increase performance come to think that they are “officially” transgressing established rules and that their behavior has in some way been sanctioned. The result is that the system migrates towards a “normal illegal” area of stabilised operation.

At the “normal-illegal” stage, violations are better termed “borderline tolerated conditions of use” (BTCUs)⁶ and may be viewed as providing management and individuals with the maximum benefit for the minimum and accepted probability of harm. These BTCUs have four features: (1) they are first seen as benefits and not as risks; (2) they enhance performance of the system or provide advantage for the individual; (3) they are tolerated by senior management and sometimes even required by it; and (4) they are associated with a variety of informal safety measures. Safety behavior, which may or may not be effective, is now operating within a social context quite independent of the rules and procedures envisaged by the designers of the system.

Normalization of deviance and reckless individuals

The third and last phase occurs after a certain amount of time has passed. The same violations may be committed as in the second phase, but these are now routine and so common as to be almost invisible to both workers and managers, echoing the normalization of deviance noted by Diane Vaughan.⁷ At this stage, any further deviance may easily result in patient harm and is generally counted as negligent or reckless conduct. A limited number of individuals, in the absence of a tight social control, are willing to violate basic procedures to the point of recklessness and actual patient harm. Moreover, these individuals are not only a danger to themselves but they may also influence the “other” workers if no action is taken to control them.

CONSEQUENCES OF THE MIGRATIONS MODEL FOR SAFETY MANAGEMENT

The arguments set out above suggest that violations pose considerable challenges for the management of safety. In most settings they are numerous, and yet comparatively few lead to harm or real danger. They are therefore tolerated and even viewed as normal occurrences in routine work. Furthermore, they are influenced by a range of personal, social and organizational factors and their occurrence may also have a distinct time course as a system migrates to the boundaries of safety or recalibrates following an adverse incident. As yet, violations are incompletely understood and the research base remains extremely slender. However, even with this limited knowledge, there are some important and immediate practical implications.

Incident reporting systems do not detect violations and migrations

Reporting systems are poor at reflecting the nature and frequency of violations, particularly when these take the form of BTCUs—that is, violations that have become routine and tolerated. As these events have become normal, they are not regarded as unusual and it is difficult to act decisively even when they are detected. Ineffective memos may be circulated reminding workers of the (old) rule, but the impact of these is short lived since the new behavior has already become socially sanctioned.

Detecting violations and system migration requires attention to progressive drift in practices (proactive control) rather than reacting to incidents (reactive control)

While it is difficult to provide definitive data on violations, except in a specially designed study, the issue can nevertheless be addressed in meetings between staff. Just as violations are in part a socially determined phenomenon, relying on complicit acceptance by the group so they can be reduced by a mutual decision that such behavior will no longer be tolerated. Such discussions can take place in a meeting of clinical staff, provided the culture is open enough to allow such conversations to take place. For such a discussion to be productive it is vital that senior clinical staff (and, ideally, managers) are also involved, both to understand the problem of BTCUs and to discuss the acceptability and elasticity of “rule interpretation”.

Violations cannot be eliminated but they can be managed. This may require adapting existing rules. A system lives and changes and these transformations must be accepted. The case reported in box 1 is enlightening because it shows that it is counterproductive to try and set up impenetrable barriers against violations. When thinking about safety we tend to think of an ideal world of clear rules and procedures, but actually these defences can be extremely fragile. The rules and procedures give a sense of reassurance, but we seldom test them in a different context such as during weekends or with low qualified staff. We need first to understand the pattern of violations and system migration, while gradually changing the behavior of the staff within these systems. In

Box 1 The difficulty of eliminating violations

In a Paris hospital in 1997 a patient died in the ICU because an alarm failed on a monitoring device. After investigation this machine proved to be new and met all European standards. These new standards, contrary to the previous French standard, allowed all alarms to be turned off; the previous French equipment standard meant that at least one alarm was always on. Once the risk was identified, the general manager chose to contain and neutralize the risk by deciding that the procurement policy would allow only the purchase of equipment that conformed to the older French standard.

However, a second death occurred for the same reason in 2000. After investigation the general manager noted two violations of the procurement policy. Firstly, a significant number of machines in operation in France had been returned to European standards after maintenance operations carried out by the manufacturer. In these cases the local maintenance department had often not been informed, as was the case for the machine blamed for the second death. Secondly, in spite of the procurement policy, some machines had been bought which complied with European standards and allowed all alarms to be turned off.
other words, it is better to manage risk than to try to eliminate it artificially because history shows that, sooner or later, defences will be overturned.

Controlling violations and system migration by limiting the triggering conditions

Several factors that trigger violations are well known in the literature, the most frequent being the setting of unachievable goals. Whenever a standard is set, some individuals or organizations will decide that the costs of compliance exceed the cost of non-compliance. As standards are made more stringent, the costs of compliance increase steeply while the cost of non-compliance remains more or less constant. Explicit discussion of this issue with those setting targets is necessary to control this. Other triggering factors operate at the level of the clinical team, as staff will sometimes violate a procedure in order to save time or to help other members of the team—for example, inadequate briefing of a new member of staff so that a clinical can proceed more quickly. Such issues must be seen not just as “one-off” instances of no consequence but as a small step in the migration of a system to its safety boundaries.

Identify and control individuals who are more prone to violations

Patient safety has made much of the role of system factors in the occurrence of error and, to a lesser extent, in the occurrence of violations. This is in some ways unfortunate in that it has meant that insufficient attention has been paid to personal factors that may lead to safe or unsafe practice. In all human groups, some individuals are always more prone to deviate than others; in such case, the control of violations largely depends on the wider clinical team. In the absence of this effective peer control, some reckless or overconfident individuals will continue migrating to the boundary of safe practice and even over it. Once that stage has been reached, the only possible way of controlling the system is to report that individual and take appropriate action. In dealing with these individuals, a certain flexibility is required depending on the problem and maturity of those concerned (see Caroll and Rudolph elsewhere in this issue).

CONCLUSION

Human beings never fully comply with rules, and deviation from procedures occurs in all industrial systems. Violations and system migration have been insufficiently studied and it is important to acquire more knowledge about the causes and evolution of violations. New methods will be required to study migration and violations in health care, both for the purposes of research and in order to observe and manage violations in clinical settings. Reporting systems are largely ineffective in monitoring more serious violations, usually only providing information after a violation has caused some harm.

Managing violations is not a trivial matter. Some flexibility with regulations and standards is probably required in complex sociotechnical work to make the system efficient and adaptive to changing circumstances. However, more extreme violations may lead to a dangerous loss of control of both individuals and systems. The balance between the two extreme positions—tolerant versus punitive approaches—is not well understood and requires fine judgment and continuous monitoring in any safety critical setting. In the present state of our knowledge it seems that regular short periods of systematic observation of practice and a continuous dialogue about practice within clinical teams are probably the best methods of managing violations and preventing extreme system migration.

Key messages

- If a system is designed with only a limited sphere of safe operation, violations are very likely to occur under the conditions of actual performance.
- Violations cannot be eliminated but they can be managed. Working conditions, staffing, and medical knowledge always evolve and change over time. Borderline tolerated conditions of use (BTCUs) are best thought of as an understandable—although not necessarily desirable—adaptation to these changes. Simply considering BTCUs as unacceptable behaviors requiring disciplinary action is unhelpful; a better strategy is to monitor performance continually and to identify both violations and system migrations at an early stage.
- Dialogue between clinicians and managers is a key factor in establishing a shared safety culture. Violations and potential system migration must be discussed openly.
- The management of violations must begin at the clinical level, in ongoing discussions between staff on standards of safe practice and acceptable and unacceptable deviations from rules and standards. When it is clear that a violation is in fact adaptive, then procedures may need to be adjusted to reflect this.
- In the absence of effective peer control, some individuals will gradually commit more extreme violations. At that stage the only way of controlling the system is to identify that individual and to take appropriate remedial or disciplinary action.

ACKNOWLEDGEMENTS

This paper is an extension of an unpublished working paper presented at The Canadian Healthcare Safety Symposium, Edmonton, Alberta, Canada, 14–16 October 2004. The authors thank Jan Davies for her helpful suggestions.

Authors’ affiliations
R Amalberti, IMASSA, Cognitive Science Department, Brétigny sur Orge, France
C Vincent, Imperial College and St Mary’s Hospital, London, UK
Y Auroy, G de Saint Maurice, Percy Military Hospital, Paris, France
Competing interests: none declared.

REFERENCES

Violations and migrations in health care