Learning From Mistakes Is Easier Said Than Done: Group and Organizational Influences on the Detection and Correction of Human Error

AMY C. EDMONDSON
Harvard University

This research explores how group- and organizational-level factors affect errors in administering drugs to hospitalized patients. Findings from patient care groups in two hospitals show systematic differences not just in the frequency of errors, but also in the likelihood that errors will be detected and learned from by group members. Implications for learning in and by work teams in general are discussed.

The evening nurse reported for work at 3 p.m. in the surgical intensive care unit of University Hospital[1] and began her first round. Checking on a patient admitted to the unit more than 24 hours earlier after a successful cardiac operation, she noticed that the bag of medication hanging upside-down in the intravenous drip was not heparin—a clot-preventing blood thinner routinely administered after heart surgery—but was instead lidocaine. Lidocaine, an anesthetic and heart rhythm stabilizer, is not likely to harm a patient for whom it is not prescribed; however, the absence of heparin might

This project was funded by the Agency for Health Care Policy and Research as part of the Prevention of Drug Related Complications Study at the Harvard School of Public Health. The author benefited from the support and advice of Lucien Leape, David Bates, and David Cullen (the three principal investigators of the study), and two other members of the project team, Kathy Porter and Martha Vander Vliet. Andy Molinsky contributed enormously to this article by conducting much of the qualitative research. Clayton Alderfer and two anonymous reviewers provided many helpful suggestions that greatly improved the final version of this article. Finally, I am indebted to Richard Hackman who provided invaluable advice and feedback at each stage of this research.

Amy C. Edmondson is a doctoral candidate in the joint program in organizational behavior at the Harvard Business School and Department of Psychology. Her current research explores issues related to organizational learning and work team effectiveness.

JOURNAL OF APPLIED BEHAVIORAL SCIENCE, Vol. 32 No. 1, March 1996 5-28

have been fatal. Fortunately the patient suffered no ill effects from the error in this case, which subsequently was investigated by the author as part of a larger interdisciplinary study seeking to understand causes of these kinds of errors in hospitals.[2]

The average hospitalized patient receives 10 to 20 doses of medication each day and stays for 5 or 6 days (Leape et al., 1991), thus risking exposure to errors such as this about 85 times. Bates, Boyle, Vander Vliet, Schneider, and Leape (1995) recently found an average of 1.4 *medication errors* per patient per hospital stay, with 0.9% of these errors ultimately leading to serious drug complications. In an earlier population study, 0.35% of 80,000 patients in New York State Hospitals suffered "a disabling injury" caused by medications during hospitalization (Leape et al., 1991). These different frequencies—0.35% to 140%—reflect a range of drug error phenomena, from infrequent but disabling injuries to less consequential but far more frequent errors in medication dosage or timing. One fact is clear: Mistakes in administering drugs in hospitals occur and some patients are harmed by these errors.

Patient injuries related to drugs are called adverse drug events, or ADEs. ADEs are classified here as either *preventable* (the result of human error) or *nonpreventable* (not involving human error, as in an unpredictable allergic reaction) (Leape et al., 1991). The present study explores underlying causes of preventable errors in drug administration. As these errors occur in organizational contexts, this study examined organizational and group influences on the process of administering drugs to patients, toward a goal of supplementing existing research that has focused on classifying types of errors and identifying which individuals or professions are more likely to make them.

The Heparin Event

To illustrate how organizational systems affect error rates, consider further the heparin case. First, we might ask who was responsible for this life-threatening mistake. The answer is not completely straightforward. Between the point at which the operating surgeon prescribed the routine administration of heparin and the evening nurse discovered the error, no fewer than six health care professionals were responsible for the patient's care. The perfusionist, a medical technician in the operating room, hung the wrong bag in the intravenous drip. In an interview with the author, he claimed to have pulled the bag off the shelf where heparin belongs, and he noted that the bags look alike. (Did pharmacy mistakenly stock lidocaine where heparin should be? Perhaps, but should not the perfusionist have checked the bag carefully without assuming it to be heparin?)

An anesthesiologist wheeled the patient to the surgical intensive care unit and failed to notice the error. Then, another nurse, on duty the afternoon the patient arrived in the intensive care unit, was responsible for checking medications. How could she have missed the perfusionist's mistake? Easily, according to cognitive psychologists. The human tendency to perceive what one expects to see rather than what is actually there is a well-documented psychological phenomenon (Norman, 1980, 1981; Reason, 1984; Rumelhart, 1980). A heparin bag hanging by the bedside of postcardiac surgery patients is utterly routine—making it all too easy for caregivers to assume the presence of heparin and miss the error. This nurse was replaced by another at 11 p.m., and another at 7 a.m. the next morning. All of them could have detected the mistake before the next evening nurse with whom this episode opened came on duty.

Had the patient experienced an adverse drug event, it would have been technically difficult to assign blame to one person, for it was in the job descriptions of several individuals to avoid and to check for such errors. Drug administration is a collective undertaking—inviting the question of whether drug error rates can be explained as a function of group or hospital unit properties, rather than focusing exclusively on individual characteristics such as ignorance, fatigue, or carelessness. Moreover, in terms of preventing ADEs, individual errors cannot ever be completely eliminated. Indeed, even when everything "goes right" in the hospital setting, errors undoubtedly have occurred without notice or consequence. When an ADE caused by human error is documented, that error is likely to have involved multiple errors—such as the repeated errors of failing to detect and correct the discrepancy, as described above. Secondary errors of not noticing are as critical as primary errors that start a potentially harmful chain of events.

In summary, the system of drug administration in a modern hospital is complex, involving multiple handoffs in the journey from physician decision making all the way through to the receipt of a medication by a patient. Bates, Leape, and Petrycki (1993) have identified 10 points at which an error can occur (or be caught): (a) physician prescription, (b) initial delivery to a unit secretary who (c) transcribes the order, which then (d) must be picked up by a nurse who (e) verifies and transcribes again and (f) hands off to the pharmacist who (g) dispenses the medication and (h) sends it back to a nurse who (i) administers to a patient who (j) receives the drug. The present study explores organizational influences on the execution of these loosely coupled tasks, all of which take place within the context of hospital patient care "units." In the next section, I review approaches to understanding errors and accidents in organizations at three levels of analysis, to set a context for the design of the present study.

THE SOCIAL PSYCHOLOGY OF ERRORS AND ACCIDENTS

Medical researchers, psychologists, and organizational theorists have conceptualized errors and accidents in at least two different ways, and their research has investigated both the causes of errors and the efficacy of preventive strategies. One approach focuses on the individual and the other on the role of the system in which individuals operate in inducing or preventing accidents. A third perspective, proposed here, integrates system and individual levels of analysis by focusing on the work group as the point where organizational and cognitive effects meet and play out in enabling or preventing errors.

Individual-Level Analysis

In general, medical researchers have tended to emphasize the role of individual caregivers in studies of errors, and to blame physiological or educational deficits for adverse events (e.g., Lesar et al., 1990; Melmon, 1971). Many studies have been conducted to detect the frequency of errors among hospitalized patients and to identify responsible individuals. For example, physicians have been identified as more often responsible for such mistakes than nurses, pharmacists, or other personnel (Bates et al., 1993). Suggested strategies for prevention include individual-focused devices such as computerized order entry to prevent errors caused by poor handwriting and physician remedial education (Classen, Pestotnik, Evans, & Burke, 1991; Cohen, 1977; Massaro, 1992).

Psychologists have offered both cognitive and affective explanations for human error. According to schema theory, perceivers' expectations, or frames, have the power to steer attention away from actual visual data, enabling perceptual processes to construct images consistent with expectations (Rumelhart, 1980). If we see what we expect to see, we can make mistakes such as administering lidocaine instead of heparin. Similarly, well-learned activities can be carried out without conscious attention, allowing us to make odd slips such as putting a cereal box in the refrigerator (Norman, 1981; Reason, 1984). This automatic quality is indeed present in some drug errors; "I wasn't thinking" or "I can do [a particular task] in my sleep" are frequent phrases used by nurses interviewed, describing how an obviously wrong medication was given or not noticed. Such cognitive explanations for error have no need for Freud's "hidden impulses" to generate "slips" (Goleman, 1985); however, the unconscious effects of emotions such as anger or anxiety also can induce error, by distracting people's attention from the task at hand. Although the current social psychological literature emphasizes cognitive explanations for human behavior (Fiske & Taylor, 1991), in examining group-level influences below, we reconsider the role of emotions.

System-Level Analysis

Certain sociologists and organizational theorists have focused on the properties of *systems* in understanding error (e.g., Perrow, 1984). Rather than trying to explain why slips are made by individuals, this approach examines the design of systems and how systems give rise to human error. The nature of the system both influences the actions of individual operators and determines the consequences of errors. Perrow (1984) describes a "normal accident" as a predictable consequence of a system that has both interactive complexity and tight coupling. *Interactive complexity* is characterized by irreversible processes and multiple, nonlinear feedback loops. Interactively complex systems thus involve hidden interactions; the consequences of one's actions cannot be seen (Perrow, 1984). Giving a patient the wrong medication has this quality; immediate feedback is typically not present. *Tightly coupled* systems have little slack; actions in one part of the system directly and immediately affect other parts. The system of drug administration to hospitalized patients offers considerable interactive complexity; however, because the procedures linking medications to patients are loosely coupled, failures in a part of the system can be caught and corrected without causing harm. Although the modern hospital thus does not fit Perrow's worst case scenario, the interactive complexity of medications does create considerable potential risk for patients.

In Perrow's model, individual human error is taken as a given; the critical question is, under what conditions is the ever-present potential for human error dangerous? When do simple slips trigger an irreversible chain of events—by virtue of properties of the system—that ends in disaster? He challenges "the ready explanation of operator error" (Perrow, 1984, p. 26) and proposes that certain systems are accidents waiting to happen.

Just as the design of a system can invite accidents, straightforward solutions—implemented to prevent specific kinds of errors that have occurred in the past—can lead to unintended negative consequences. Organizational systems tend to resist straightforward solutions to problems (Forrester, 1971). When only the superficial symptoms of complex problems are addressed, the underlying problem typically remains unsolved, and even can be exacerbated if the solution feeds into a vicious cycle (such as providing food as direct aid, which relieves starvation but perpetuates the problem of population growth in inhospitable climates; Senge, 1990). This perspective suggests that strategies such as remedial education of physicians who have written incorrect prescriptions in the past, or use of warning stickers to remind caregivers of allergies, may have limited effectiveness in preventing drug errors, as they fail to consider the nature of the system that consists of many individuals and technical systems interacting in administering drugs to patients.

Organizational systems also transmit broader social forces that affect attitudes and behaviors related to error—including social forces that preclude "embracing error" (Michael, 1976) and thus inhibit learning. A widely held view in society of error as indicative of incompetence leads people in organizational hierarchies to systematically suppress mistakes and deny responsibility (Michael, 1976). Hierarchical structures thus discourage the kind of systematic analysis of mistakes that would allow people to better design systems to prevent them.

In summary, systemic approaches expand the scope of analysis of errors and accidents beyond the idiosyncrasies of individual behavior. Moreover, they introduce the variable of how errors that have been made are treated in the system, and how systems for error prevention can be designed. Some of these theories focus on technical and structural features of macro systems that primarily influence the outcomes of error and thwart simple preventative strategies; others focus on societal forces that shape attitudes toward error. However, these theories are less useful for understanding variance in behavior within a given organization or system. Thus we turn to the social psychology of groups for insight into how small social systems such as work teams may operate to prevent or catch errors.

Group-Level Analysis

Individual skills, motivation, and cognition are imperfect. Organizational systems are inevitably flawed. Hackman (1993) proposes that, facing this ever-present potential for error, teams in organizations can act as "self-correcting performance units." Members of a superb team have a way of coordinating tasks, anticipating and responding to each other's actions, and often appearing to perform as a seamless whole.

Consider, for example, the study by Foushee, Lauber, Baetge, and Acomb (1986) on the effects of fatigue on flight crew errors. These researchers found, to their surprise, that crews who had just logged several days flying together (fatigue condition) made significantly fewer errors *as teams* than well-rested crews who had not yet worked together. As expected, the fatigued individuals made more errors than others; however, functioning as teams they were able to compensate for these, presumably because they were better able to coordinate and to catch each other's mistakes.

Such differences in group behavior and performance have been examined by two research traditions in the social psychology literature. One research tradition has focused on social, affective, and unconscious influences on groups and their members, whereas another tradition is rooted in cognition, goals, and structures.[3] The former tradition finds its origins in the late 19th-century work of Gustave Le Bon, who proposed that unconscious processes of a crowd or group could manifest themselves in the actions of individual members (Alderfer, 1987). Similarly, the Hawthorne studies reported that emotions and tacit group norms could exert a greater influence on performance than working conditions or economic incentives (Roethlisberger & Dickson, 1939), and sociotechnical theorists later described the importance of a strong social unit in motivating workers (Rice, 1958; Trist & Bamforth, 1951). Finally, the intergroup perspective has explored how membership in identity groups (such as gender or race) or organizational groups (such as rank or function) can affect communication and motivation in task groups (Alderfer, 1987), which could lead to errors due to lack of coordination. When members of a work team communicate across tacit boundaries imposed by rank or identity group, this can inhibit the transfer of valid data (Argyris, 1985). Along these lines, nurses and physicians working as part of the same team (hospital unit) face identity group boundaries confounded with status differences that can affect within-team communication and thereby influence the process of administering drugs to patients.

The other research tradition emphasizes intellectual formulations such as structure and design, minimizing attention to emotional processes. The emphasis is on identifying conditions that help teams work together and solve problems (e.g., Hackman & Morris, 1975; Maier, 1967), such as early research that examined what leadership style would enable positive group outcomes (Lewin, Lippitt, & White, 1939). Recent work in this tradition has focused on the importance of practice (Senge, 1990) and members' interpersonal skills (Helmreich & Foushee, 1993) as influences on team performance. The structure of the team, the design of its task, and the supportiveness of the organizational context (reward, information, and educational systems) have also been central concerns in this tradition (Hackman, 1987). Both research traditions share the understanding that a group is more than the sum of its members and that group-level phenomena exist and influence task performance (Alderfer, 1987; Hackman, 1990). Although the design of the present study drew almost exclusively from the latter tradition, the results, as discussed more fully below, suggest that the story of drug errors cannot be told fully without attention to the constructs of the former.

Hospital units are one kind of work group, and group research may help explain differences in error rates. Hospital units are primarily managed and staffed by nurses, whereas physicians and pharmacists affiliated with each unit interact with the nursing

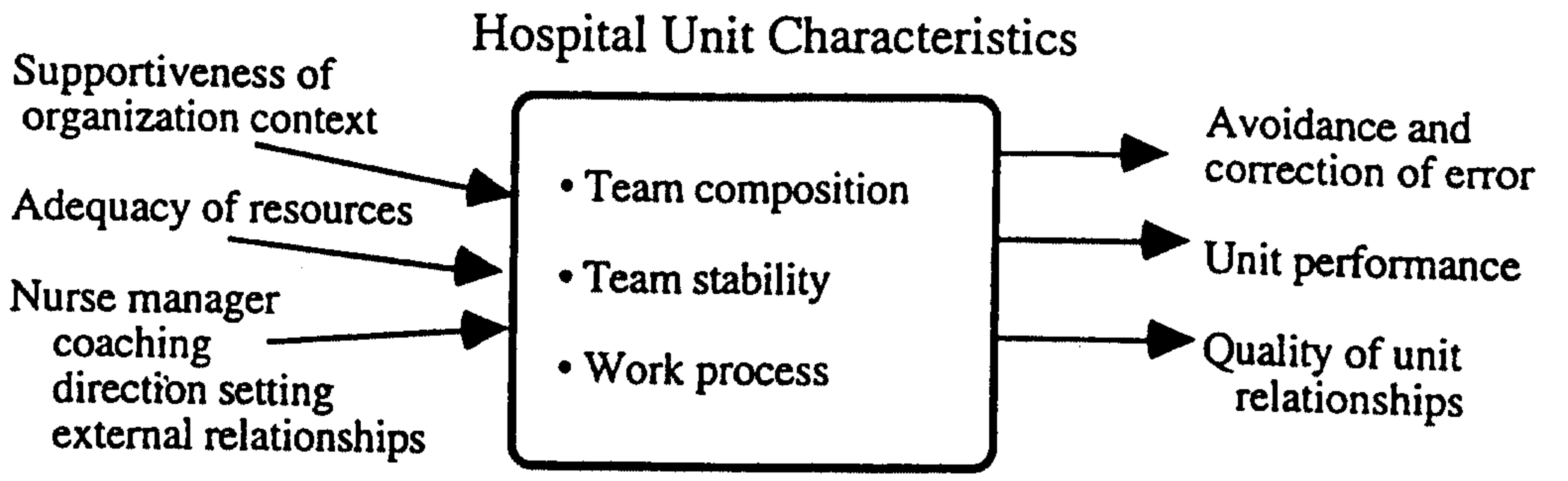

FIGURE 1: Model of Unit-Level Influences on Drug Error Rates and Other Performance Outcomes

teams. Preliminary observation revealed considerable flexibility in carrying out work processes; units vary in the way work is done and in how people work together. When and by whom the various tasks will be accomplished is left largely up to the discretion of the unit members, suggesting that critical performance outcomes such as the prevalence of drug errors can vary. Based on this perspective, the present study focused on hospital units as the unit of analysis in attempting to gain insight into causes and prevention of medication errors. The guiding question—rather than what causes people to make mistakes—was thus, are some work groups better than others at catching and correcting human error before it becomes consequential? Further, are such groups better able to learn from the inevitable errors that do occur and to avoid making the same ones in the future?

THE STUDY

Previous research suggested that error rates vary widely across units within the same hospital (Bates et al., 1993). Similarly, unit error rates obtained in the current study of drug complications range from 2.3 to 23.7 errors per thousand patient days. (Table 6, discussed in the Results section, displays the data for each unit.) What accounts for these differences? One possibility is that unit characteristics (such as team stability, norms, and work structure) influence error rates. Starting with the proposition that team behaviors are influenced by organizational context, team leader behaviors, task design, resource adequacy, and team composition (Hackman, 1987), the present study tests the hypothesis that error rates vary with these unit characteristics. Figure 1 depicts a model of proposed influences on, and outcomes of, hospital unit work processes. In this model, unit outcomes such as error rates and members' assessments of unit performance are group-level characteristics.

The research question guiding the present study was, are differences in work group (unit) properties associated with differences in error rates? And do members' perceptions of how their unit performs and of the quality of unit relationships vary with the "hard" data of drug error rates?[4] Thus the study explores the extent to which certain group-level properties contribute to understanding differences in drug error rates in hospital units.

Method

The study used a comparative nonexperimental research design. Eight hospital unit teams were randomly selected for study from two urban teaching hospitals affiliated with the same medical school. No known differences in unit composition, skill, professionalism, or workload existed between the two hospitals.

Three parallel data collection activities were conducted in the units studied. Independent of the present study, potentially harmful drug-related errors were identified over a 6-month period by trained medical investigators through daily chart review, daily informal visits to each unit to inquire about unusual drug events, and a confidential system to allow unit members to report incidents in writing. Previous research (Bates et al., 1993) found this combination of activities to uncover more errors than other methods.

Dependent Variables

Data were collected for four variables related to patient adverse drug events (ADEs), each expressed as a number of incidents per thousand patient days: (a) nonpreventable ADEs (adverse drug events unrelated to error), (b) preventable ADEs, (c) potential ADEs (or PADEs) (consequential errors that did not harm patients despite having the potential to cause injury, such as the heparin incident described above), and (d) interceptions (errors caught and corrected before reaching patients). The first two variables (nonpreventable ADEs and preventable ADEs) were measured by a combination of patient chart review and voluntary reporting. The second two variables (PADEs and interceptions) could only be measured through voluntary reporting of events by unit members. To capture the construct of human error, the primary dependent variable for the present study ("detected error rates") is the sum of preventable ADEs and potential ADEs.

Work Group Measures

To address the main research question of how unit properties influenced these outcome variables, two independent methods were employed (survey and observation) to develop a full picture of how the hospital units studied functioned and how they differed. First, the author developed a survey to assess the social and organizational properties of hospital units, based on a prior instrument designed to study the performance of cockpit crews (Hackman, 1990). The new survey measures include leadership behaviors of nurse managers, organization context (adequacy of training, information, and equipment), team characteristics (stability, composition, quality of unit relationships, and performance outcomes), and individual satisfaction and motivation. Figure 1 depicts relationships among these variables.

Next, another researcher, blind to both error rates and survey results, observed nursing teams in the eight units for several days each over a 2-month period; his goal was to observe behavioral dynamics in support of the larger project goal of understanding conditions surrounding drug errors.[5] He interviewed all nurse managers, several nurses, and members of the support staff in each unit, using a semistructured

TABLE 1
Summary Statistics for Survey-Based Measures

Scale	*Mean*	SD	*Internal Consistency Reliability*
Leadership behaviors			
Nurse manager direction setting	3.78	0.86	.90
Nurse manager coaching	3.60	0.88	.90
Nurse manager external relations	3.79	0.88	.90
Organization context			
Supportiveness of organization context	4.96	0.94	.74
Unit characteristics			
Team composition	4.43	1.15	.67
Team stability	4.76	0.95	.68
Unit outcomes			
Unit performance outcomes	5.06	0.93	.77
Quality of interpersonal relationships	4.12	0.84	.80
Individual satisfaction			
Internal motivation	6.14	0.71	.70
General satisfaction	4.63	0.57	.73
Satisfaction with work relationships	5.74	0.66	.68
Satisfaction with growth opportunities	5.51	0.85	.72
Single items			
"Drug-related errors occur frequently in this unit"	2.88	1.25	
"In this unit, drug-related errors are always reported"	4.43	1.54	
"If you make a mistake in this unit it is held against you"	3.24	1.20	

interview format that included open-ended questions to elicit interviewees' own descriptions of how mistakes are handled and perceived in their unit.

Sample

Eight units in two urban hospitals were included in the study. Five units from one, referred to as Memorial Hospital, include three intensive care units and two general care units. The three units from the other, referred to as University Hospital, are all general care units. Two hundred and eighty-nine surveys were distributed to nurses, physicians, and pharmacists identified by hospital personnel as full- or part-time members of the eight units. Fifty-five percent (159) of the surveys were completed and returned.[6] Surveys were filled out at the end of the second month of the 6 months during which drug errors were tracked.[7]

Survey Variables

Survey items include unit descriptions such as "This unit operates as a real team" as well as individual satisfaction items such as "I am satisfied with the amount of pay I receive." (Table 2 shows items for selected scales as illustration.) Each descriptive statement is followed by a 7-point Likert-type scale (7 = *strong agreement*, 1 = *strong disagreement*), except for the items describing nurse manager behaviors, which are followed by a 5-point scale ranging from *never* to *frequently*. As shown in Table 1, the

TABLE 2
Survey Items for Selected Scales

Selected Scale	*Items*
Nurse manager direction setting	The nurse manager takes initiatives to establish strong standards of medical excellence and professionalism in this team.
	The nurse manager sets clear goals and objectives for this team.
	The nurse manager is clear and explicit about what he or she expects from unit members.
	The nurse manager actively encourages nurses in this unit to stretch their level of performance.
Nurse manager coaching	The nurse manager takes initiatives to build the unit as a team.
	The nurse manager actively coaches individual unit members.
	The nurse manager shares leadership with other experienced members of the unit.
	The nurse manager is an ongoing "presence" in the unit—someone who is readily available.
Quality of interpersonal relationships	Members of this unit care a lot about it and work together to make it one of the best in the hospital.
	Working with the members of this unit is an energizing and uplifting experience.
	Some people in this team do not carry their fair share of the overall workload. (R)
	Every time someone attempts to straighten out a team member whose behavior is not acceptable, things seem to get worse rather than better. (R)
	There is a lot of unpleasantness among members of this team. (R)
Unit performance outcomes	Recently, this unit seems to be "slipping" a bit in its level of performance and accomplishments.
	Patients often complain about how this unit functions. (R)
	The quality of care this unit provides to patients is improving over time.
	Drug-related errors occur frequently in this unit. (R)
	This unit shows signs of falling apart as an organization.
	Attending physicians often complain about how this unit functions.
	In this unit, drug-related errors are always reported.

NOTE: (R) = reverse scored.

individual items combine to create 12 variables in five categories: leadership behavior, organization context, unit characteristics, unit outcomes, and individual satisfaction. (These categories map to the research model in Figure 1.) Single-item variables include members' perceptions of the frequency of drug errors, of the degree of reporting of such errors, and of the consequences of making a mistake in their unit.

Analytic Strategy

Preparatory analyses assessed (a) the adequacy of the psychometric properties of the new survey instrument, and (b) whether the survey scales were meaningful group-level variables. As the research questions examine relationships between unit attributes and unit error rates, it is important to establish that the survey variables in question are meaningful as unit-level properties (Kenny & LaVoie, 1985). Spearman rank order correlations were then used to examine substantive relationships between unit properties and detected error rates, interceptions, and nonpreventable ADEs. Finally, qualitative data were analyzed independently to understand ways in which unit climates and nurse behaviors differ, and units were ranked according to openness in discussing mistakes, a variable that surfaced during observation and analysis as differing across units. These qualitative results subsequently were compared to the quantitative survey and drug error data.

Survey psychometrics. Analysis of the psychometric properties of the survey results yielded satisfactory results, demonstrating the validity and internal consistency reliability of the scales as shown in Table 1. Internal consistency reliabilities range from 0.67 for team composition to 0.90 for each of the three nurse manager behaviors.

Unit-level properties. To assess the degree to which the organizational attributes measured by the survey are meaningful as group-level variables, two complementary measures of within-unit agreement about unit properties were computed. Measures of within-unit agreement also provide an indication of interrater reliability in measuring unit characteristics with this survey. The first, the intraclass correlation coefficient (ICC), uses one-way analysis of variance (ANOVA) to compare between- and within-unit variance. The second, the interrater reliability coefficient (IRR) derived by James, Demaree, and Wolf (1984), compares actual variance to a measure of "expected variance" to assess within-group agreement, without between/within comparisons.

ICCs for survey variables assessing unit characteristics were examined first; Kenny and LaVoie (1985) maintain that an ICC greater than zero indicates that a variable is meaningful at the group level. In many organizations, however, institutional forces restrict variance, leaving strong similarities across work groups, which suggests that allowing the magnitude of between-group differences to determine whether a variable is meaningful at the group level may not always be sensible in real organizational settings. For example, in these data, several variables designed to measure unit-level attributes have very low positive ICCs, such as team composition (.08) and unit performance outcomes (.06). Does this indicate lack of within-team agreement? Perhaps not. IRRs for the same variables (.80 and .88, respectively) provide a measure of within-group agreement that is unaffected by between-group similarities, and these values suggest high levels of within-team agreement (see Table 3). As many organizational features are similar across units, the critical sources of variance among units may be largely due to those features that are noticeably different, such as leadership behaviors of nurse managers.

The IRR is appealing as an absolute measure of within-unit agreement; however, its expected variance term is derived based on an assumption that survey responses are rectangularly distributed, thus posing a different limitation. Positive leniency in responding to items about unit performance is likely, which limits variance and can inflate estimates of agreement (see James, Demaree, & Wolf, 1984).[8] Thus response bias makes it difficult to assess absolute levels of unit agreement. However, examining both coefficients shows which variables have greater between-group differences while

TABLE 3
Intraclass Coefficients (ICC) and Interrater Reliability Coefficients (IRR) Compared Across Units

	ICC	*Median IRR*
Supportiveness of organization context	.12*	.85
Nurse manager direction setting	.22*	.88
Nurse manager external relationships	.21*	.89
Nurse manager coaching	.18*	.82
Team composition	.08*	.80
Team stability	.10*	.83
Unit performance outcomes	.06	.88
Satisfaction with growth opportunities	.04	.91
"Drug-related errors occur frequently"	.08*	.67

*One-way ANOVA for each of these variables by unit is significant ($p < .05$).

having similar within-group agreement levels. To facilitate comparison of the two kinds of unit-agreement coefficients, ICCs and median IRRs are shown for eight variables in Table 3.[9]

The ICC's between/within comparison is useful in showing which variables stand out as attributes that vary across, and hence best distinguish among, units. For example, nurse managers' direction setting (.22) has the highest ICC in the sample, reflecting the greater variance between units in perceived behavior of nurse managers, compared to variables describing less tangible unit characteristics such as performance outcomes (ICC = .06). ICCs for each of the three leadership scales are significant ($p < .001$). This result is consistent with the author's on-site observations of the nurse managers' behaviors; each unit has a different nurse manager, in some cases with strikingly contrasting management styles. In contrast, the variable "unit performance outcomes" has low between-unit variance (ICC = .08) despite having the same level of within-unit agreement (IRR = .88) as nurse manager direction setting.

In sum, IRRs and ICCs provide different, complementary measures of within-group agreement. In these data, the IRRs reveal the high within-unit agreement levels for variables, such as unit performance outcomes, which lack high between-unit differences and thus have low ICCs. Together these coefficients allow us to have confidence that these measures are meaningful as group- or unit-level variables.

RESULTS AND DISCUSSION

As discussed above, this study employed two distinct methods to investigate organizational factors that may account for variance in drug error rates across hospital units. In this section, I first review the quantitative results, and then present results from the qualitative research. These two lenses provide distinct and complementary pictures of the phenomenon.

TABLE 4

Organizational Influences on the Detection of Error: Spearman Rank Order Correlations Between Survey Variables and Drug Error Variables

	Correlation
Detected error rates and . . .	
Nurse manager coaching	.74*
Nurse manager direction setting	.74*
Unit performance outcomes	.76*
Quality of unit relationships	.74*
Willingness to report errors	.55
Mistakes (are not held against you)[a]	.44
Intercepted errors and . . .	
Nurse manager coaching	.71*
Nurse manager direction setting	.83*
Unit performance outcomes	.71*
Quality of unit relationships	.76*
Willingness to report errors	.62
Mistakes (are not held against you)	.45
Nonpreventable drug complications and . . .	
Nurse manager coaching	–.10
Nurse manager direction setting	–.09
Unit performance outcomes	.11
Quality of unit relationships	–.07
Willingness to report errors	.09
Mistakes (are not held against you)	.07

NOTE: All correlations are between unit means for each survey variable and the dependent drug error variable.
a. The survey item that reads "If you make a mistake in this unit it is held against you" is reverse scored in analyzing the data.
*$p < .03$, two-tailed.

Relationships Between Error Rates and Unit Characteristics

Correlational analyses of the relationships between unit characteristics and error rates yielded unexpected results. We expected to find higher error rates to be associated with lower mean scores on perceived unit performance, quality of unit relationships, and nurse manager leadership behaviors. Exactly the opposite result was found. As shown in Table 4, detected error rates are strongly associated with high scores on nurse manager direction setting ($r = .74$), coaching ($r = .74$), perceived unit performance outcomes ($r = .76$), and quality of unit relationships ($r = .74$), with $p < .03$ in each case. These relationships between nurse manager behaviors and the independently collected error rates are noteworthy, and at first glance an odd result indeed. Do better coached teams make more mistakes?

An alternative interpretation of these findings also merits careful consideration: In certain units, the leaders may have established a climate of openness that facilitates discussion of error, which is likely to be an important influence on detected error rates. Awareness of the labor intensiveness and difficulty of tracking drug-related errors—

and of how easy and natural it is for human beings to underreport error—suggests that measuring errors in each unit is not a trivial undertaking. The difficulty of assessing actual error rates is shown in a recent study in which a hospital in Salt Lake City was able to increase the number of ADEs identified forty-fold after instituting a new computing system to predict and track errors (Evans et al., 1992). The magnitude of this increase indicates how few of the errors made in hospital units are reported, and suggests also that variance in error rates caused by differences in patient severity and complexity across general and intensive care units can be overwhelmed by the influence of lack of reporting of errors (see Note 10). In short, detected error rates are a function of at least two influences—actual errors made and unit members' willingness to report errors. Moreover, in organizational settings in which errors are consequential, willingness to report may be a greater influence on the error rates obtained than is variance in actual errors made. These observations suggest that positive correlations between error rates and nurse manager coaching, perceived unit performance, and quality of unit relationships may be explained by examining the role of members' willingness and ability to catch and report drug errors.

The Role of Willingness to Report Errors

Higher detected error rates in units with higher mean scores on nurse manager coaching, quality of unit relationships, and perceived unit performance may be due in part to members' perceptions of how safe it is to discuss mistakes in their unit. Several other survey variables provide support for this hypothesis. First, a variable labeled "willingness to report errors" was computed for each unit; this is the difference between two item means—"in this unit, drug errors are always reported" and "drug errors occur frequently in this unit." This difference score, which ranges from –4 to 5 in these data, measures perceived willingness to report errors controlling for perceived frequency of error. Detected error rates then were found to be correlated with willingness to report errors ($r = .55$). Second, a single-item variable measuring unit members' perceptions that making a mistake in their unit will not be held against them is also correlated with detected error rates ($r = .44$; see Table 4). Moreover, one-way ANOVA reveals significant between-unit differences for both of these single-item variables ($p < .05$), suggesting that unit climates vary significantly in perceptions of the risk of discussing mistakes.

Examining relationships between these two survey variables and another dependent variable, the number of interceptions in each unit, provides additional support for the hypothesis that willingness to report error influences ability to detect errors in a unit (see Table 6 for means and standard deviations of interceptions and other drug error data). Recall that the concept of self-correcting performance units highlights the role of catching errors before they become consequential; the number of interceptions in a unit provides an indication of unit members' attentiveness to their interdependence in caring for patients, as well as a measure of their ability to function as a self-correcting performance unit. Analysis of the data shows that units with a greater number of interceptions (team self-correcting behaviors) also tend to have relatively higher scores on nurse manager direction setting ($r = .83$), coaching ($r = .71$), performance outcomes