

An Ethogram to Quantify Operating Room Behavior

Laura K. Jones, PhD¹  · Bonnie Mowinski Jennings, PhD, RN, FAAN² · Ryan M. Goelz, BS^{3,4} · Kent W. Haythorn, MSN, RN⁵ · Joel B. Zivot, MD^{6,8} · Frans B. M. de Waal, PhD⁷

Published online: 26 January 2016
© The Society of Behavioral Medicine 2016

Abstract

Background The operating room (OR) is a highly social and hierarchical setting where interprofessional team members must work interdependently under pressure. Due primarily to methodological challenges, the social and behavioral sciences have had trouble offering insight into OR dynamics.

Purpose We adopted a method from the field of ethology for observing and quantifying the interpersonal interactions of OR team members.

Methods We created and refined an ethogram, a catalog of all our subjects' observable social behaviors. The ethogram was then assessed for its feasibility and interobserver reliability.

Results It was feasible to use an ethogram to gather data in the OR. The high interobserver reliability (Cohen's Kappa coefficients of 81 % and higher) indicates its utility for yielding

largely objective, descriptive, quantitative data on OR behavior.

Conclusions The method we propose has potential for social research conducted in healthcare settings as complex as the OR.

Keywords Health communication · Health services research · Worksite health

Because a surgical team's underperformance can be costly in terms of patients' lives and financial liabilities, hospitals are seeking to better understand teamwork [1–5]. Decades ago, Donabedian [6] identified two elements comprising clinical performance—technical (i.e., knowledge and skill) and interpersonal (i.e., information exchange via communication). Today, quality officers are interested in understanding both the technical and interpersonal interactions of clinicians [7–10]. The dynamics of non-technical, interpersonal interactions such as non-protocol driven social behaviors ranging from pleasantries and gossiping to flirtation and verbal confrontations, remain largely unknown. Knowing these social behaviors is an important step toward enhancing teamwork. For this article, we focus on social behaviors that are relatively common within the operating room (OR).

Teamwork requires the promotion of constructive and the mitigation of destructive social behaviors [11]. Although the dynamics are not fully understood, it is known that gossip, exclusion, threats, and unjustified criticism instigate friction [11]. In one study, the authors reported that an average of four conflicts erupt among the OR team per procedure [12]. Moreover, conflicts may initiate errors that increase the morbidity and mortality of patients [13, 14]. Estimates indicate medical errors are responsible for at least 98,000 avoidable deaths annually in the USA alone [15], and 60 % of all medical errors

✉ Laura K. Jones
lkjones@emory.edu

¹ Emory University, Atlanta, GA, USA

² Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, GA, USA

³ Emory Healthcare Surgical Services, Atlanta, GA, USA

⁴ South University Anesthesiologist Assistant Program, Savannah, GA, USA

⁵ Emory Healthcare Surgical Services, Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, GA, USA

⁶ Department of Surgery, Emory University School of Medicine, Atlanta, GA, USA

⁷ Living Links, Yerkes National Primate Research Center and Professor at Emory University, Atlanta, GA, USA

⁸ Department of Anesthesiology, Emory University School of Medicine, Atlanta, GA, USA

occur in the OR [13]. Emotional abuse, a possible consequence of conflict, may impair clinicians' mental and physical health [16, 17]. The Joint Commission, which accredits US hospitals, affirms that discordance and distractions, behaviors that leads to conflict within OR teams, are hazardous to the well-being of both patients and their care teams [14, 18].

Despite the need to better understand social behavior in biomedical settings, methodological challenges as fundamental as being granted access to care settings restrict much needed behavioral research. Because the OR environment is sterile and confined, it is a particularly inaccessible space, rendering certain approaches to behavioral research profoundly flawed and/or difficult to execute [19]. To date, investigators have used surveys, interviews, videography, and small-scale, narrowly focused observational methods to investigate OR behavior [20–27]. Such approaches to data collection are needed yet, with the exception of videography, they may not provide a full picture of OR behavior inasmuch as events are filtered through the OR team's own perceptions.

Here, we propose a method for examining the balance (or imbalance) between behavior that promotes or signals cooperation and behavior that may disrupt cooperative outcomes in the OR. With this new method, we are able to examine the whole spectrum of social behaviors, from the most constructive to the most destructive. Currently, we cannot address the influence of this behavior on patient safety, but we believe patient safety is at stake whenever conflict arises in the OR.

Limitations of Existing Methodologies

Surveys are a common means of gathering data on behavior in the OR. It is a challenge, however, to avoid response bias. Given the potential legal liability of OR team members, it is reasonable to expect distorted responses on the critical issue of their own behavior in the OR [28, 29]. According to medical anthropologists, physicians tend to respond to self-reports so as to appear competent [30].

Post hoc interviews are a common means through which social and behavioral scientists have attempted to examine OR behavior [31]. Interviews conducted outside of the OR, however, lack the raw truthfulness that real-time, in situ interviews might provide. In situ interviews, however, would likely cause distractions during a surgical procedure and therefore are not recommended. OR team members might be able to participate in interviews between cases, known as room turn-over in the OR vernacular. Interviews conducted during the room turn-over, when re-sterilization and restocking occur, would not only be retrospective but would also be curtailed and not confidential because OR team members generally continue working through this interval [21]. Interviews could occur before or after a shift but would probably have to occur away from the OR.

Participant observation is difficult to conduct in the OR because of information sensitivity. Furthermore, Institutional Review Board (IRB) and hospital-imposed equipment restrictions are also in place. Although select researchers have used audiovisual recorders in the OR [25, 26], such recordings are often prohibited to protect patient privacy, as outlined by the Health Insurance Portability and Accountability Act [32, 33]. Thus, generally, observers must compose handwritten field notes that may result in fragmentary documentation of the oftentimes chaotic OR setting [19].

Investigators relying on surveys and interviews strive to overcome these obstacles by conducting multimethod and/or ethnographic studies. Such research engages two or more types of data to enhance the validity of each. Closest to a multimethod study in the OR, Lingard et al. [22] paired observations with coded interviews, to quantify OR behavior without relying on self-reports. The Lingard team tracked 48 OR procedures to examine communication failure by counting the incidence of four types of communication breakdowns among OR team members: (a) occasion (i.e., communication occurring before or after the appropriate time), (b) content (i.e., loss of important information), (c) audience (i.e., absence of individuals necessary to the conversation), and (d) purpose (i.e., appropriate response; [22]). Missing from the counts, however, was information about which team members were involved in the (mis)communications. Moreover, the study was conducted using a pen and paper method without time-stamping or interobserver reliability testing.

The strategy we propose is to document OR behavior as it occurs in real-time, using quantifiable and replicable methods. We borrow a technique widely used not only for the observation of animal behavior but also applied to human interactions, known as the ethological method (from "ethology:" the biological study of human and animal behavior in natural or naturalistic settings). This method is open to accuracy assessment and promises to fill a critical gap in our knowledge by focusing on actual behavior in the OR rather than the participants' interpretation or recall of their own behavior.

The Ethological Method

The field of human ethology was founded by Eibl-Eibesfeldt who postulated that humans are not so evolved as to no longer be animals; hence, ethological methods may be applied to document their behavior [34]. He himself used this quantitative methodology to measure behavior in preliterate human societies, as have evolutionary anthropologists and behavioral ecologists [35–37]. Anthropologist Cloak argued that cultural anthropologists as well should consider using ethological methods [38].

Human ethology has been used in medical settings, too. For example, psychiatrists studying patients with mental

health problems have found ethograms useful for tracking the course of their patients' illness [39, 40]. While developing these tools for behavioral assessment, psychiatry has not, to the best of our knowledge, problematized issues of construct and content validity. In another medical context, the ethological method has been used to look for behavior patterns demonstrative of pain in neonates [41, 42]. Similar methods have not been applied to systematically record social behavior in the OR.

Ethologists start by creating an “ethogram”—an inventory of carefully defined behavior patterns that can be recognized and reliably documented. Ethograms are typically designed around research questions and the variables that researchers deem relevant to address these questions. The behavior patterns of interest are defined in an explicit manner to ensure the method of data collection is replicable. In studying human behavior, an ethogram may include physical interactions (e.g., shaking hands), vocalizations (e.g., shouting at someone), or verbal content (e.g., complimenting another person). Despite their specificity, ethograms are flexible and can be reused in the future. For instance, a research team may change its interests over time, perhaps increasing the detail of the observations in a given area, yet an original ethogram may serve as the basis for the increased detail. Similarly, published ethograms are often adapted and/or modified by others [43, 44].

The OR ethogram was adapted from existing ethograms used for primatological studies of bonobo and chimpanzee communication (Table 1; 43). These primate ethograms cover a wide range of behavior, from locomotion (e.g., running, walking, climbing) and foraging (e.g., the kind of edibles consumed) to social behavior (e.g., playful interaction or grooming). Social behavior is classified using broad categories, such as aggression, play, sex, and affiliation, with special attention to communication patterns. Typically, behavior is recorded in the triplet format of “Who does What to Whom.” The behavioral descriptions are kept as neutral as possible, without interpretation or suggested function. The intent of

the ethogram is to document behavior, leaving interpretation to take place later, during data analysis.

The specific methodology we followed while studying OR teams was a mixture of what Altmann terms “all-occurrence sampling” and “scan sampling” [45]. With all-occurrence sampling, the observer records every observed instance of a behavior using the ethogram. For example, primatologists studying conflict may use an ethogram to record each facial expression and vocalization defined as “aggressive” [46]. All-occurrence sampling is typically used for “point behaviors,” those with a short duration that occur at seemingly random intervals. Scan sampling, by contrast, is used to capture behavioral states with longer duration (e.g., the location of a subject) that are recorded at pre-set time intervals (e.g., every 10 min).

The OR ethogram we describe is not meant to be the product of more classical instrument development but rather was fashioned after typical primatological models [47]. Development of the ethogram can also be likened to observational studies conducted by Coiera et al. who assessed communication in the emergency department [48]. They developed the “Communication Observation Method (COM) . . . to measure communication patterns within clinical organizations” [49]. Similarly, we developed the OR ethogram to measure social behavior in the OR. To enhance the content validity of the ethogram, we consulted with two experts—an anesthesiologist (JBZ) and the vice president of surgical services for the healthcare system/former OR nurse (KWH)—in its creation and revision.

Methods

Sample and Setting

The ethogram was developed as part of an investigation of non-simulated cooperation and conflict in the ORs of three large, urban, academic hospitals, all of which are part of a

Table 1 Excerpt from a chimpanzee/bonobo ethogram

Behavior	Definition
Silent teeth-baring	“Retraction of the lips resulting in partial or complete exposure of the teeth and gums, with mouth practically closed and without vocalizing. The face is oriented toward the partner but the eyes may make evasive movements. Subtypes of the pattern can be distinguished dependent upon the amount of teeth exposure, and upon whether the lips arc curled outward or pulled flat.” p.190
Play face	“The mouth is opened with the lips either a) in a relaxed position, covering the upper teeth completely and the lower teeth partially, or b) retracted, without however pulling back the mouth corners, resulting in baring of both the upper and lower frontal teeth. The second variant occurs both silently and in combination with the ‘panting laugh.’ The two variants alternate and blend so frequently that it is hard to draw a line between them.” p. 197

Note: This excerpt was extracted from de Waal's “The communicative repertoire of captive bonobos (*Pan paniscus*), compared to that of chimpanzees.” [47]

multi-hospital healthcare system. Many team members rotated among these three hospitals, located only a few miles apart. Most of the attending surgeons were on the faculty at the academic institution. Some of the senior surgeons, however, worked primarily at community hospitals. In total, social behavior in the OR was recorded for 200 procedures. The procedures involved over eight surgical subspecialties (Table 2). Sixty-eight percent were open procedures and 32 % were laparoscopic.

Participants included surgeons, surgical residents, anesthesia providers (anesthesiologists, anesthesiologist assistants (AAs), certified registered nurses anesthetists (CRNAs)), registered nurses (RNs who were circulating or scrub nurses), and surgical technicians. One hundred twenty-one different attending surgeons participated, between one and four times each, with always at least a week between observed cases. No attending surgeon was observed more than four times to maximize the diversity of surgeons in the sample. The attendance of other team members was unpredictable and could not be used when selecting cases to observe. There were at least 87 surgical residents, 85 different anesthesia providers as well as 83 different circulating nurses and 61 different scrub nurses or surgical technicians. Exact numbers for all groups were difficult to derive because these team members' schedules were inaccessible, they generally wore face covers, and they seldom used last names.

Our methodological exploration was supported by the highest levels of administration at the three hospitals because of the organizational commitment to improve quality and reduce OR errors. The study protocol was reviewed by the IRB for this university/hospital system, which determined the study to be exempt research. Patients were not a part of the data collection, and thus, the IRB deemed their consent unnecessary. Still, at one hospital, it was requested that a member of the surgical team acquire verbal assent from the patient.

Although not IRB-stipulated, the researchers also obtained verbal assent to observe each team member who was in the OR just prior to the start of the procedure. Of the 200+

procedures that we attempted to observe to assess the ethogram, team members declined to be observed 2 % of the time. When team members declined participation, the researcher left the OR without collecting data and did not return during that procedure. The only data saved of these procedures were field notes describing the environment of the OR prior to the arrival of the team member who did not give assent, as well as field notes indicating the reason why permission to observe was declined or revoked.

Data Collection

Preparing to Construct the Ethogram

During this phase, the primary observer (LKJ), who had no previous experience working in the OR, learned its unwritten rules and determined study parameters. The rules she learned were those that are largely taken for granted by those who work there, such as the: (a) obligatory wearing of attire called "scrubs," as well as shoe and hair covers; (b) prohibition of food or beverages in the surgical suites or the hallways; and (c) prohibition against touching certain areas, such as the scrub nurse/technician's table, unless one is sterilely gowned and gloved. The observer also learned that the OR team preferred that she sit on a stool near the circulating nurse who had a computer station, a position that allowed the observer to see procedures while remaining unobtrusive. She was moved to another location if space became constrained. In addition, the circulating nurse, the guardian of the OR, could watch her and prevent mistakes (e.g., touching sterile equipment).

Sometimes team members, most commonly the circulating nurse, would talk to the observer, initially asking questions and providing supervision. Once familiar with the observer, the conversations between the nurses and observer increasingly featured friendly chatting. The observer would be polite but quiet, doing her best to minimize distraction.

Along with learning about the OR culture, the primary observer established various study parameters. For example, selection of which operating rooms to observe was based on the procedures and the attending surgeon. The goal was to derive the most robust sample distribution possible, using a form of maximum variation sampling to observe the widest variety of care teams and surgical procedures [50]. Likewise, the beginning and ending of the OR procedure were defined. Data collection began when the first incision was made and ended when the drapes were removed. Once the drapes were removed, too many new team members entered the room to accurately track their behavior. Postoperative handoffs are important yet they were not a goal of this study, as they have been studied meticulously elsewhere [51, 52].

In addition to becoming familiar with the OR and establishing study parameters, the primary observer initiated *de novo* observations to develop a list of non-protocol driven

Table 2 Cases observed

Department	Total
General	40
Neurology	29
Gynecology	24
Cardiothoracic	23
Otolaryngology	18
Urology	17
Orthopedics	15
Vascular	14
Other	20
	200

social behaviors. This was accomplished by doing a month of informal observations and recording all typical social behavior, from small talk to confrontation, from playfulness to courtship. She kept recording behaviors until no new or unique behaviors were observed, thus achieving a form of data adequacy or saturation [53]. These behaviors then served as the basis for designing the ethogram.

Designing and Using the Ethogram

We began developing the OR ethogram by organizing the non-protocol driven social behaviors seen during preliminary observations of OR team members into 10 behavioral categories (Table 3). Category development was based on grouping similar actions into the same sets. Each of the 10 behavioral categories included two to four subgroups that listed specific actions, making for a total of 28 specific actions (Table 4).

We developed instructions for using this behavioral index: First, observers indicate the intensity of the action with “1” reflecting the lowest level of intensity and “4” as the highest (Table 4). Second, the observers use the behavioral categories to collect data using the format: A does X to B, while always recording the highest level of action intensity if multiple levels within the same category occur during the same interaction. Third, instructions stipulate that observers re-document the same behavior directed at the same target only after an interruption without the same behavior of 3 min or longer. For example, suppose a surgeon yells a profanity at his or her resident at 8:03:09. If the behavior is repeated at 8:04:58 (under the 3-min margin) at the same intensity level, this second instance is not recorded. The same behavior occurring at 8:06:09 or after is recorded as a new instance, however. The same behavior occurring at a lower intensity within the 3-min margin is recorded.

Additional instructions include how to document various types of interactions. If there are multiple independent interactions occurring simultaneously from different categories or within the same category but at varying intensities, then each is recorded. Conversely, if different behaviors within the same category but at varying intensities are observed, the highest-ranking category in the ethogram is recorded. For example, if the circulating nurse and the scrub nurse are engaged in friendly conversation while the attending surgeon is teaching his/her resident, two different interactions are recorded. If these events occur too quickly in succession for the observer to record, the observer records the behavior with the greatest potential social impact and/or intensity first. For example, if a subset of the team is discussing their weekend plans while another subset is having an argument over equipment needs, the observer prioritizes his/her attention to the latter. In cases of ambiguous behavior that could be misinterpreted, for example, due to face masks that cover facial expressions and muffle voices, observers refrain from recording the interaction. All interlocutors in the room contend with these problems and generally find ways to compensate for lost cues.

The ethogram is designed using pre-assigned letter codes for all 28 behaviors (Table 4) and each of the 7 most common types of OR team members (Table 5), with upper case letters used for persons (e.g., “S1” designated the attending surgeon) and lower case letters for behaviors (e.g., “c”) with the accompanying intensity level (e.g., 1 indicated low level conflict). The attending surgeon and his/her assistants are distinguished (i.e., “S1” or “S2”). Everyone in the same role is given the same code (Table 5). If there is more than one circulating nurse in the room, for instance, both are coded as “N1.” The names of subjects are recorded to preserve anonymity. To document that the

Table 3 Operating room ethogram categories

Behavioral category	Definition ^a
Confrontational	Confrontational, aggressive, or insulting behavior
Submissive	Non-confrontational, submissive or appeasing responses to confrontational behavior
Protesting	Non-submissive reactions to confrontational or aggressive behavior
Playful	Playful, happy interaction
Mudslinging	Negative verbal behavior about third parties that indicates conflict or may lead to conflict
Friendly	Outgoing, affiliative behavior that is not as casual as “Playful Behavior”
Supportive/teambuilding	Behavior that is generally affiliative and helps the team to function as a team
Disintegrating	Behavior that threatens the functioning of the team
Courtship	Interactions of a flirtatious, sexual, or romantic nature
Reaction to courtship	Acceptance level of the above courtship behavior

^a For the precise content, see the full ethogram in Table 4

Table 4 Full operating room ethogram

Behavioral category	Code/intensity	Example
Confrontational	c1	Interrupting or ignoring another who is talking. Acting curt, irritated, demanding, being dismayed, or raising one's voice while talking to the other. Calling someone by their function, such as "Anesthesia" or "SOSI (Surgical Operating Services, Inc.)"
	c2	Grabbing equipment. Taking supplies from another when they are not offering them, or taking them from another in a hasty manner demonstrating impatience. Stomping
	c3	Cursing, yelling, name-calling, use of demeaning titles (e.g., "Dude," "Miss")
	c4	Making it clear that one does not want to work with others (e.g., asking for a mediator or superior). Storming out of the OR
Submissive	r1	No reaction to conflict-provoking behavior. Ignoring or acting as if unnoticed. Becoming quiet (conversation stops abruptly)
	r2	Using formal titles, being apologetic ("I'm sorry, sir")
	r3	Cowering, moving away to create distance, slumping, or bowing in posture
Protesting	p1	Head shaking, eye-rolling, or other gestures of disagreement behind the back of the initiator of a confrontation, verbally stating disagreement with the initiator
	p2	Gestures of protest visible to the initiator, raising one's voice while verbally disagreeing with the initiator
Playful	w1	Joking, making funny remarks, laughing with others. Lighthearted conversation
	w2	Singing, dancing, humming, whistling, shooting hoops (i.e. throwing garbage into trash from afar and celebrating if it goes in)
Mudslinging	m1	Bad-mouthing an absent non-team member to an ally (e.g., calling the blood bank technologists lazy), making fun of the anesthetized patient (e.g., "She's gigantic. We're going to need ten people to move her")
	m2	Bad-mouthing an absent team member (e.g., the anesthesiologist complains about the surgeon's tardiness)
	m3	Bad-mouthing a present team member to an ally
Friendly/solicitous	f1	Giving pleasantries (e.g., "Thank you!"), friendly greetings (e.g., "How are you?"). Polite conversation with interest in the other
	f2	Personal compliments (e.g., "Good work, doc!"). Prolonged non medical conversation that does not include joking or laughing (including innocuous gossip)
	f3	Patting on the back, hugging
Supportive/teambuilding	t1	Roll call, assembling
	t2	Teaching (e.g., telling medical student/fellow/etc. what to do to the patient, how the limbic system works, how to pass board exams)
	t3	Motivational talk, complimenting a team member or the team in general for their teamwork, debriefing (e.g., "Good job today, team")
Disintegrating	d1	Remarks that suggest a team member does not want to be in the room. Discussing quitting, retirement (e.g., "I can't wait until I'm 65. Just one more year of this place!")
	d2	Quitting, calling in a mediator or supervisor
Courtship	x1	Flirting: behaving in a manner that increases the intimacy between two individuals one or both of whom seem to be sexually interested (e.g., a smiling/giggling technician teasing a nurse about his/her unflattering scrubs)
	x2	Telling jokes with sexual content, light touching
	x3	Prolonged or frequent touching, stroking of the other's lower back, waist, thigh, up and down the arm, neck, shoulders. Using pet names (e.g., "Baby")
Reaction	y1	Accepting, reciprocating, or favorably acknowledging (e.g., to Courtship smiling back)
	y2	Neutral (the flirtee neither reciprocates nor seems reluctant)
	y3	Rejecting the overtures by not reciprocating or smiling, reproaching the other, or moving away

Note: Each behavior and the subjects involved were recorded in real-time

attending surgeon (S1) cursed (c3) at the scrub nurse (B1), the observer records "S1 c3 B1." The "3" in "c3" represents a high intensity conflict.

In addition, the observers collect 10-min scan samples to record the general atmosphere in the room as hyper (o), neutral (u), or negative (g; Table 6). "Hyper" indicates that the OR

Table 5 Codes for operating room team roles

Code	Definition
S1	Attending surgeon
S2	All other surgeons (usually fellows and residents), surgical physician assistants, medical students at the operating table
N	Circulating nurse
A	Anyone from the anesthesia team
B	Scrub nurse or surgical technician (manages surgical equipment and supplies at the bedside during a procedure)
Z	Anybody who was not with nursing, surgery, or anesthesia (radiology technician, medical student who was not at the operating table or doing anesthesia, perfusionist, medical sales representative, maintenance, blood bank technician, phone caller or texter, trainee who was just observing, all idle observers, information technology technicians, anyone that could not be categorized with the other subject codes)
99	Recipients from multiple services (e.g., a joke by the attending surgeon intended to be heard by the scrub nurse and the fellow, or the whole room)

environment is rowdy. The team members are highly engaged in social behavior, such as gossiping, joking, or singing. “Neutral” indicates a calm state in which technical and social behaviors are in balance, and everyone is focused on the procedure. “Negative” indicates tensions and/or open conflict, adversely affecting the atmosphere in the room. The observers use the ethogram to document all occurrences of the behaviors and interim social atmosphere conditions within a surgical procedure by typing the correct code for the behavior and the team member into an iPad; the software (Neukadye’s Timestamped Field Notes™; Fig. 1) automatically timestamps the entry at the moment the enter key was hit.

These study procedures were pilot-tested by sampling 10 surgical procedures representing various surgical departments (e.g., cardiothoracic surgery, general surgery). Data were collected using the same equipment we planned to use in the study: an iPad, a pen, and paper. Once the study procedures were solidified, we proceeded to tests of interobserver reliability, a major goal of this study.

Examining Interobserver Reliability

Interobserver reliability (IOR) is a means by which observational data, which may be influenced by subjective interpretation, can be assessed for reliability or consistency. IOR may demonstrate that agreement among multiple independent observers is not due to chance. Acceptable IOR

Table 6 Scan sample codes collected every 10 min

Code	Description of operating room atmosphere
u	Normal, neutral, professional, appropriate
o	Hyper, fun, lively, distractive
g	Negative, tense, stressful, agitated

is established if independent observers are able to apply the same ethogram to the same observed behavior and report consistent sets of events occurring in real-time or via video recording [54].

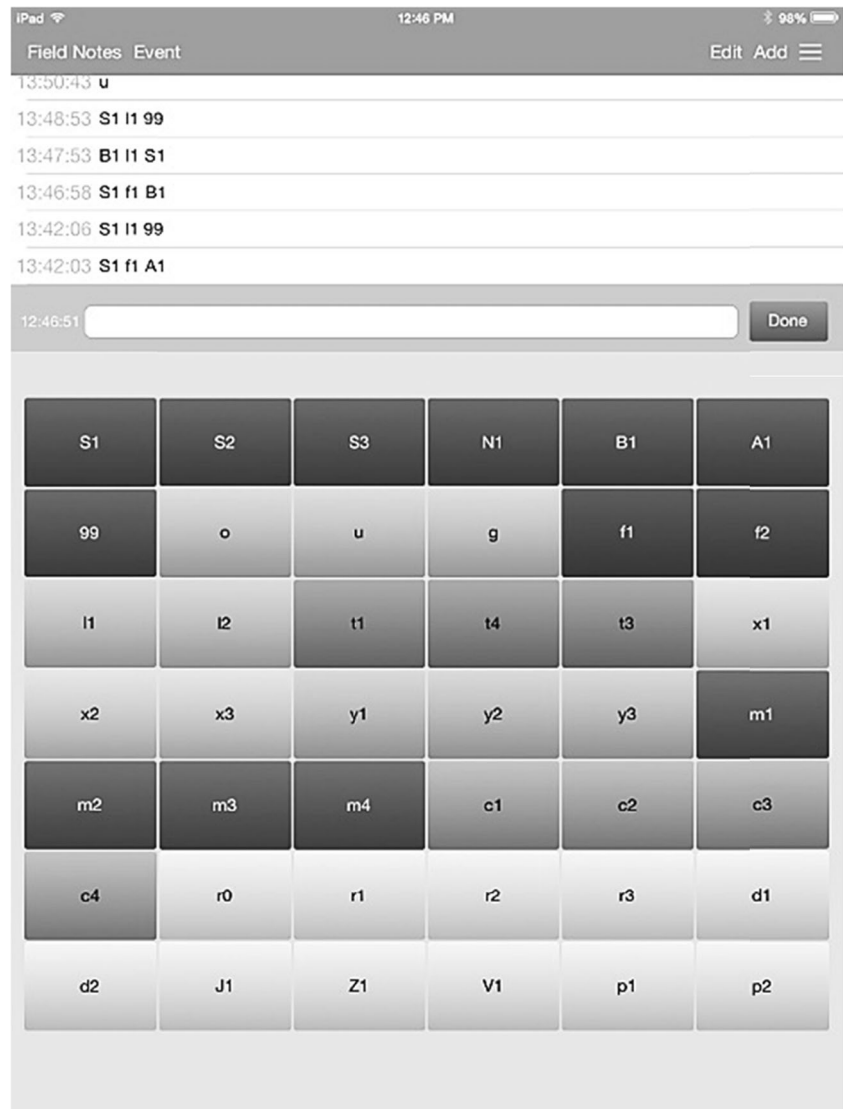
A second observer was engaged to work with the primary observer in assessing IOR. After a period of training, the two observers (LKJ and RMG, whose backgrounds were in the social sciences and the health sciences, respectively) collected OR data simultaneously from different angles while avoiding speaking or looking at each other. Their iPads (on which the ethogram’s codes were displayed) were enclosed in extra-large binders to guarantee the independence of their observations.

The two observers jointly collected behavioral data on 537 min of OR procedures, during which 483 mutually observed social interactions were recorded. Cohen’s Kappa coefficient (a measure of agreement corrected for chance levels) was 89 % for the recorded actor identities and 81 % for the behavior patterns [55]. Kappa’s of this magnitude are considered “almost perfect” [56].

Discussion

The achievement of such high IOR indicates the reliability of using the ethological method to assess social behavior in the OR. Previous OR studies in which behavioral data were recorded invested far less observation time and provided no rate of IOR [22, 57]. There are no official criteria of how many observations are needed to establish IOR, but we believe our sample size to be adequate. IOR was based on nearly 500 jointly observed social interactions, with the interactions constituting the unit of analysis. Our approach was modeled after that of de Waal [47] whose ethogram is the standard for our closest primate cousins. The high level of interobserver

Fig. 1 The image above is an iPad Screen Shot of Neukadye's Timestamped Field Notes™ (<http://www.neukadye.com>). The top of the screen shot lists timestamps corresponding to each coded event. The bottom section of the screen shot shows the keys for each code in the ethogram. The shading is for aesthetic purposes and has no meaning beyond its utility for the coders



agreement indicates the reliability of the method for other social and behavioral scientists conducting studies in healthcare settings as complex as the OR.

Our findings furthermore demonstrate the feasibility of using an ethogram to record OR social behavior. Feasibility was conceived before the study as comprising two elements: (a) our ability to get permission to observe the majority of procedures we requested to observe; we were granted permission 98 % of the time; and (b) our ability to apply the methods consistently and reliably for 200 procedures. Given the significant obstacles to conducting behavioral research in the OR, feasibility was not guaranteed. We therefore believe that our experience demonstrates that it is feasible to be granted permission and apply ethological methods for using the ethogram in the OR.

Along with the value in the method described, we do recognize inherent limitations. We cannot guarantee that

our presence did not have a Hawthorne effect that may have altered the team members' behavior [58], although the magnitude or even the existence of this effect remains under debate [59]. Many OR team members initially expressed a desire to demonstrate their "best behavior," which may have reduced the amount of conflict observed. It should be noted, however, that very soon OR team members began to ignore the observers, barely paying attention to them. The behavioral effects of observer presence, if any, seemed to fade rapidly. Regardless of whether abated episodes of conflict were observed, conflict was observed and, using the ethogram, we will be able to determine under which circumstances and between which OR team members it more typically occurs.

The observers' vantage point in the room, by the circulating nurse's station, also may have influenced data collection. The anesthesia team, specifically, could be

difficult to see and hear at times. The surgery and nursing team members contended with the drape obscuring the anesthesia workstation as well. We are planning future experiments to address how these drapes influence communication.

The data collected using the ethogram can be used to discover numerous facets of OR behavior, such as correlations between certain hospitals, departments, types of surgical procedures, and the quality of teamwork. The data may also reveal associations between the gender composition in the OR and teamwork, or the possible effects of the presence of multiple surgeons (or alpha individuals, in primatological language). Furthermore, we can use these data to develop sociograms that map the interconnectedness of various team members and actions that are either supportive or conflict-provoking. Such social analyses have the potential to serve as the basis for developing interventions to improve OR behavior, hospital administrators' ultimate goal when engaging social and behavioral scientists. The advantage of the ethological method is that it also can be used to determine the effect of interventions by measuring OR behavior before and after their implementation. We encourage the pairing of the ethological method with other methods, such as questionnaires, interviews, or videography (where possible), to increase external validity and achieve a more nuanced understanding of OR behavior.

Continued research on methods apropos for conducting rigorous, large-scale studies of OR team members' behavior is imperative. Such research is possible because ethological methods are feasible in the OR yielding reliable data, as we have shown here. If we want to improve teamwork and mitigate errors, we must augment our methodological repertoire to deepen our understanding of how teams function.

Conclusion

The goal of our study was to construct an ethogram and determine whether it was feasible and reliable for the study of OR behavior and whether it could be used reliably when multiple observers were involved. We unexpectedly surpassed the level of IOR for which we originally aimed. We believe the OR ethogram is replicable from site to site and observer to observer. Still, continued research on methods apropos for conducting rigorous, large-scale studies of OR team members' behavior is imperative. If we want to improve teamwork and mitigate errors, we must augment our methodological repertoire to deepen our understanding of how teams function.

Acknowledgments Thank you to Dr. William Bornstein, Chief Medical Officer and Chief Quality Officer at Emory Healthcare, as well as President/CEO John Fox, and Fred Sanfilippo, Director, Healthcare Innovation Program. This research was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR000454. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This research was also supported by grant number R03 HS23403-01 and its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Agency for Healthcare Research and Quality.

Compliance with Ethical Standards

Author's Statement of Conflict of Interest and Adherence to Ethical Standards Authors Laura K. Jones, Bonnie Mowinski Jennings, Ryan M. Goelz, Kent W. Haythorn, Joel B. Zivot, and Frans B. M. de Waal declare that they have no conflict of interest involving anyone/anything associated with this manuscript. The highest ethics were practiced and all IRB stipulations were met.

References

1. Rogers DA, Lingard L, Boehler ML, et al. Surgeons managing conflict in the operating room: Defining the educational need and identifying effective behaviors. *Am J Surg.* 2013; 205(2): 125-30.
2. Rogers DA, Lingard L. Surgeons managing conflict: A framework for understanding the challenge. *J Am Chem Soc.* 2006; 203(4): 568-74.
3. Rogers DA, Lingard L, Boehler ML, et al. Foundations for teaching surgeons to address the contributions of systems to operating room team conflict. *Am J Surg.* 2013; 206(3): 428-32.
4. King HB, Battles J, Baker DP. Teamstepps™: Team strategies and tools to enhance performance and patient safety. *Advances in Patient Safety: New Directions and Alternative Approaches.* Vol 3. Rockville, MD: AHRQ; 2008. Retrieved from <http://teamstepps.ahrq.gov>.
5. Saltman DC, O'Dea NA, Kidd MR. Conflict management: A primer for doctors in training. *Postgrad Med J.* 2006; 82(963): 9-12.
6. Donabedian A. The quality of care. How can it be assessed? *JAMA.* 1988; 260(12): 1743-8.
7. Gilbert HV, Yan J, Hoffman SJ. A WHO report: Framework for action on interprofessional education and collaborative practice. *J Allied Health.* 2010; 39((Supplement 1)): 196-7.
8. Reeves S, Tassone M, Parker K, et al. Interprofessional education: An overview of key developments in the past three decades. *Work.* 2012; 41(3): 233-45.
9. Six leading health education associations unite to form a new organization focused on interprofessional education and practice. AAMC. [news release]. February 15, 2012. <https://ipecollaborative.org/uploads/IPEC-PR-2-14-12-Updated-Version.pdf>.
10. Yule S, Flin R, Maran N, et al. Surgeons' non-technical skills in the operating room: Reliability testing of the NOTSS behavior rating system. *World J Surg.* 2008; 32(4): 548-56.
11. Randall P. *Adult Bullying: Perpetrators and Victims.* London: Routledge; 1997.
12. Booji L. Conflicts in the operating theatre. *Curr Opin Anaesthesiol.* 2007; 20(2): 152-6.
13. Gawande AA, Zinner MJ, Studdert DM, et al. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery.* 2003; 133(6): 614-21.

14. Rosenstein AH, O'Daniel M. A survey of the impact of disruptive behaviors and communication defects on patient safety. *Jt Comm Qual Pat Saf.* 2008; 34(8): 464-71.
15. Kohn LT, Corrigan J, Donaldson MS. *To Err is Human: Building a Safer Health System.* Washington, D.C.: National Academies Press; 2000.
16. Helmreich RLDJ. Culture, threat and error: Lessons from aviation. *Can J Anesth.* 2005; 51(6): R1-4.
17. Niedhammer I, David S, Degioanni S. Association between workplace bullying and depressive symptoms in the French working population. *J of Psychosom Res.* 2006; 61(2): 251-9.
18. Joint Commission on Accreditation of Healthcare Organizations, The [JCAHO] (2015). Sentinel event statistics data: Root causes by event type (2004 - 3rd Quarter 2015); 2015. Retrieved from <http://www.jointcommission.org/>
19. Pope C. Conducting ethnography in medical settings. *Med Educ.* 2005; 39(12): 1180-7.
20. Awad SS, Fagan SP, Bellows C, et al. Bridging the communication gap in the operating room with medical team training. *Am J Surg.* 2005; 190(5): 770-4.
21. Lingard L, Reznick R, Espin S, et al. Team communications in the operating room: Talk patterns, sites of tension, and implications for novices. *Acad Med.* 2002; 77(3): 232-7.
22. Lingard L, Espin S, Whyte S, et al. Communication failures in the operating room: An observational classification of recurrent types and effects. *Qual Saf Health Care.* 2004; 13(5): 330-4.
23. Mitchell L, Flin R, Yule S, et al. Thinking ahead of the surgeon. An interview study to identify scrub nurses' non-technical skills. *Int J Nurs Stud.* 2011; 48(7): 818-28.
24. Parker SH, Yule S, Flin R, et al. Surgeons' leadership in the operating room: An observational study. *Am J Surg.* 2012; 204(3): 347-54.
25. Parker S, Flin R, McKinley A, et al. Factors influencing surgeons' intraoperative leadership: Video analysis of unanticipated events in the operating room. *World J Surg.* 2014; 38(1): 4-10.
26. Hazlehurst B, McMullen CK, Gorman PN. Distributed cognition in the heart room: How situation awareness arises from coordinated communications during cardiac surgery. *J Biomed Inform.* 2007; 40(5): 539-51.
27. Guerlain S, Adams RB, Turrentine FB, et al. Assessing team performance in the operating room: Development and use of a "black-box" recorder and other tools for the intraoperative environment. *J Am Chem Soc.* 2005; 200(1): 29-37.
28. Hamasaki T, Hagihara A. A comparison of medical litigation filed against obstetrics and gynecology, internal medicine, and surgery departments. *BMC Med Ethics.* 2015; 16(1): 72-29. Adams A, Soumerai S.
29. Studdert DM, Mello MM, Gawande AA, et al. Claims, errors, and compensation payments in medical malpractice litigation. *N Engl J Med.* 2006; 354(19): 2024-33.
30. Forsythe DE. Using ethnography to investigate life scientists' information needs. *Bull of Med Libr Assoc.* 1998; 86(3): 402-9.
31. Cassell J. *The Woman in the Surgeon's Body.* Cambridge, Mass: Harvard University Press; 1998.
32. Zhou M, Kelly, JJ, Video Recording: Responsibility and Liability. In *The SAGES Manual of Quality, Outcomes and Patient Safety*; 2012: 547–552. Springer.
33. Campbell S, Sosa JA, Rabinovici R, et al. Do not roll the videotape: effects of the health insurance portability and accountability act and the law on trauma videotaping practices. *Am J Surg.* 2006; 191(2): 183-90.
34. Eibl-Eibesfeldt I. *Human Ethology.* New Brunswick, N.J.: Aldine Transaction; 2007.
35. Hawkes K. Showing off tests of an hypothesis about men's foraging goals. *Ethol Sociobiol.* 1991; 12: 29-54.
36. Kaplan H, Hill K, Cadeliña RV, et al. Food sharing among ache foragers: Tests of explanatory hypotheses. *Curr Anthropol.* 1985; 26(2): 223-46.
37. Mulder MB, Caro TM, Chrisholm JS, et al. The use of quantitative observational techniques in anthropology. *Curr Anthropol.* 1985; 26(3): 323-35.
38. Cloak FT. Is a cultural ethology possible? *Hum Ecol.* 1975; 3(3): 161-82.
39. Troisi A. Ethological research in clinical psychiatry: The study of nonverbal behavior during interviews. *Neurosci Biobehav Rev.* 1999; 23(7): 905-13.
40. Schelde T. Ethological research in psychiatry. *Ethol Sociobiol.* 1994; 15: 349-68.
41. Warnock F. An ethogram of neonatal distress behavior in response to acute pain (newborn male circumcision). *Infant Behav Dev.* 2003; 26(3): 398-420.
42. Warnock FF, Allen M. Ethological methods to develop nursing knowledge. *Res Nurs Health.* 2003; 26(1): 74-84.
43. Scherer KR, Ekman P. *Handbook of Methods in Nonverbal Behavior Research.* Cambridge: Cambridge University Press; 1982.
44. Immelmann K, Beer C. *A Dictionary of Ethology.* Cambridge, Mass: Harvard University Press; 1989.
45. Altmann J. Observational study of behavior. *Sampling Methods Behav.* 1974; 49(3): 227-67.
46. Parr LA, Cohen M, de Waal FBM. Influence of social context on the use of blended and graded facial displays in chimpanzees. *Int J of Primatol.* 2005; 26(1): 73-103.
47. de Waal FBM. The communicative repertoire of captive bonobos (Pan paniscus), compared to that of chimpanzees. *Behaviour.* 1988; 106(3-4): 183-251.
48. Coiera EW, Jayasuriya RA, Hardy J, et al. Communication loads on clinical staff in the emergency department. *Med J Aust.* 2002; 176(9): 415-8.
49. Spencer R, Logan P, Coiera E. (2002). *Communication observation method manual.* Centre for Health Informatics, University of New South Wales.
50. Patton MQ. *Qualitative Research & Evaluation Methods.* 3rd ed. Thousand Oaks, CA: Sage; 2002.
51. Manser T, Foster S, Flin R, et al. Team communication during patient handover from the operating room: More than facts and figures. *Hum Factors.* 2013; 55(1): 138-56.
52. Gardiner TM, Marshall AP, Gillespie BM. Clinical handover of the critically ill postoperative patient: An integrative review. *Aust Crit Care.* 2015.
53. Morse JM. *Qualitative nursing research : A contemporary dialogue.* Rockville, Md.: Aspen Publishers; 1989.
54. Midford PE. Ontologies for behavior. *Bioinformatics.* 2004; 20(18): 3700-1.
55. Bakeman R, Gottman JM. *Observing Interaction: An Introduction to Sequential Analysis.* 2nd ed. New York: Cambridge University Press; 1997.
56. Viera AJ, Garrett JM. Understanding interobserver agreement: The kappa statistic. *Fam Med.* 2005; 37(5): 360-3.
57. Greenberg CC, Regenbogen SE, Studdert DM, et al. Patterns of communication breakdowns resulting in injury to surgical patients. *J Am Coll Surg.* 2007; 204(4): 533-40.
58. McCarney R, Warner J, Iliffe S, et al. The Hawthorne effect: A randomised, controlled trial. *BMC Med Res Methodol.* 2007; 7: 30.
59. Adair JG. The Hawthorne effect: A reconsideration of the methodological artifact. *J Appl Psychol.* 1984; 69(2): 334-45.