BMJ Open

"Put the what, *where*? Cut *here*?!" challenges to coordinating attention in robot-assisted surgery: a microanalytic pilot study

Antara Satchidanand ,[1,2] Jeff Higginbotham,[1,2] Ann Bisantz,[1,3] Naif Aldhaam,[1] Ahmed Elsayed,[1] Iman Carr,[1] Ahmed A Hussein,[1] Khurshid Guru[1]

To cite: Satchidanand A, Higginbotham J, Bisantz A, *et al*. "Put the what, *where*? Cut *here*?!" challenges to coordinating attention in robot-assisted surgery: a microanalytic pilot study. *BMJ Open* 2021;**11**:e046132. doi:10.1136/bmjopen-2020-046132

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-046132).

Received 30 October 2020
Accepted 25 June 2021

[1]ATLAS Program, Roswell Park Comprehensive Cancer Center Department of Urology, Buffalo, New York, USA
[2]Communication Disorders and Sciences, State University of New York at Buffalo, Buffalo, New York, USA
[3]Industrial and Systems Engineering, State University of New York at Buffalo, Buffalo, New York, USA

Correspondence to
Dr Khurshid Guru;
Khurshid.guru@roswellpark.org

ABSTRACT

Introduction During robot-assisted surgery (RAS), changes to the operating room configuration pose challenges to communication by limiting team members' ability to see one another or use gesture. Referencing (the act of pointing out an object or area in order to coordinate action around it), may be susceptible to miscommunication due to these constraints.

Objectives Explore the use of microanalysis to describe and evaluate communicative efficiency in RAS through examination of referencing in surgical tasks.

Methods All communications during ten robot-assisted pelvic surgeries (radical cystectomies and prostatectomies) were fully transcribed. Forty-six referencing events were identified within these and subjected to a process of microanalysis. Microanalysis employs detailed transcription of speech and gesture along with their relative timing/sequencing to describe and analyse interactions. A descriptive taxonomy for referencing strategies was developed with categories including references reliant exclusively on speech (anatomic terms/directional language and context dependent words (CD)); references reliant exclusively on gesture or available aspects of the environment (point/show, camera focus/movement in the visual field and functional movement); and references reliant on the integrated use of speech and gesture/environmental support (integrated communication (IC)). Frequency of utilisation and number/percent 'miscommunication', were collated within each category when miscommunication was defined as any reference met with incorrect or no identification of the target.

Results IC and CD were the most frequently used strategies (45% and 26%, respectively, p≤0.01). Miscommunication was encountered in 22% of references. The use of IC resulted in the fewest miscommunications, while CD was associated with the most miscommunications (42%). Microanalysis provided insight into the causes and nature of successful referencing and miscommunication.

Conclusions In RAS, surgeons complete referencing tasks in a variety of ways. IC may provide an effective means of referencing, while other strategies may not be adequately supported by the environment.

Strengths and limitations of this study

- ► Study methodology (microanalysis) addresses aspects of communication in robot-assisted surgery (RAS) previously unstudied due to limitations in traditional methodologies.
- ► Microanalysis affords the ability to examine how the communication constraints of the RAS operating theatre affect surgeons' integrated use of speech and gesture and how this impacts communicative efficiency in referencing.
- ► Microanalysis allows for examination miscommunication in RAS beyond quantification of breakdowns revealing its root causes.
- ► Rigorous statistical analysis of this study's quantitative results was not possible due to its small sample size.
- ► The descriptive taxonomy for referencing strategies developed in this study is likely not comprehensive due to the study's limited scope.

INTRODUCTION

Robot-assisted surgery (RAS) is fast becoming an integral part of the armamentarium to surgical sciences. But traditional operating theatre layout may pose challenges to communication among the surgical team.[1] With the surgeon seated at a remote console, it is often difficult for team members to hear one another, while necessary equipment (ie, video monitors, the surgical console) disrupts sight-lines, limiting the use of visible behaviour (ie, gesture, eye gaze) in communication (figure 1).[2 3] One communication act that may be susceptible to miscommunication in this environment is *referencing*, defined as drawing a partner's attention to a target object or area, usually to coordinate action (eg, The surgeon says 'Wash here', and the assistant locates the area).[4–6] Often in referencing, speakers rely on gestures along with speech to make targets salient.[4]Though referencing is common in daily conversation,

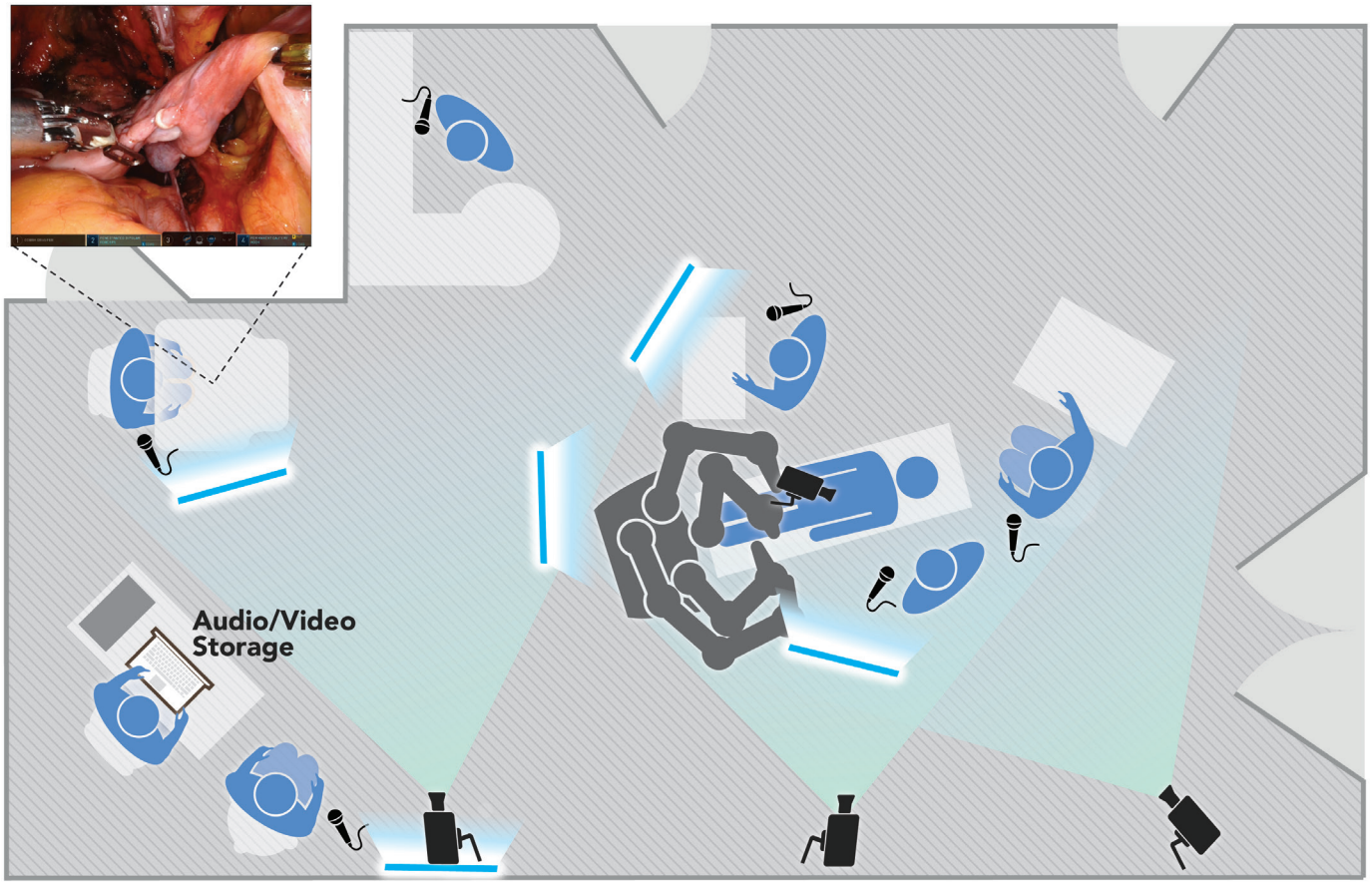


Figure 1 Operating theater layout in robot-assisted surgery.

it is also central to the completion of complex, coordinated tasks in high-stakes environments.[7] Given the long-demonstrated impact of communication efficiency on healthcare outcomes, understanding how constraints on gesture in RAS might impact referencing is vital.

Randell *et al*[1] and Raheem *et al*[8] examined communication and teamwork in RAS, by focusing on verbal requests by surgeons. Both studies demonstrated that precisely worded requests specifying the addressee or the requested item were generally more successful than less precise requests, but these findings varied with task complexity and familiarity among team members.[4] Use of gesture in RAS has also been studied with up to 87% of interactions classified as 'non-verbal'.[9 10] Referencing in RAS has not been examined thoroughly, though Elprama and colleagues identified it as a source of miscommunication in current RAS practice.[11]

While the aforementioned studies provide a foundation for understanding communication in RAS, they are all limited by addressing communication dichotomously, as consisting of either 'verbal' or 'non-verbal' acts. None of them takes into account the way integrated speech and gesture are commonly deployed.[12] Studies of communication, and referencing specifically, in laparoscopy and open surgery have used microanalysis to capture integrated communication's (IC's) role.[13] Microanalysis employs detailed transcription to examine various communication modalities simultaneously (ie, speech, gesture, timing and sequencing of communication acts, use of environmental props to aid communication).[14] Microanalysis of referencing in the context of laparoscopy has shown that negotiating the accurate identification of targets is central to helping team members coordinate their actions, and that the integrated use of speech and gesture is key to this process.[15]

This pilot study examines the feasibility of using microanalysis in conjunction with time sampling and discourse categories (ie, referencing) to locate and understand communication weaknesses in the virtual RAS environment. It provides a descriptive taxonomy for referencing in RAS and outlines a viable methodology for future larger-scale studies that describe and analyse the root causes of miscommunication with an eye to the development of both technological and training solutions.

METHODS

Data for this study were recorded in the operating room at Roswell Park Comprehensive Cancer Center as part of the Technofields project, launched with the aim of studying the operative environment in RAS. Technofields was funded by the Roswell Park Alliance Foundation (no grant number available) and the University at Buffalo's Office of Research and Development's Innovative Micro-Programs Accelerating Collaboration in Themes grant. Seventy-nine surgeries performed between 2013 and 2015

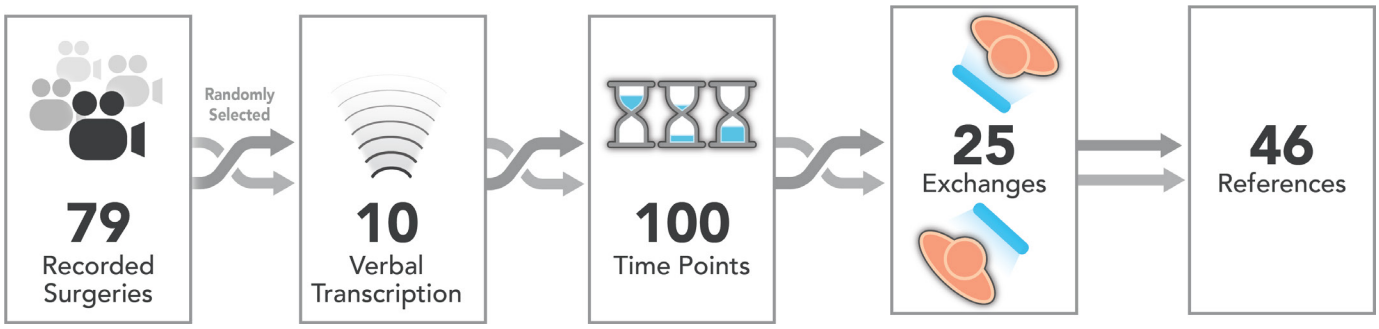


Figure 2 Sample selection process.

were video recorded, and 10 were selected at random for full verbal transcription. Within these, 100 time points were randomly selected, and topically bounded communication exchanges containing these points identified. For the current study, 25 exchanges were selected at random and investigated for referencing, yielding 46 referencing events (figure 2). These events were then subjected to microanalysis, and a descriptive taxonomy for referencing strategies used in RAS was developed.

All surgeries were performed by a mentor surgeon (15 years or more of experience) or surgical fellow engaged in training using the da Vinci Si Surgical System. The console surgeon's view of the surgical field was shared throughout the operating theater via video monitors. Microphones and speakers within the surgical console and on the monitor closest to the bedside assistant relayed audio between the surgeon and bedside team.

Our methodology for recording the OR environment has been previously described.[9] Three overhead cameras provided a view of surgical team members' relative positions in the OR, and a fourth feed recorded console footage to provide surgical context for the communication. Audio was recorded using lapel microphones worn by the surgeon, assistant surgeon (trainee), bedside assistant, scrub nurse and up to four additional trainees/shift replacement personnel. Output from the surgical system's microphones and speakers was not directly available to the research team but was picked up by participants' lapel microphones. Four trained graduate students, one PhD candidate and a full professor of communicative disorders and sciences (CDS) performed all transcription using the Eudico Linguistic Annotator.[16] Transcription time per minute of video was approximately 18:1.

For the current study, definitions of key terms and taxonomic descriptions were drawn from the literature in communication science. 'Referencing events' were defined as interactions in which one participant sought to draw another's attention to a target area or object.[4–6] Only references made to targets within the surgical field were used in this study. A reference was classified as 'successful' when it contained three observable parts: verbal production of a reference, identification of the target through words or action, and confirmation of accuracy.[3 4 17] For example, the surgeon requests that an area be 'washed', the assistant washes what he believes to be the target, and the surgeon confirms that he is correct by saying 'Yeah', or by simply continuing with the next expected action. References were considered 'miscommunication' when the recipient asked for clarification, identification of the target was met with an observable rejection or the recipient failed to respond within one second.[3 17 18] The final criterion is based on well-established evidence that speakers across cultures interpret 'gaps' of longer than 1 s between a contribution and its response to be indication that the recipient had difficulty with contribution's uptake.[18] For references in which the initial attempt was classified as miscommunication, each subsequent attempt was counted and analysed as a separate reference (figure 3).

Microanalysis was completed by the PhD candidate on the transcription team through two additional transcription passes and a process of case verification including at least one surgical fellow or surgeon. The first additional transcription pass captured surgical tool movement within the body cavity, and the second, timing/sequencing of speech and tool movement.

Based on the results of microanalysis, researchers developed six taxonomic categories to describe referencing strategies observed in RAS.[12 17 19] The categories anatomical terms (AT) and context dependent words (CD) included speech dependent references; (point/show (PT), functional movement (FM) and camera focus/movement (CF) included references using gesture or other action involving elements of the environment (ie, the camera's capacity to zoom in or out); and IC included references made through the coordinated use of speech and gesture/other action (table 1).

Taxonomic categories were also grouped dichotomously into Speech Only (SO) and Gesture Inclusive (GI) references, reflective of the 'verbal/non-verbal' classifications used in prior studies of communication in RAS. AT and CD comprised the SO group, and the remaining categories (PT, FM, CF, IC) formed the GI group.[10]

Case verification was used for transcription and microanalysis, with discrepancies resolved through work sessions attended by between two and four members of the research team. Inter-rater agreement was calculated for microanalysis and taxonomic classification using 10 randomly selected referencing events representing just under 22% of the sample. Review was provided by the

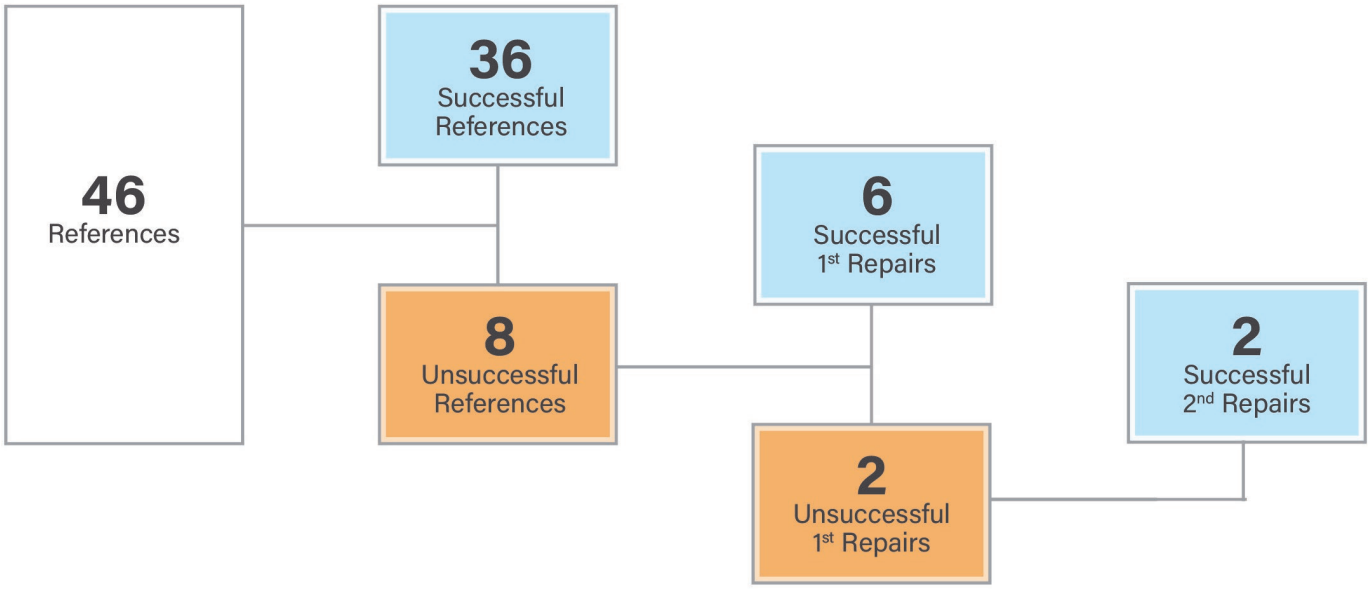


Figure 3 References: initial attempts and repairs.

team's CDS faculty member. Agreement for each was 90%.

Statistical analysis

Frequency counts were used to describe the frequency of use and effectiveness of referencing strategies. Comparisons across referencing strategy categories were performed using non-parametric testing (ie, binomial test, Fischer's exact test). A two-tailed alpha level was set at $p<0.05$. All statistical analyses were performed using SAS (V.9.4, SAS Institute).

Patient

Patients and the public were not involved in the design of the research questions, methods, outcome measures or dissemination plans for this study. Dissemination to these groups is not applicable.

RESULTS

Analysis of referencing strategies, revealed that IC was the most frequently used (45%, n=21) and was the most effective strategy with 90% (n=19) of references successfully completed (binomial test, $p<0.01$) (figure 4). The next most commonly used strategy, CD (26%, n=12) was not shown to be particularly effective (58%; n=7 successful), however, a Fischer's exact test of effectiveness was not statistically significant ($p=0.77$).

Miscommunication occurred in 22% (n=10) of attempted references. A total of eight initial references resulted in miscommunication. Six of these were resolved with one repair and the remaining two required an additional repair. None remained unresolved.

To better understand the effects of a dichotomous analysis of referencing, reflective of the analysis commonly

Table 1 Referencing strategies in robot-assisted surgery

Strategy	Classification	Description
Anatomy terms/verbal direction (AT)	Speech Only	► Speaker uses directional words without an accompanying gesture (ie, 'go left') to direct attention. ► Speaker uses anatomic terms to direct attention (ie, 'toward the umbilicus').
Context dependent words (ie, here/there) (CD)	Speech Only	► Speaker says 'Stop there' without gesturing to specify the point at which the listener should come to rest in the surgical field.
Point/show (PT)	Gesture Inclusive	► Speaker uses instrument to point at referent. ► Speaker uses instrument to grasp referent. ► Speaker uses instrument to clear view of referent.
Functional movement (FM)	Gesture Inclusive	► Speaker retracts, cauterises or blunt dissects tissue and uses that touch to identify a referent.
Camera focus/movement (CF)	Gesture Inclusive	► Speaker zooms the camera in on referent. ► Movement visible on the monitor is treated as salient by the speaker (ie, dripping blood).
Integrated communication (IC)	Gesture Inclusive	► Use of AT or CD in conjunction with PT, CF or FM.

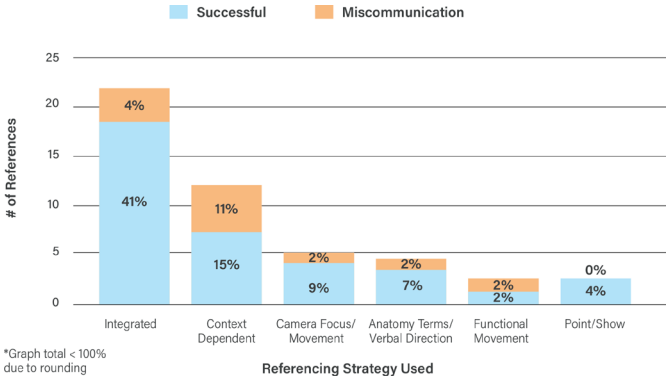


Figure 4 Frequency and effectiveness: referencing strategies.

used in prior research addressing communication in RAS, the SO and GI groups were compared (figure 5).

The SO group of strategies comprised 35% (n=16) of references and was 62% effective, while GI accounted for 65% (n=30) of references and was 87% effective, though a Fischer's exact test of the difference in effectiveness was not statistically significant (p=0.07).

Microanalysis of miscommunication

Microanalysis provided insight into the nature of both successful and unsuccessful referencing in RAS. The following example, 'Cut *Where?*' highlights the use of the CD strategy resulting in miscommunication. Taken together, the transcript, written analysis and video that follow provide a window into synthesis of the elements of interaction captured in microanalytic transcription and case verification (figure 6 and online supplemental file 1). The use of various formats to present microanalytic data often becomes part of an iterative process, with translation into each new format revealing new elements of interaction. Below, the trainee surgeon (Trainee) and his mentor (Mentor) attempt to mitigate the mentor's lack of access to gesture which impedes his use of IC. A misplaced cut results.

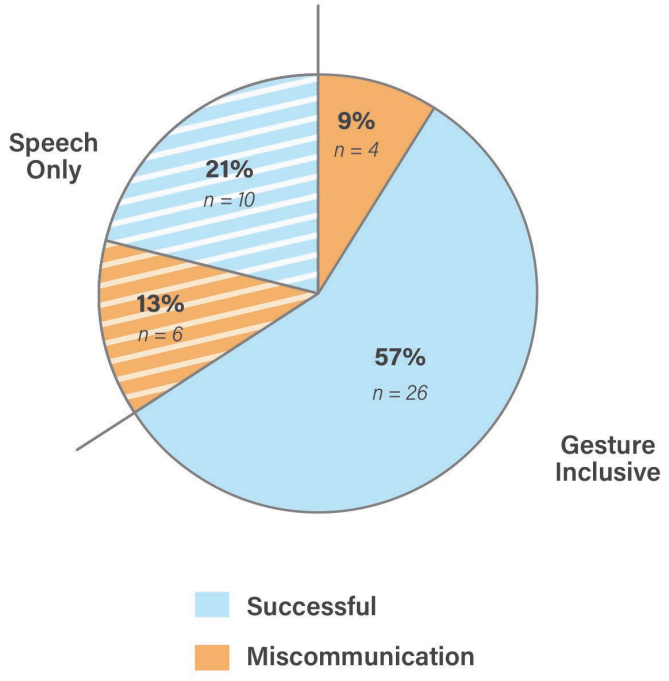


Figure 5 Frequency and effectiveness: Gesture Inclusive Speech Only.

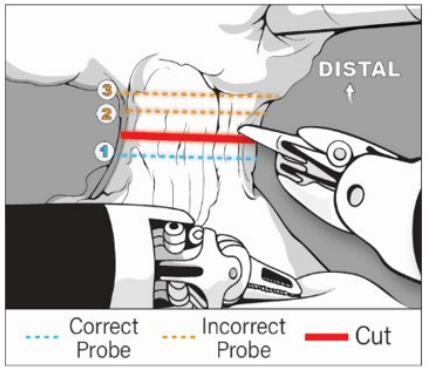
To begin, Trainee is seated at the console while Mentor watches from the observation area. Mentor directs Trainee to 'clean … back' tissue from the urethra prior to anastomosis (0 s). In deciding where to cut, Trainee probes the area 'on the right-hand side' of the urethra three times in rapid succession. Each probe lands distal to the last. Because Mentor cannot point or otherwise gesture within the surgical field, he must rely on speech to draw Trainee's attention to the correct location. To do this, Mentor attempts to coordinate his words with Trainee's surgical tool movements, using those movements as proxy point gestures. Mentor offers verbal approval of the location of probe 1 with, 'Yeah', then repeats this approval as Trainee begins probe 2. Once probe 2 is complete, however, Mentor assesses its placement, distal to probe 1, as problematic. Mentor's approving utterance and Trainee's movements have become discoordinated. Mentor initiates a correction of probe 2's placement (6.4 s), but before he can finish his corrective utterance, Trainee executes probe 3 (6.5 s) even further distal. While it was initially unclear whether Mentor's approval indicated that probe 1 was in the correct general area for dissection (ie, right of the urethra) or in the precise location to be cut, his correction of probe 2 clarifies that only a cut at the position of probe 1 or proximal to it is acceptable. In the time it takes for Trainee to adjust his retraction in preparation to cut, (6.5 s) he hears and accepts Mentor's correction of probe 2. He repositions his scissors proximal to probes 2 and 3 as directed, but still unaware that his placement of probe 1 was significant, he chooses a point distal to it and cuts. He is met with rejection from Mentor and withdraws his scissors.

This example illustrates how unequal access to gesture within the body coupled with errors in the sequencing and interpretation of talk and action lead to miscommunication and difficulty in communication repair.

DISCUSSION

In 2015, The Joint Commission identified miscommunication as the third leading root cause of sentinel incidents in US healthcare.[20] The Department of Defense and Agency for Healthcare Research and Quality also point to communication as one of four core areas of medical team training in need of improvement.[21] With the growing use of RAS, understanding of how communication occurs in this environment is vital for patient safety.

The comparison of GI and SO references in RAS showed that 65% of references employed some type of gestural communication. This is consistent with prior studies showing that an average of 67% of interactions were found to be 'non-verbal'.[10] However, analysis of



Here, the Mentor Surgeon (MS) attempts to coordinate his talk with the Trainee Surgeon (TS)'s actions to specify where a cut should be made. Temporal misalignment between TS's actions and MS's speech causes ambiguity as to which probe MS has granted approval of. An erroneously placed cut by TS results at 9.6 seconds.

Time (sec)	Person	Utterance/Action
0	MS	So on the right hand side [you want to clean that back a little bit.
1.5	TS	[(Moves scissors right of urethra)
4.6	TS	(Probes @ 1)
4.9	MS	Yeah.
5.6	TS	(Probes @ 2)
5.8	TS	(Continues to move scissors distal)
6.4	MS	You're [getting a little [too far out distally on the right.
6.5	TS	[(Probes @ 3) [Adjusts retraction right.
8.2	TS	(Approaches tissue slowly with scissors)
9.1	MS	On the left [I like [where you are.
9.6	TS	[Places scissors. [Cuts
10.9	MS	No you don't want to do that!
11.3	TS	(Withdraws scissors)

Figure 6 Microanalysis of miscommunication in referencing.

referencing by taxonomic category revealed that 70% (21/30) of GI references fell under IC, meaning that they were not reliant exclusively on gesture, but on the use of speech and gesture in combination. This finding is consistent with a body of literature indicating that the combined use of speech and gesture is central to both professional and non-professional interactions.[19] In prior studies of communication in RAS, the treatment of communication as either strictly 'verbal' or 'non-verbal' obscured this relationship leaving the contribution of IC unaccounted for. Our data revealed IC to be the most frequently used and effective referencing strategy.

The use of microanalysis in this pilot study revealed causal factors of miscommunication, including temporal misalignment, and suggested various avenues for further research, technology development and training. For example, the TeamSTEPPS programme for improved healthcare team interaction recommends the use of closed-loop communication but provides little guidance regarding the contexts in which it can be used or what its use might look like.[22] Microanalysis can provide just such teachable moments. In 'Cut *Where*?' the trainee surgeon moved his scissors slowly into place before making his erroneous cut, leaving time both to request and be given instruction. Mentor surgeons and trainees can be made aware of opportunities like these and of the value of pausing in order to maintain temporal coordination. These are skills not currently addressed in surgeons' non-technical education.

This study's limitations, including its small sample size, are in part related to its exploratory nature and the use of microanalytic techniques. The descriptive and inferential analyses used were meant to augment the findings provided by the microanalytic investigation, therefore issues of statistical power and generalisation could not be validly applied in this feasibility study. Future quantitative research may benefit from additional statistical analyses. Because this study used a pre-recorded video database, it was impossible to evaluate the impact of referencing training on surgical performance. Future research could include refinements to the taxonomy, microanalysis of specific surgical tasks, consideration of interaction differences between learning exchanges and general surgical procedure, and the impact of training on effective referencing.

CONCLUSION

Integrated speech and gesture play a key role in completing referencing tasks during RAS, and microanalytic study of this multimodal communication provides an important methodology for understanding both successful interaction and problematic communication in this environment.

Contributors Concept: AS, JH, AB, KG. Data curation: AS, NA, AE. Formal analysis: AS and JH, AE. Funding acquisition: JH, AB, KG. Methodology: JH, AB, AS. Validation: JH, AS, NA, AE, KG. Visualisation: IC, AS. Writing-original draft: AS. Writing-review and editing: AAH, KG, AE, JH.

Funding Funding for this study was provided by the State University of New York at Buffalo Office of Research and Economic Development's Innovative Micro-Programs Accelerative Collaboration in Themes (IMPACT) grant (grant number 1137639-1-000077) with support from The Roswell Park Alliance Foundation, Roswell Park Comprehensive Cancer Center. Funders played no role in study design, data collection, analysis or interpretation. Researchers were independent of funders and all authors had unfettered access to all data in the study and can take responsibility for the integrity and accuracy of the data and analysis. The lead author affirms that this manuscript is an honest, accurate and transparent account of the study being

reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval This study received ethics approval from Roswell Park Comprehensive Cancer Center's review board (I 244113). All patients and operating room staff gave informed consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. Not applicable.

ORCID iD

Antara Satchidanand http://orcid.org/0000-0002-3300-779X

REFERENCES

1 Randell R, Honey S, Hindmarsh J, *et al*. A realist process evaluation of robot-assisted surgery: integration into routine practice and impacts on communication, collaboration and decision-making. *Health Services and Delivery Research* 2017;5:1–140.
2 Cavuoto LA, Hussein AA, Vasan V, *et al*. Improving teamwork: evaluating workload of surgical team during robot-assisted surgery. *Urology* 2017;107:120–5.
3 Clark H, Brennan SE. Grounding in communication. In: Resnick JML LG, Teasley SD, eds. *Perspectives on socially shared cognition*. Washington, DC, US: American Psychological Association, 1991: 127–49.
4 Clark H, Bangerter A. Changing ideas about reference. In: Noveck LA, Sperber D, eds. *Experimental pragmatics*. London: Palgrave Macmillan UK, 2004: 25–49.
5 Levinson SC. *Pragmatics*. Cambridge, England: Cambridge University Press, 1983: 434.
6 Clark HH, Wilkes-Gibbs D. Referring as a collaborative process. *Cognition* 1986;22:1–39.
7 Hindmarsh J, Heath C. Embodied reference: a study of deixis in workplace interaction. *J Pragmat* 2000;32:1855–78.
8 Raheem S, Ahmed YE, Hussein AA, *et al*. Variability and interpretation of communication taxonomy during robot-assisted surgery: do we all speak the same language? *BJU Int* 2018;122:99–105.
9 Tiferes J, Bisantz AM, Bolton ML, *et al*. Multimodal team interactions in robot-assisted surgery. *Proc Hum Factors Ergon Soc Annu Meet* 2016;60:518–22.
10 Tiferes J, Hussein AA, Bisantz A, *et al*. Are gestures worth a thousand words? verbal and nonverbal communication during robot-assisted surgery. *Appl Ergon* 2019;78:251–62.
11 Elprama S, Kilpi K, Duysburgh P, *et al*. Identifying barriers in telesurgery by studying current team practices in robot-assisted surgery. *Proceedings of the ICTs for improving Patients Rehabilitation Research Techniques* 2013.
12 Goodwin C. Pointing as situated practice. In: Kita S, ed. *Pointing: where language, culture and cognition meet*. New York: Psychology Press, 2003: 225–50.
13 Koschmann T, Goodwin C, LeBaron C, *et al*. Dissecting common ground: examining an instance of reference repair. *Proceedings of the Annual Meeting of the Cognitive Science Society* 2001.
14 Higginbotham DJ, Engelke CR. A primer for doing talk-in-interaction research in Augmentative and alternative communication. *Augment Altern Commun* 2013;29:3–19.
15 Koschmann T, LeBaron C, Goodwin C. The mystery of the missing referent: objects, procedures, and the problem of the instruction follower. *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work; 2006: ACM* 2006:373–82.
16 Elan (Version 5.1) [Computer software]. Nijmegen: max Planck Institute for Psycholinguistics. The Language Archive.. Available: https://tla.mpi.nl/tools/tla-tools/elan
17 Clark H. *Using language*. Cambridge: Cambridge University Press, 1996.
18 Roberts F, Francis AL, Morgan M. The interaction of inter-turn silence with prosodic cues in listener perceptions of "trouble" in conversation. *Speech Commun* 2006;48:1079–93.
19 Goodwin C. *Co-Operative action*. Cambridge, England: Cambridge University Press, 2018: 521.
20 Schumacher P. The Joint Commission online a complementary publication of The Joint Commission. The Department of Communications.2015 https://www.jointcommission.org/-/media/deprecated-unorganized/imported-assets/tjc/system-folders/joint-commission-online/jconline_april_29_15pdf.pdf?db=web&hash=DEFFBC41623A360F1C1428A5E9602773
21 Clancy CM, Tornberg DN. TeamSTEPPS: assuring optimal teamwork in clinical settings. *Am J Med Qual* 2007;22:214–7.
22 Weld LR, Stringer MT, Ebertowski JS, *et al*. TeamSTEPPS improves operating room efficiency and patient safety. *American Journal of Medical Quality* 2016;31:408–14.