PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Put the What, Where? Cut Here? Challenges to Coordinating Attention in Robot-Assisted Surgery: A Microanalytic Pilot Study
AUTHORS	Satchidanand, Antara; Higginbotham, Jeff; Bisantz, Ann; Aldhaam, Naif; Elsayed, Ahmed; Carr, Iman; Hussein, Ahmed; Guru, Khurshid

VERSION 1 – REVIEW

REVIEWER	Thanigasalam, Ruban Chris O'Brien Lifehouse
REVIEW RETURNED	07-Jan-2021

GENERAL COMMENTS	Authors have used a novel way to assess co-ordinating attention in robotic surgery. It would have been preferable to have two cohorts to assess the micro-analysis of the reference events, such as a standard group (where the surgeons/fellows perform the surgery as normal) and a reference group (where the surgeons/fellows perform the surgery following a debrief on effective communication and the descriptive taxonomy-PT/CF/AT/FM/CD/IC described in the article for the robotic procedure), and then assess the outcomes in terms which provided the most effective or least effective strategy in achieving the task. The microanalysis of miscommunication paragraph in the results section should be reviewed and revised. The binomial exact test of the difference in effectiveness is not statistically significant (p=0.07). The discussion reflects that this is a pilot study and the sample size is small, however integrated speech and gesture are important in completing keys steps of the robotic procedure. The clinical utility of the study to practice would be more useful with a larger sample size, however the authors should be commended on their novel work in this area.

REVIEWER	Tsafrir, Ziv Tel Aviv Sourasky Medical Center
REVIEW RETURNED	16-Jan-2021

GENERAL COMMENTS	This inovative study shed a light on an important challenge introduced in robotic assisted surgeries: the ability of team members to communicate effectively. The authors analyzed the efficiency of diffferent communication strategies and concluded that integrated use of speech and

	gesture may be the most reliable.
	The weakness of this study stem from it's small sample size.
	My comments are as follows:
	1. Was power analysis performed? On what ground did the authors decide to randomly select 10 surgeries only?
	2. Did all team-members wear microphones?
	3. How many team members participated in these cases? Who was the leading surgeon at the console (Mentor / trainee)? What was the experience of the mentor and trainee in robotic surgery?
	4. Which model type of the robotic platform was used (The Si type has microphone at the console site only, compared to the Xi)? This aspect may influence the efficiency of speech dependent references.
	5. Was the Mic turned on at the console site?
	6. In a case which the initiator did not vocally confirm the recipient's act, and the procedure went on, would the authors categorize this reference as "successful'?
	7. Defining "miscommunication" when the recipient failed to respond within 1 second may be too crude, especially if the trainee surgeon is inexperienced.
	8. If one or the participants couldn't hear clearly the vocal reference, was it classified as "miscommunication"?
	9. Fig. 4 should be amended. Columns are not summurized to 100%.
	10. Page 10. Paragraph 5, line 49 - This paragraph is redundant, (Fig. 6 and supplemental video already discuss it).

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Prof. Ruban Thanigasalam, Chris O'Brien Lifehouse, The University of Sydney

- It would have been preferable to have two cohorts to assess the micro-analysis of the reference events, such as a standard group (where the surgeons/fellows perform the surgery as normal) and a reference group (where the surgeons/fellows perform the surgery following a debrief on effective communication and the descriptive taxonomy-PT/CF/AT/FM/CD/IC described in the article for the robotic procedure), and then assess the outcomes in terms which provided the most effective or least effective strategy in achieving the task.;

    Thank you for your comment. We were unable to create cohorts based upon training on effective referencing because our data were recorded a year prior to our work as part of a larger study. This is clarified in revisions to the organization of the methods section on page 6 (ln 19-20) and page 7 (ln 2) and addressed in the revised discussion of our study's limitations on page 15 (ln 8-9).

- The microanalysis of miscommunication paragraph in the results section should be reviewed and revised.

    We have clarified our approach to the written microanalysis on page 12 (ln 5– 12) by providing an explanation of its role in laying the foundation for others who may want to use our methods.

2

- The binomial exact test of the difference in effectiveness is not statistically significant (p=0.07).

    > On pages 11 (lns 11 and 20) "binomial test" has been corrected to "Fischer's Exact Test". In addition, on page 12 (ln 1-2) we have revised our report of the results to read "the difference in effectiveness was not statistically significant"

- The discussion reflects that this is a pilot study and the sample size is small, however integrated speech and gesture are important in completing keys steps of the robotic procedure. The clinical utility of the study to practice would be more useful with a larger sample size, however the authors should be commended on their novel work in this area.

    > We have added mention of the way our methods impact sample size and have provided more detailed discussion of possible future directions in the paragraph addressing this study's limitations on page 15 (ln 3-4 and 9-12).

    > Thank you for your commendation of our work.

Reviewer: 2
Dr. Ziv Tsafrir, Tel Aviv Sourasky Medical Center

- Was power analysis performed?

    > Thank you for this clarifying question. Due to the study's reliance on microanalytic techniques, issues of statistical power and generalization could not be validly applied. We have added discussion of this and the generalizability of our work to the paragraph on study limitations page 15 (ln 4-7) of the marked main document.

- On what ground did the authors decide to randomly select 10 surgeries only?

    The decision to select 10 surgeries only was made prior to the start of our study. This is clarified in the revisions to our methods section on page 7 (ln 1-3).

    The number of surgeries was limited to 10 due to the time and labor involved with full transcription. Surgeries were an average of 1 hour and 55 minutes in length with full verbal transcription taking approximately 18 minutes per one minute of video. A total of approximately 20,000 utterances were transcribed as part of the larger study from which our video data were drawn.

- Did all team-members wear microphones?

    Four key members of the surgical team, the surgeon, assistant surgeon (trainee), bedside assistant, scrub nurse wore lapel microphones during all cases. Up to four additional trainees / shift replacement personnel wore them as needed. Please see revisions to our method section on to page 7 (ln 18-19).

- How many team members participated in these cases?

    Only mentor surgeons, trainee surgeons, bedside assistants and scrub nurses were represented in our referencing events. Each of these roles were filled by various individuals over the course of the study. No formal data was recorded on how many different individuals served in each role.

- Who was the leading surgeon at the console (Mentor / trainee)?

The console was manned by both mentor surgeons and trainee surgeons during the surgeries in our corpus. We have added this detail in revisions to our methods on page 7 (ln 10-11), however we do not have data on how many of the referencing events in our sample took place with either mentors or trainees at the console.

- What was the experience of the mentor and trainee in robotic surgery?

  Mentor surgeons had more than 15 years of experience. Trainee surgeons were surgical fellows with varying experience. This information has been added on page 7 (ln 10 - 11)

- Which model type of the robotic platform was used (The Si type has microphone at the console site only, compared to the Xi)?

  The da Vinci Si model was used for these surgeries, and this has been to our methods on page 7 (ln 11) along with description of how surgical team members were able to hear one another in the OR.

- Was the Mic turned on at the console site?

  All of the system's microphones were on allowing team members to hear one another during surgery as described on page 7 (ln11) of the revised methods description.

- In a case which the initiator did not vocally confirm the recipient's act, and the procedure went on, would the authors categorize this reference as "successful'?

  This important detail has been clarified in revisions to page 8 (ln 15-16) of the methods section which now includes an example of confirmation of a successful reference through continued action.

- Defining "miscommunication" when the recipient failed to respond within 1 second may be too crude, especially if the trainee surgeon is inexperienced.

  Thank you for this clarifying comment.

  There is a well-established body of literature in communication science demonstrating that when people are engaged in conversation, a lag of longer than 1 second in duration between the end of a speaker's utterance and the response from his/her listener, whether in language or other action, is typically interpreted by the speaker as indication that there was some difficulty with uptake of that contribution (Jefferson, 1989; Roberts, Francis, & Morgan, 2006; Sacks et al., 1974). This includes references. Based on the high stakes and temporally demanding nature of surgery, and the need to provide feedback to prevent surgical error, we deemed a one second gap to be an appropriate marker of difficulty. We have provided a brief grounding in this theory to address concerns other readers may have on page 8 (ln 21-23) of the revised methods.

- If one or the participants couldn't hear clearly the vocal reference, was it classified as "miscommunication"?

  Yes, a non-response or request for clarification due to difficulty perceiving a reference would be classified as a miscommunication under our criteria. We have made this more clear with the addition of "asking for clarification" as an indication that miscommunication has occurred on page 8 (ln 19-20).

- Fig. 4 should be amended. Columns are not summurized to 100%.

  A note has been added below the graph in Figure 4 indicating that the total of <100% is due to rounding.

4

- Page 10. Paragraph 5, line 49 - This paragraph is redundant, (Fig. 6 and supplemental video already discuss it).

> The primary purpose of this study was to explore the feasibility of using microanalysis to study communicative interaction in context. Our written analysis provides a window into the components of microanalysis that may support others in learning to apply our techniques. The written analysis allows the reader to examine an example of microanalytic reasoning without relying on the supplemental video. On page 12 (ln 5 – 10) we have provided an explanation of the role written analysis, transcription and video can play in the iterative process of microanalysis and the presentation of microanalytic data to audiences.

References

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L.G. Resnick, J.M. Levine, & S.D. Teasley (Ed.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC, US: American Psychological Association.

Jefferson, G. (1989). Preliminary notes on a possible metric which provides for a "standard maximum" silence of approximately 1 second in conversation. In D. Roger & P. Bull (Eds.), *Conversation: An Interdisciplinary Perspective* (p. 166). Multilingual Matters.

Hindmarsh, J., & Heath, C. (2000). Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics, 32*(12), 1855–1878.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language, 50*(4), 696–735.

Roberts, F., Francis, A. L., & Morgan, M. (2006). The interaction of inter-turn silence with prosodic cues in listener perceptions of "trouble" in conversation. *Speech Communication, 48*(9), 1079–1093.

VERSION 2 – REVIEW

REVIEWER	Tsafrir, Ziv Tel Aviv Sourasky Medical Center
REVIEW RETURNED	30-May-2021

GENERAL COMMENTS	The authors fully and apropriatelly addressed reviewers concerns. This novel study shed a light on an important communication challenge of robotic assisted surgery.