
Cultivating Psychological Safety in Human-Robot Teams with Social Robots

Sarah Strohkorb

Yale University
New Haven, CT 06511, USA
sarah.strohkorb@yale.edu

Brian Scassellati

Yale University
New Haven, CT 06511, USA
brian.scassellati@yale.edu

Abstract

Research in human teams has exposed the importance of emotional intelligence and psychological safety in group functioning and overall task performance. As social robots are increasingly incorporated in human teams, we seek to provide them with the ability to positively contribute to the team's social environment. We draw from a diverse body of work on the topics of vulnerable disclosure, illness support groups, and improvisational theater as inspiration as we look to design robot behaviors that enhance group psychological safety.

Author Keywords

Human-Robot Interaction, Psychological Safety, Human-Robot Group Collaboration

ACM Classification Keywords

I.2.9 [Computing Methodologies]: Robotics; H.1.2 [Information Systems]: User/Machine Systems

Introduction

Organizations of all sizes are interested in group performance. Work in academia and industry increasingly relies on collaboration and teamwork between both physically present and virtual team members. Unsurprisingly, many organizations place a premium on finding the top talent to fill their teams. In response to the growing importance

Paste the appropriate copyright statement here. ACM now supports three different copyright statements:

- ACM copyright: ACM holds the copyright on the work. This is the historical approach.
- License: The author(s) retain copyright, but ACM receives an exclusive publication license.
- Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single spaced in a sans-serif 7 point font.

Every submission will be assigned their own unique DOI string to be included here.

of group performance, Woolley et al. set out to determine what, if any, factors determine group success, much like cognitive ability (measured by intelligence tests like the IQ test) reliably predicts individual success on mental tasks [5]. Woolley et al. demonstrated evidence for a collective intelligence, that is "not strongly correlated with the average or maximum individual intelligence of group members" but rather individual social sensitivity, the distribution of turn taking among group members, and the number of females in the group [18]. These findings strongly suggest that although emotional/interpersonal intelligence may not determine individual success, it is essential for group success. As Howard Gardner proposed in his theory of multiple intelligences, human intelligence is multi-dimensional, including an interpersonal dimension which is characterized by individuals who communicate effectively and easily empathize with others [7].

Work by Edmondson aids our understanding of the connection between emotional intelligence and team performance. Edmondson defined a concept called psychological safety: "a shared belief held by members of a team that the team is safe for interpersonal risk taking." After thoroughly examining the dynamics of 51 diverse teams from a large manufacturing company, Edmondson discovered that psychological safety predicts team performance when mediated by team learning behavior (e.g. seeking feedback, discussing errors, and learning from mistakes) [6]. In other words, psychologically safe teams who express mutual respect, trust, and care, are more likely to express behaviors that produce higher performance.

If psychological safety is critical for the success of human-robot teams, it is essential to discover mechanisms that team members, as fellow peers, can employ to build an environment where all team members feel safe to take in-

terpersonal risks. It is well established in psychological and management literature that the behavior of team leaders can significantly influence the performance of their teams [15]. Even though team member relationships have received less attention, the behavior of a single team member on team behavior and performance has demonstrated a similar effect. Barsade has shown that the pleasantness of one member of a group has significant effects on group mood [1]. He has named this spreading effect an 'emotional contagion' and has shown not only an impact in group mood, but also increased cooperation, less conflict, and greater perceived task performance in a group.

Thus, robots in human-robot teams can reasonably influence group dynamics using their own emotional contagions. In order to define more specifically which behaviors robots can use to grow group psychological safety, we examine research and take inspiration from a variety of topics and sources including vulnerable disclosure, illness support groups, and improvisational theater.

Vulnerable Disclosure

Psychological study on self-disclosure has revealed a reciprocity effect: individuals are more likely to self-disclose after an interaction partner has revealed intimate information [3]. Building psychological safety within a group is characterized by team members feeling safe to take interpersonal risks, and it is likely that this interpersonal safety can be built using vulnerable disclosure. Since vulnerable disclosure has been shown to be reciprocal for humans, is it also possible that a vulnerable disclosure by a robot could elicit reciprocal vulnerability in a human?

Martelaro et al. tested the responses of participants interacting with robots having either high/low vulnerability and either high/low expressivity, and discovered that participants

trusted the robot that was more vulnerable and felt that they disclosed more to the expressive robot [13]. One example of a high vulnerable and high expressive disclosure from the robot in this study was, "They reset my memory this morning, so my day has been a little rough" (where the robot was frowning, moved its arms downward, and exuded a blue color). Although robot disclosures of vulnerability will certainly be of a different nature than typical human vulnerable disclosures, this study suggests that a social robot in a group could engender psychological safety through the use of vulnerable disclosures.

Learning from Support Groups

Illness support groups have shown a promising effect on mental and physical health of participants, with groups ranging from the more stigmatized diseases (e.g. AIDS, alcoholism, breast and prostate cancer) to less embarrassing but similarly detrimental illnesses (e.g. heart disease, stroke, lung cancer) [4]. Ussher et al. have found that cancer support groups, as compared with other supportive relationships, provide participants with an accepting community where members feel an increased sense of empowerment and facilitate positive relationships with friends and family [17]. One notable feature of these support groups is that they provide a safe place for participants to express the emotions they experience as they learn to cope with their diagnosis. As cancer patients in cancer support groups express their negative affect and adopt an attitude of realistic optimism, they experience an overall reduction of distress [2].

Although the teams that robots interact with will likely not resemble cancer support groups, emotional expression is relevant for psychological safety in groups. Members in psychologically safe groups believe that team members will care about them as a person, respect their competence,

and not reject or judge them [6]. Thus, team members in psychologically safe groups feel comfortable expressing anger, frustration, or tension they experience through group work. In human-robot teams, social robots could facilitate and influence the expression of emotion. Robotic systems are growing in their ability to detect affect expression, and could likely invite expression of observed affect among group members.

Examining Improvisational Theater

Improvisational theater, or improv, is a form of theater where a script is created and acted out at the time of performance. Actors simultaneously create, act within, and embellish story lines by advancing the contributions of the other actors into often bizarre directions. The significance of improv relative to our focus on psychologically safe teams is found in the supportive, risk-taking, and communicative characteristics of improv groups. Both professors in academic settings that want to promote good teamwork within student project teams and companies that desire effective communication and idea generation within teams and departments have implemented improv workshops and activities with promising results that include helping to create collaborative classroom environments, stimulating innovative thinking in requirements engineering, enhancing the effectiveness of brainstorming, and teaching skills necessary for product design [8, 9, 10, 12].

We have identified four improvisation principles from which we believe a robot could derive behaviors to build group psychological safety: (1) agreement, (2) moving the scene forward, (3) celebrating failure, and (4) supporting team members. We discuss each of these and the implications the inspired robot behaviors may likely have on group collaborative behavior.

Agreement

One of the most common 'rules' in improv is "do not block". A block refers to any action that obstructs the ideas, moves, or offers of other players. For example, if John says, "nice day, isn't it?" to Alice, Alice could block John by responding with, "Oh, I don't think so." [11]. Johnstone further explains that when watching good improvisers, it seems as if everything is pre-scripted due to the fact that they accept (and do not block) all offers made during the performance [11]. Social robots within teams should perhaps speak up in agreement, whenever they agree with the speaker. This will likely build the morale and confidence of those the robot is agreeing with and spur on others to agree as well.

Although blocking seems like a behavior to always prohibit, Sonalkar et al. suggest that although some team members may block one another, this does not prohibit them from collaborative success [16]. In order for these teams to reach their goal of concept generation, however, they had to resolve the blocks that occurred in their creative process [16]. As social robots may detect blocks from other human members of the team, it may be an effective strategy to suggest solutions or ask a question to try and get the team to resolve such blocks.

Moving the scene forward

Good improvisers not only agree with the offers made by their team members, but also advance the scene by adding additional content without waffling (speaking without driving the plot line forward) [14]. One classic improv game named "Yes, and..." has actors repeatedly make statements building on each other by starting each sentence with 'Yes, and' to build the skills of both agreement and progressing the narrative [11]. Social robots, upon detecting that their team is stuck or not making much progress could encourage the team to continue pursuing the goal, asking questions to

spur on new ideas and discussion, or by proposing new content on its own to maintain forward momentum in the group.

Celebrating failure

In order to promote risk-taking that may possibly lead to achieving great success, improvisers must develop a comfortability with failure. Improvisers worried about failure are slowed down by additional conscious thought, where improvisers who adopt a 'playful' mindset are able to make decisions more rapidly and fluidly [11]. Social robots could easily respond when someone reports a failure by supporting their admission of failure, validating the person who made the error, and encouraging the team to learn from the mistake made. A social robot could also perhaps make mistakes and model admitting mistakes and failures when they occur.

Supporting team members

One important virtue of improv is that of supporting one's fellow improvisers [14]. Support requires an improviser to be attentively listening to fellow improvisers and willing to forfeit their comfort and ideas for the sake of their peers and the show at large. Social robots should be programmed to have appropriate nonverbal behaviors, showing attention to teammates fairly and in a way that communicates that the robot is actively listening to what is being said. Social robots are also uniquely positioned in a human-robot team as the only agents without 'selfish' desires, thus, robots can and should always act in the best interest of the group.

Conclusion

Social robots are increasingly becoming members of human-robot teams and must be equipped with behaviors to contribute effectively to team success. Human teams are most productive when they are psychologically safe, i.e. charac-

terized by a trusting, caring, respectful environment where team members feel safe taking interpersonal risks. We seek to design social robot behavior to promote psychological safety in human-robot teams by applying research on vulnerable disclosure, illness support groups, and improvisational theater. We believe that robots that make vulnerable disclosures, invite emotional expression, promote agreement, move the team forward, celebrate failure, and support fellow team members will help human-robot teams develop psychologically safe environments leading to optimal team performance.

REFERENCES

1. Sigal G Barsade. 2002. The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly* 47, 4 (2002), 644–675.
2. Matthew J Cordova, Janine Giese-Davis, Mitch Golant, Carol Kronnenwetter, Vickie Chang, Sarah McFarlin, and David Spiegel. 2003. Mood disturbance in community cancer support groups: The role of emotional suppression and fighting spirit. *Journal of psychosomatic research* 55, 5 (2003), 461–467.
3. Paul C Cozby. 1973. Self-disclosure: a literature review. *Psychological bulletin* 79, 2 (1973), 73.
4. Kathryn P Davison, James W Pennebaker, and Sally S Dickerson. 2000. Who talks? The social psychology of illness support groups. *American Psychologist* 55, 2 (2000), 205.
5. Ian J Deary. 2000. *Looking down on human intelligence: From psychometrics to the brain*. Vol. 34. Oxford University Press.
6. Amy Edmondson. 1999. Psychological safety and learning behavior in work teams. *Administrative science quarterly* 44, 2 (1999), 350–383.
7. Howard Gardner. 2011. *Frames of mind: The theory of multiple intelligences*. Basic books.
8. Elizabeth Gerber. 2007. Improvisation principles and techniques for design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1069–1072.
9. Elizabeth Gerber. 2009. Using improvisation to enhance the effectiveness of brainstorming. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 97–104.
10. Julie Sheldon Huffaker and Ellen West. 2005. Enhancing learning in the business classroom: An adventure with improv theater techniques. *Journal of Management Education* 29, 6 (2005), 852–869.
11. Keith Johnstone. 2012. *Impro: Improvisation and the theatre*. Routledge.
12. Martin Mahaux and Neil Maiden. 2008. Theater improvisers know the requirements game. *IEEE software* 25, 5 (2008), 68.
13. Nikolas Martelaro, Victoria C Nneji, Wendy Ju, and Pamela Hinds. 2016. Tell Me More: Designing HRI to encourage more trust, disclosure, and companionship. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 181–188.
14. Tom Salinsky and Deborah Frances-White. 2013. *The improv handbook: the ultimate guide to improvising in comedy, theatre, and beyond*. A&C Black.

15. John Schaubroeck, Simon SK Lam, and Sandra E Cha. 2007. Embracing transformational leadership: team values and the impact of leader behavior on team performance. *Journal of applied psychology* 92, 4 (2007), 1020.
16. Neeraj Sonalkar, Ade Mabogunje, and Larry Leifer. 2013. Developing a visual representation to characterize moment-to-moment concept generation in design teams. *International Journal of Design Creativity and Innovation* 1, 2 (2013), 93–108.
17. Jane Ussher, Laura Kirsten, Phyllis Butow, and Mirjana Sandoval. 2006. What do cancer support groups provide which other supportive relationships do not? The experience of peer support groups for people with cancer. *Social science & medicine* 62, 10 (2006), 2565–2576.
18. Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. 2010. Evidence for a collective intelligence factor in the performance of human groups. *science* 330, 6004 (2010), 686–688.