Validating the Use of Assistive Models in Social Robots

AYURG | Natural Sciences and Engineering (NSE) | *Tags: Computational/Mathematical Modeling; Design/Build*

This cover page is meant to focus your reading of the sample proposal, summarizing important aspects of proposal writing that the author did well or could have improved. ***Review the following sections before reading the sample****. The proposal is also annotated throughout to highlight key elements of the proposal's structure and content.*

Proposal Strengths	Areas for Improvement
While the methodology and subject of this proposal are extremely technical in nature, the researcher does a great job of writing for a general audience and reviewing literature from relevant fields.	The proposal is missing a distinct "preparation" section, which is typically included at the end of the proposal. This section outlines the classes, work experience, previous research experience, and technical experience relevant to the project. Additionally, the section includes 1-2 sentences showing how this work connects to the researcher's future goals. It is important to include such a section in your proposal.
The recruitment strategy and eligibility criteria for human subject participants are clearly defined and justified.	The specific aims of the project would be better situated within the proposal if the researcher added a sentence or two connecting the aims and methods back to an explicit research question.
	The student should give a clear picture of what the data analysis will look like for their project.
	Whether broadly early in the proposal or more detailed at the end of the background section, the inclusion of an explicit research question (ending with a question mark) or statement of objectives would strengthen the proposal.

Other Key Features to Take Note Of
All Academic Year URGs require a budget. There is no required format; however, we do provide a template on our website. The scope of the proposal should focus on what the funding covers, and there are some funding limitations to your request. For example, funding of "durable equipment," like the Cozmo robot, may not exceed $100.
If you are proposing a study that is a continuation from previous work, you must focus the proposal on how your new study will build upon what was established by your work and others' work.
AYURG proposals require IRB submission at time of application; you will need to enter the IRB number during the application process. Please see the Human Subjects Research section of our website for additional details.

The focus of the project is clear within the first paragraph.

A statement including your chosen methods and your research question should be included in the first paragraph.

Gaze cues, defined as eye movement patterns, are important in human social interactions for conversation management and mutual understanding. They are necessary for communication and feedback; most people inherently give off and interpret such cues seamlessly. For example, looking at a particular object signals attention towards that object. A sizeable body of work has been done researching gaze cues in human-human interaction, andwork is emerging that focuses on these patterns in human-robot interactions. Gaze cues in the context of a human-robot collaboration are important considering the rising popularity of using robots in assistance tasks. For example, assisting older adults with day to day tasks. A seamless user interaction with an assistive robot could increase positive perceptions of the robot as well as usability. This work seeks to address this need by validating the use of gaze cues for use in social robotics.

Background draws on work from other relevant fields with clear logic flow.

Justifies the importance of looking at gaze.

Recently there is substantial interest in socially assistive agents [1], like robots, that provide assistance in a social manner, such as companionship or task-related feedback. These robots are used in a variety of applications, some of which include physical therapy or rehabilitation [2, 3], companionship for older adults [4], helping children with autism [5, 6, 7] and medication management for older adults [8]. Because more people are utilizing such agents, it is important to understand how we can build better systems for these agents by using human social cues, specifically gaze. Thus, it is important to consider previous work which incorporatesgaze into the human-robot interaction, either through robots producing gaze cues or robots programmed to interpret meaning from their human counterpart's gaze.

Robots which have been programmed to produce gaze cues by using animated faces on a screen show that the fluidity of the interaction as well subject engagement increases when robots utilize gaze [11]. Similar gaze-producing robots are more successful in turn taking within conversations and are perceived as more thoughtful by the subjects [10]. Work done on robots that interpret gaze cues shows similar results; agents that wait for a gaze cues before handing over an object to their human subject are much more successful at object handover than agents which do not use the human gaze cue [12]. Lastly, robots developed to do both the interpreting and producing of gaze cues showed more fluid interactions and were understood better by the human subjects [13].

Typically, URGs focus on the independent contribution of the student. For this reason, limit the use of "we" language.

There is a connection missing on why this particular task would be a useful model for gaze.

Evaluating a model of gaze for predicting if assistance is needed can have important implications within this field because it would improve the fluidity and seamlessness of this interaction. Additionally, it would allow the robot to better support the human subject in the task by understanding when the subject needs this assistance, and when they do not. Through URAP, I worked alongside Professor Wilson to develop and publish work in this field. We show that models of gaze are good predictors for whether a person needs assistance in one particularhuman robot collaboration task [9]. A model of eye gaze utilizes the direction and pattern of where someone is looking to determine if a person needs assistance in a task. In a medication sorting task, participants were asked to use the instruction on the pill bottles to sort the medication into a weekly grid. The assistive agent, a Nao robot (Fig. 1) provided assistance when the subject made a mistake. The Nao is an anthropomorphic robot with a face and is capable of speech and movement. By analyzing this interaction, we built computational models which successfully predicted instances where the user needed assistance in the task. In order to generalize these models to other tasks and other robots I plan to conduct a study as part of my Independent Study courses my senior year. Implications of validating these models across tasks and agents include the development of systems which can readily and accurately respondto a user when they need help.

The background section should leave the reader with an explicit understanding of the research question.

Justifies two populations used for study.

Forty participants will be recruited from the Northwestern campus and compensated $10 for their time. Half of these participants will be younger adults (18-30 years old) and half will be older adults (50+ years old). The justification for this split is that while we are not testing population differences, feedback from both younger and older individuals will allow us to see if user perception or other ideas surrounding this agent which need to be considered as the use of social robotics expands to broader age groups. While older individuals are the target populationfor

Explicit recruitment strategy and guidelines are stated.

socially assistive robots, our previous study was done on younger adults, so it is important to generalize these models to both groups. Participants will be recruited from campus. Flyers will be posted in campus buildings, with the largest focus of flyers posted at Norris, as there are often older adults from Evanston who come to Norris for mini courses or events. Only individuals without any cognitive issues will be considered, as they will need to knowingly and actively consent to this study. Audio/video recording is a necessary part of my study, and this will be made explicit to all participants so that they can actively consent to the recording. They will be recorded during a 30 to 45-minute-long study. They will be asked to complete two puzzletasks with the help of a small robot on the desk in front of them. These two tasks will be counterbalanced for order effects and are described below.

Justifies the type of robot being used.

The robot which I plan to use in this study is the Cozmo robot. This robot is a small truck, with a display for eyes (Fig. 2). Cozmo is different from the Nao robot in that it is not humanoid and lacks some of the more advanced features like face and hands. In order to see if these models work across different conditions, the Cozmo is ideal because of these differences.
However, it is important to note that while the Cozmo is different in many regards, it still has eyes, which is vital for the expression of eye gaze patterns, as a robot without eyes would be unlikely to illicit social eye contact from human subjects. In order to test the gaze models againstother tasks, the participants will take part in two different tasks with the goal of recording their eye gaze for input to the gaze models. The first task is the Towers of Hanoi task (Fig. 3) where participants will be asked to move all disks from one peg to another peg, maintaining the tower order. The rules of this task are that one disk may be moved at a time, and larger disks may not be placed on smaller disks. The second task is the Lego Building Task (Fig. 4), where participants will be given a semi constructed Lego dinosaur, and expected to construct the rest of the pieces given an image of the completed dinosaur as a reference. In both tasks, the robot will provide verbal assistance when the user makes a mistake. Such assistance can take the form of "try looking at peg 1 again" or "try using the blue Lego piece" and the content of these verbalizations will be the same across participants.

It would be beneficial for the researcher to elaborate on data collection. For example, where will the camera be placed for the eye tracking software to be used during analysis?

Justifies use of independent annotators.

Analysis will be done on the videos recorded from these interactions. There are two categories of data that are vital to this study (Fig 5). First, eye tracking software will be used to derive gaze from the videos. This data will be fed into the models which will produce Yes/No output for gaze events in the videos. This output consists of predictions for if the person needs assistance. E.g., the models may output "Yes" for event 5 in the video, which means the person needs assistance at that moment. This is the first type of data, which I will call the Predicted Outcome. The second type of data I will call the Ground Truth. Because people are innately good at understanding social gaze cues, Ground Truth data is important to compare the computational model to a human baseline. This data will be collected from 2-3 independent annotators who will watch all the videos and select times where the person needs assistance inthe task. I will provide the training and instruction materials for all the annotators, they will be told to annotate moments of confusion, hesitation, frustration, and other markers that show a person is struggling and needs assistance in the task. It is necessary to have multiple annotators because we are collecting subjective data and we want it to be as accurate and unbiased as possible. Majority agreeance across annotators will be used, as we have done it inour previous work. This data will be the Ground Truth. Then, we will compare the Predicted Outcome with the Ground Truth to measure the accuracy, precision, and recall of the gaze models. For example, if the Predicted Outcome for an event is "Yes", and the Ground Truth for that instance in the video is also "Yes", then our model has successfully identified an instance where the subject needs assistance and we have a true positive, this data will be used to run statistical tests, such as t-tests and F score. In this way, I will be able to statistically measure how well models of gaze can be extended to scenarios other than the medication sorting task with a Nao robot.

There is no clearly defined paragraph describing student preparation. It provides the

It would be good to include a description of how the model analyzes the data.

References

1. David Feil-Seifer and Maya J. Mataric. 2005. Defining socially assistive robotics. In *Rehabilitation Robotics, 2005. 9th International Conference of Rehabilitation Robotics.*456-468.
2. Jonathan Casas, Nathalia Cespedes Gomes, Emmanuel Senft, Bahar Irfan, Luisa F Gutierrez, Monica Rinco, Marcela Munera, Tony Belpaeme, and Carlos A Cifuentes. 2018. Architecture for Social Assistive Robot in Cardiac Rehabilitation. In *2018 IEEE 2ndColombian Conference on Robotics and Automation (CCRA).* IEEE, 1-6.
3. Maja J Mataric, Jon Eriksson, David J Fiel-Seifer, and Carolee J Winstein. 2007. Sociallyassistive robotics for post-stroke rehabilitation. *Journal of NeuroEngineering and Rehabilitation* 4, 5 (2007).
4. Kazuyoshi Wada and Takanori Shibata. 2007. Robot therapy in a care house – Changeof relationship among the residents and seal robot during a 2-month long study. *IEEE Transations on Robotics* 23, 5 (2007), 972-980.
5. David Feil-Seifer and Maja J Mataric. 2009. Toward socially assistive robotics for augmenting interventions for children with autism spectrum disorders. In *Experimental Robotics*, Oussama Khatib, Vijay Kumar, and George J. Pappas (Eds.). Springer BerlinHeidelberg, 201-210.
6. Jillian Greczek, Edward Kaszubski, Amin Atrash, and Maja Mataric. 2014. Graded cueing feedback in robot-mediated imitation practice for children with autism spectrumdisorders. In *The 23rd IEEE International Symposium on Social and Human InteractiveCommunication.* 561-566.
7. Hideki Kozima, Marek P. Michalowski, and Cocoro Nakagawa. 2009. Keepon: A Playfulrobot for research, therapy, and entertainment. *International Journal of Social Robotics* 1, 1 (2009), 3-18.
8. Jason R. Wilson, Nah Young Lee, Annie Saechao, Linda Tickle-Degnen, and MatthiasSchuetz. 2018. Supporting Human Autonomy in a Robot-Assisted Medication Sorting Task. *International Journal of Social Robotics* (February 2018).
9. Ulyana Kurylo and Jason R. Wilson. 2019. Using human eye gaze patterns as indicatorsof need for assistance from a socially assistive robot. In proceedings of *9th International Conference on Social Robotics.*
10. Sean Andrist, Xiang Zhi Tan, Michael Gleicher, and Bilge Mutlu. 2014. Conversational gaze aversion for humanlike robots. In Proceedings of the 2014 ACM/IEEE internationalconference on Human-robot interaction. ACM, 25–32
11. Allison Bruce, Illah Nourbakhsh, and Reid Simmons. 2002. The role of expressivenessand attention in human-robot interaction. In Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292), Vol. 4. IEEE, 4138– 4142.
12. Grigore, E. C., Eder, K., Pipe, A. G., Melhuish, C., & Leonards, U. (2013, November). Joint action understanding improves robot-to-human object handover. In *2013 IEEE/RSJInternational Conference on Intelligent Robots and Systems*(pp. 4622-4629). IEEE.
13. Andrist, S., Gleicher, M., & Mutlu, B. (2017, May). Looking coordinated: Bidirectional gaze mechanisms for collaborative interaction with virtual characters. In *Proceedings ofthe 2017 CHI conference on human factors in computing systems* (pp. 2571-2582). ACM.

Figures

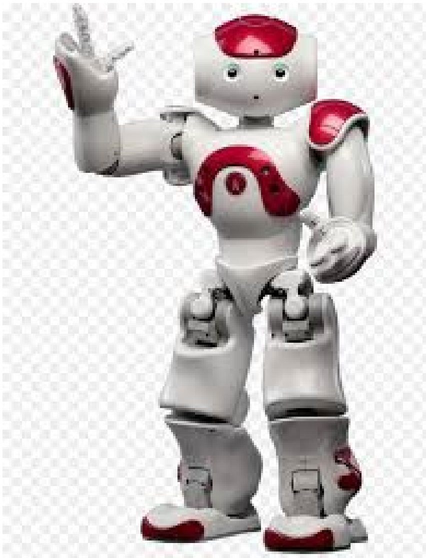

Figure 1. Image of Nao humanoid robot.

Figure 2. Image of Cozmo Robot

Figure 3. Towers of Hanoi game.

Figure 4. Lego Building task.

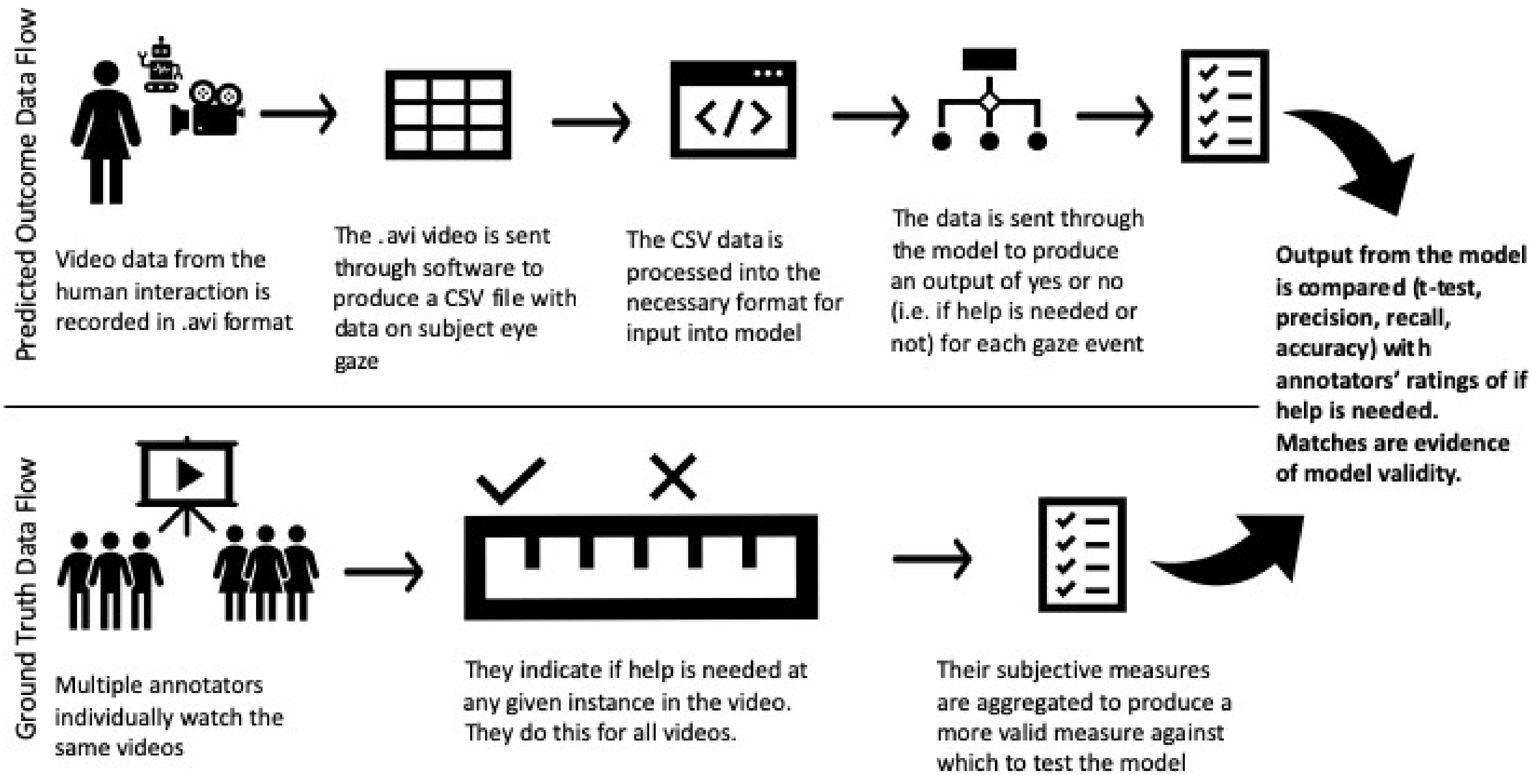

Figure 5. Data flows for Predicted Outcome and Ground Truth

Budget

A. Research-Related Expenses (Data Collection; Analysis)

TYPE	COST	NOTES
1. Consumable Materials		
2. Non-Consumable Materials	$60.00	Games for use in task
3. Equipment/Durable Goods	$100.00	Cozmo robot (from eBay)
4. Research Subject Compensation	$400.00	$10 X 40 participants
5. Fees		
6. Transcription Services		
7. Tuition/Mandatory Fees		
8. Instructional Materials		
9. Living Expenses		
10. Other	$400.00	2 annotators each at $200

B. Travel-Related ExpensesN/A

C. International-Related ExpensesN/A

TOTAL EXPENSES

TYPE	COST	NOTES
Total Research Expenses (A)	$960.00	
Total Travel Expenses (B)	$0.00	
Total International Expenses (C)	$0.00	
TOTAL EXPENSES	**$960.00**	

D. POTENTIAL FUNDING

SOURCE	AMOUNT	NOTES
possible URG funding	$960.00	
TOTAL FUNDING	**$960.00**	