Vicbots as a Lens for Social Robots: This project suggests that 'vicbots' may pose a risk to interhuman empathy and may need to be restricted from the commercial market unless further research demonstrates their effects are generally harmless

❖ **Social Robot Issues:**
- Current lack of consensus about whether social robots have the ability to affect interhuman interactions
- Social robots are unlike any other possessions to date [1, 2]: *they introduce an embodied, subordinate intelligent agent into one's own environment*
- Human robot interaction (HRI) literature indicates that humans tend to view social robots as social partners [1, 10, 11, 12, 22]: *highlighting our perception of them as socially significant*
- Manufacturers strive to make social robots increasingly emotionally appealing, relatable, and trustworthy [3, 18, 19]: *our perception of them as socially significant will likely increase and become more complex*
- It is important to determine whether what we <u>know</u> about social robots and what we <u>feel</u> about social robots are in alignment [12, 22]: *the literature indicates a disconnect*
- Because of the above issues, and because social robots are not yet a mainstream component of society, now is the time to better understand their potential to impact interhuman relationships [14]: *to what extent might they enhance our relationships and to what extent might they pose a risk to human relationships?*

❖ **Vicbot Challenges and Issues:**
- Vicbots, <u>humanoid autonomous social robots capable of verbally pleading for the cessation of or protesting harm</u>, emerged in tandem with some of the first examples of social robots [9, 15]: *they will likely continue to exist in varying forms*
- Vicbots may pose a public risk: *whether the user becomes desensitized to protests of harm from an entity viewed as a social partner but thought of as a machine may in turn cause the user to become desensitized to similar protests from other social partners – their fellow humans*
- A verbal articulation of suffering is a uniquely human attribute of distress; facial expressions, non-verbal utterances, or body language are typically human attributes [17]: *there may be some instances in which it is better for the social bot to respond unemotionally and impersonally as replying in such a manner would discourage 'bad' behavior while reducing the likelihood of a decline in empathy to victim-like responses*
- Vicbots could be the result of social bots with response customization options or as innocuous as social bots pre-programmed to help teach children not to hit by saying "Ouch, that hurt, please don't hit me", ex.: *many users are likely to find this distressing given their attachment to their social bot while others may enjoy/become excited by such responses, potentially exacerbating underlying issues or a predisposition to violence*
- Humans view social bots as social partners but the lack of legal standards guiding interactions with social bots differentiates them from other social partners [9, 10, 12, 22]: *lack of consequence for behavior towards social robots that plead for the cessation of harm may subsequently decrease human empathy for such articulations in human social partners*
- Some studies suggest social robots are the ideal targets for abusive behavior, as embodied, subordinate, and intelligent agents in one's own home [1, 2]: *thus increasing the likelihood of humans interacting with vicbots*
- Humans may abuse any social robot but not all abused social bots are necessarily vicbots, if incapable of pleading for the cessation of or protesting harm: *moderating the way the social bot responds, rather than the behavior of the human, would reduce likelihood of a decline in empathy to victim-like responses*
- Lack of scholarly consensus on the influence of violent media on human aggression [13]: *the effects of interacting in a violent manner in one's own environment with a subordinate social robot capable of pleading for the cessation of harm also lacks consensus*
- Because social robots are not yet a mainstream component of society, neither are vicbots [14]: *now is the time to establish regulatory processes for mitigating potential harm by determining whether vicbots are correlated with a decline in interhuman empathy and whether such a decline would pose a public risk*

❖ **Recommendations for Vicbot Oversight –**
- Unless further research focused on embodiment, environment, and vicbot responses in a humanoid bot demonstrates that vicbots are generally harmless and do not pose a public risk by reducing interhuman empathy in their users, this project recommends a regulatory approach in alignment with that proposed in the literature [6, 7, 20, 21, 25, 27], ex.: *a social robots certification process, requiring the absence of verbal vicbot response capabilities, where certified manufacturers face limited tort liability*
- However, the literature does suggest that vicbots may be appropriate in rehabilitative contexts, or in prevention of human abuse [9, 15]: *medical professional's prescription/recommendation would be necessary, if further research finds vicbots detrimental to interhuman empathy and the above certification process is in place*
- Utilize existing scales for human perception of humanoid social robots (such as PERNOD) and incorporate findings from further research on vicbots [5, 16, 17, 24, 26]: *integrate cognitive results to yield an empirical scale combining qualitative and quantitative results to help guide policy makers in determining liability based on level of risk inherent in manufacturer design*
- If vicbots are found to be generally harmless regarding their impacts on interhuman empathy, no further vicbot regulatory steps required

* References on
following slide*