

# Evaluating Social Robots

Insights from the Diagnosis of Autism  
Spectrum Disorders

**Bill Smart**

Willow Garage & Washington University  
wds@willowgarage.com



# Main Idea

People will evaluate human-robot interactions  
in the way they evaluate human-human  
interactions.

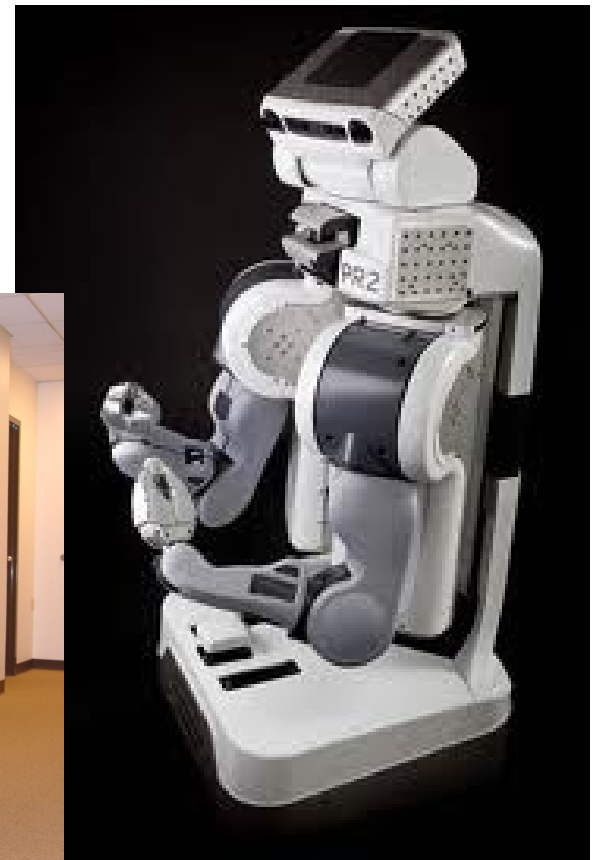
So should we!

# Main Idea

- Specifically, the public
  - are interested in interacting with the general public
  - have no technical background (or interest) in robots
  - don't care about our typical metrics (eg for SLAM)
  - do care about “quality” of the social aspects of the interaction
- Social interaction metrics are in addition to task-specific interaction
  - assuming interaction isn't the task
  - metrics might be in opposition

# Assumptions

- Somewhat “creature-like” robot
  - “head” with “eyes”
  - probably with arms
  - probably mobile



# Measuring Interaction Quality

- If people compare human-robot interactions to human-human interactions, we should measure them in the same way we measure human-human interactions.
- One way of doing this is to draw inspiration from the diagnosis of Autism Spectrum Disorders (ASDs).

# Autism Spectrum Disorders

- Complex developmental disability
  - manifests as deficits in social skills
    - verbal and non-verbal communication
    - social interactions
    - leisure and play activities
  - varying levels of severity, from mild to profound
- Unfortunately, there is no set diagnosis method
  - but there are guidelines

# Diagnosing ASDs

- The Diagnostic and Statistical Manual, 4<sup>th</sup> edition (DSM-IV) defines a number of markers for diagnosis of ASD
  - need at least six symptoms for a positive diagnosis
  - not all apply to our robots
- Markers can suggest how to evaluate our robots
  - and also give us a roadmap for what to implement on our robots, to make them appear more social

# Category 1: Social Interaction

Qualitative impairment in social interaction, manifest by at least two of the following:

- Marked impairment in the use of multiple non-verbal behaviors, such as eye-to-eye gaze, facial expression, body postures and gestures, to regulate social interaction.
- Failure to develop peer relationships appropriate to developmental level.
- Lack of spontaneous seeking to share enjoyment, interests, or achievements with other people.
- Lack of social or emotional reciprocity.



# Category 2: Communication

Qualitative impairment in communication, as manifest by at least one of the following:

- Delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime).
- In individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others.
- Stereotyped and repetitive use of language, or idiosyncratic language.
- Lack of varied, spontaneous make-believe, or social imitative play appropriate to developmental level.

# Category 3: Behavior

Restrictive, repetitive, and stereotypical patterns of behavior, interests, and activities, as manifested by at least one of the following:

- Encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus.
- Apparently inflexible adherence to specific non-functional routines or rituals.
- Stereotyped and repetitive motor mannerisms (e.g. hand or finger flapping or twisting, or whole-body movements).
- Persistent pre-occupation with parts of objects.

# So, How Does the Apply to Our Robots?

- Our robots don't actually have these symptoms
  - since they're not human
- But, the public will tend to interpret behavior in terms of the symptoms
  - since they don't actually know what the robot is doing
  - since they're not aware of the sensors, other than a movable head with cameras
  - all about the *perception* of manifesting the markers

# Social Interaction

Marked impairment in the use of multiple non-verbal behaviors, such as eye-to-eye gaze, facial expression, body postures and gestures, to regulate social interaction.

- robots often do not move, or make eye contact, unless it's needed by the task
- we use extensive secondary actions to shepherd interactions

make eye contact when not  
otherwise using the cameras

# Social Interaction

Lack of social or emotional reciprocity.

- robots don't wave back

detect human social gestures, such as waving, and respond to them

# Communication

In individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others.

- most robots that vocalize do so in declarative statements
- no sense of back-and-forth conversation, scaffolded by secondary actions

avoid back-and-forth speech interaction,  
until we know how to do it right

# Communcation

Stereotyped and repetitive use of language, or idiosyncratic language.

- most vocalizations are from a (very) limited set
- speech synthesis adds to the impression of repetition, since it's (generally) exactly the same each time

add some prosody or variation to speech synthesis, and a richer set of synonymous phrases, or avoid speech altogether

# Behavior

Encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus.

- our robots tend to stare intently at objects, and ignore the rest of the world

move apparent focus when not  
actually capturing data from  
the cameras



# Behavior

Apparently inflexible adherence to specific non-functional routines or rituals.

- even though movements are purposeful, they might look non-functional to a non-technical observer

add variation to repetitive movements,  
or have a set of equivalent movements  
to cycle through

# Behavior

Stereotyped and repetitive motor mannerisms (e.g. hand or finger flapping or twisting, or whole-body movements).

- many grasps start from a safe, known initial pose
- our robots still get stuck in doorways

add variation to repetitive movements,  
or have a set of equivalent movements  
and admit failure early

# Behavior

Persistent pre-occupation with parts of objects.

- door-opening robots are often fixated on handles

add secondary actions when doing  
computation on captured data

# Final Thoughts

- Methods for the diagnosis of ASDs can evaluate social interaction quality
  - no set scheme, or rating scale
  - we (probably) can't benchmark HRI as precisely as we can benchmark manipulation or navigation
  - at the very least, we can count the number of markers our robots have
- They also suggest a roadmap for the development of more social robots
  - or, at least, robots that seem to be more socially present

# Final Final Thoughts

- Improving social interaction might lead to a degradation in task performance
  - need to share resources between task achievement and interaction gestures
  - leads to a multi-dimensional performance metric
  - two dimensions are (probably) coupled
- Or it might improve it
  - through social manipulation