ARTS22 Breakout Session Title:
353-Adapting War Games to Explore Safety Measurement: An Interactive Exercise

Session Contact/Organizers:
Laura Fraade-Blanar, Senior Safety Researcher, Waymo (session contact)
Jack Weast, Intel Fellow, Intel (session contact)

TRB Sponsor/Partner Committees (if any):
AV Safety Metrics Working Group from the TRB Committee on Vehicle-Highway Automation (ACP30)

Session Description
The safety metrics session at ARTS 2021 provided a broad overview of safety metrics, discussing how they met (or failed to meet) the needs of Infrastructure Owner-Operators (IOOs), regulators, researchers, enforcement organizations, the traveling public, and others. Since this session, interest from stakeholders has accelerated around pre-crash measures of safety, also known as leading measure or measures of driving behavior (e.g., near misses, hard braking, etc.). This interest is fueled by the challenges with using outcome measures (also known as lagging measures) such as crashes and the promise of leading measures.

This session will feature a collaborative conversation to increase understanding of how leading measures of driving safety can be used to inform stakeholders on AV safety. Attendees will engage in role play as they apply these metrics to real situations.

Goals/Objectives/Outputs
- Define and disambiguate terms (e.g., leading metrics, lagging metrics, time to collision, etc.)
- Discuss how these terms are currently being used and by whom (e.g. federal regulators, state operators and enforcement, private sector, traveling public, academics, etc.) and challenges around communicating safety measures to said groups.
- Facilitate a guided, but open and interactive activity where attendees can engage with these metrics to get a better understanding a feel for what the variety of metrics proposed actually do and what they do not convey about the safety of AV

Agenda

00:00:00 - 00:10:00 an introduction to leading measures
Laura Fraade-Blanar and Jack Weast

00:10:00 - 00:40:00 detailed presentations on the top leading measures, including strengths, weaknesses, and use cases to-date
Slot 1: Jack Weast
Slot 2: Maria Elli
Slot 3: Robert Heilman
Slot 4: Kevin Gay - Organizational, Operational, & ADS Safety Metrics
Slot 5: Chris Schwarz
Slot 6: Laura Fraade-Blanar

00:40:00 - 00:50:00 an introduction to war gaming and an explanation of the activity
Laura Fraade-Blanar

00:50:00 - 00:55:00 attendees separate into stakeholder groups (e.g., one group of AV companies, one group of local regulators, one group of safety advocates, etc.). Note that attendees will be randomly assigned to their groups. Suggested groups: ADS developer; Commercial fleet manager; municipality/regulator; public advocacy grp; insurance.

00:55:00 - 0:10:00 Within each stakeholder group, attendees debate which metrics best fit their stakeholder goals. Session organizers will circulate and act as advisors if groups have questions about the metrics.
Group leaders, organizers, and safety advisors include:
Laura Fraade-Blanar
Jack Weast
Maria Elli
Robert Heilman
Chris Schwarz

01:10:00 - 01:30:00 dice roll: groups may change their plan, modify their plan, or keep their plan (see appendix for description of dice activity). Groups may continue rolling their dice to encounter new scenarios if desired
Led by Laura Fraade-Blanar

01:30:00 - 02:10:00 each group reports out
Led by Jack Weast, recorders (TBD)

02:10:00 - 02:20:00	attendees change groups such that there are now representatives from each type of stakeholder in each group (e.g., each group has at least one AV company, local regulator, safety advocate, academic, federal regulator, etc.)
02:20:00 - 02:55:00	groups discuss how they wish to proceed in terms of what metrics best fit their new, multistakeholder goals. One-third of groups decide collaboratively. One-third roll for order (e.g., every stakeholder in the group rolls and that is how they order the decision. So, if the AV company rolls a 3, the safety advocate rolls a 1, and the regulator rolls a 4, then the advocate says what they think is best, the AV company can use that or use their own, and then the regulator can use what the AV company says or use their own ideas.) One-third rolls and although all stakeholders state their views, the stakeholder who rolls the highest number is the one who decides what metrics to use. Group leaders, organizers, and safety advisors include: Laura Fraade-Blanar Jack Weast Maria Elli Robert Heilman Chris Schwarz
02:55:00 – 03:25:00	each group reports out Led by Jack Weast, recorders (TBD)
03:25:00 - 3:30:00	Wrap up and next steps Laura Fraade-Blanar

Session Logistics

Date: July 20, 2022 Time: 1:30 pm – 5:00 pm

Appendix

War games are analytical exercises. They simulate real world scenarios at the political, operational, or tactical levels to explore how planning and choices can affect outcomes. Gaming methods have long been used to understand situations of great importance, complexity, and uncertainty. Games provide a pathway to explore the potential consequences of important policy decisions. The most traditional type of gaming, military wargames, are tabletop exercises that can involve players moving tanks or ships around on board to test different strategies or the implications of new weapons or tactics. But the field has evolved to problems far removed from the battlefield, with gaming methods applied to concepts including school budgeting, stakeholder reactions to

cybersecurity crises, and building capabilities at the community level to deal with climate change. These are collaborative games with no winners or losers. Playing can span hours or days. Within the game – essentially a parallel world with policies, rules, incentives, etc.– the players work with one another, making choices within the game's "rules" that reflect their goals and interests.

Because of time constraints we will offer a simple, collaborative game. Individuals will be randomly assigned to stakeholder groups where they will work together to decide which leading metrics are the most usable and aligned with their interests. Dice rolls will present them with new situations (e.g., if they roll a 6, there was a cybersecurity attack on automated trains; if they roll a 3 a new foreign company has entered the domestic market, claiming their level 2 AV is a level 5, etc.) and they will have to decide what these events mean in terms of their preferred metrics (referred to as their plans) – should they change to other metrics? Keep the ones they have but tweak them? This game will allow participants to gain a deep understanding of how leading metrics can and cannot provide assurances of AV safety.

More information can be found here.