Measuring Team Metrics in Cooperative Gameplay: Standardizing Teamwork Analysis

Thomas Walker, Emma Edgar, Eli Whitney
Yvonne Farah, M.S., Michael Dorneich, Ph. D.

Background

- Teamwork is the interdependent actions and processes of individuals working toward a common goal (Salas et al., 2014).
- Cooperative games are an event-based situation where a group of two or more players make a team to work together towards a common goal (Kyniss et al., 2010).
- Behaviors support teamwork with cooperative game mechanics (Farah et al., 2022).
- Behavioral markers are cooperative actions of teammates, performed to meet a goal for the team (Ottatiere et al., 2011), and allows for objective evaluation.
- Cooperative video game footage was annotated to evaluate team performance.

Methods

- Annotate gameplay using a codebook of behavioral markers (Fig. 1) and cooperative features (Fig. 2).
- Iterate codebook as needed to improve ability to annotate all game genres and improve clarity.
- Annotate a select amount of video and regroup to discuss ideas and issues.
- Calculate inter-rater reliability (IRR) among annotators to assure data quality. IRR ratings ranged from 50% to 90% which is considered a moderate to a very strong range of agreement.
- Research team splits into pairs to annotate multiple teams playing cooperative games across different genres, annotating a total of six games with two teams of players per game.
- Compile collected data and analyze to develop themes and relationships.

Results

- The “Heat Map” tables show average behavior-mechanic correlation frequency. Green indicates greater frequency, red represents little or no frequency.
- Figure 5 shows a large variety of correlations whereas Figure 6 is concentrated around one or two correlations.
- The “Star Plot” graphs and associated tables highlight differences within the constructs between two teams playing the same game.
- Figure 7 star plot has a mix of similarities and differences and Figure 8 star plot shows a similar correlation but with different values.
- In SnipperClips (Fig. 7 and Fig. 9), Team 1 performed better than Team 2. In Don’t Starve Together (Fig. 8 and Fig. 10), Team 1 survived more days but also had more deaths than Team 2.
- The behavioral marker frequency graphs in Figures 11 & 12 compare specific behavioral marker frequencies for each team.

Discussion and Limitations

- The behaviors must be objective. Assuming player motivation leads to subjective conclusions. During the data collection process, care was taken to prevent implicit bias.
- The nuanced semantics of the behavior definitions were changed many times in order to accurately describe varied teamwork behaviors. The definitions must be precise to clearly describe a specific behavior yet broad enough to be applicable in any cooperative gaming scenario.
- Standard annotating conventions were written to assist annotators and increase IRR rating.
- Importing the data from the video annotating software and preparing it for spreadsheet analysis was a lengthy process, with room for error.

Future Work

Development of a prototype game using the correlation data found to test team performance. Insights for a prototype:
- Asymmetrical roles in games incite high levels of explicit coordination in teams.
- The story and aesthetics of games encourage team cohesion and interpersonal relationships between teammates.
- Expert teams generally have more implicit coordination and less interpersonal relationships.
- Implicit coordination can be difficult to measure in some games because annotators struggle to identify nonverbal behaviors.

Funded by NSF Grant 1757900