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 (Mynatt 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 (Sottilare 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 split 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.

Behavioral Marker	Description	Qualifications	Examples	Code
Explicit Coordination: Sequencing or Synchronizing	Teammates organizing their movement through verbal communication	Must be verbally expressed Must have clear sequencing or synchronizing	"You go first and I'll follow you." "1,2,3, go!"	EC-S

Fig. 1 Example of a behavior in the codebook, Explicit Coordination - Sequencing

Cooperative Mechanic	Description	Code
Shared Puzzle	A challenge that requires problem solving and is associated with a set of actions in order to be cleared. In some contexts, a shared puzzle can be a series of shared obstacles that have the same end goal.	SP

Fig. 2 Example of a mechanic in the codebook, Shared Puzzle

Fig. 3 Team 17, Ghost Town Games. (2018). *Overcooked 2* [Video game].

Fig. 4 Hazelight Studios. (2021). *It Takes Two* [Video game]. Electronic Arts.

Results

	Team Leadership	Monitoring	Backup	Analysis and Planning	Explicit Coordination	Implicit Coordination	Failure	Failure Recovery	Cohesion	Interpersonal
Aesthetics	0	2.5	0.7	0	0	0	0	0	0	16.3
Boss Fight	2.3	1	1.7	11.1	21.1	12.5	13.1	2	8	1.7
Competitive Challenge	0	0	0.3	1.2	0	0	0	0	1.5	3.2
Common Enemy	0	0	0.5	0.5	1	1.5	0	0	0	0.5
Complementary Obstacle	0.7	2.2	3.3	17.4	19.6	20.5	8	1	1.5	2
Complementary Puzzle	1	0.7	0.7	12.8	10	6.8	11.3	3	2.5	0
Common Risk	0	0.3	0	0	0	0	3.8	0.3	0	1
Individual Obstacle	0	0.3	2.8	3	1.7	0	6.3	1.5	0	0.5
Interactive Object for Fun	0	0.8	0	0	1	0.5	0	0	0	2.5
Story	0	0	0	0	0	0	0	0	0	17.3
Shared Environment	0	5.8	0.5	1.7	2.2	0	0	0	0	2.3
Shared Obstacle	1.5	0.5	1.3	9.6	6.2	23.1	1.7	0.3	0.5	0
Shared Puzzle	2.7	2.2	4.7	26.9	18.1	13.3	15.8	2.8	2.5	0.3

Fig. 5 *It Takes Two* Cooperative Behaviors and Mechanics Correlation Heat Map

	Team Leadership	Monitoring	Backup	Analysis and Planning	Explicit Coordination	Implicit Coordination	Failure	Failure Recovery	Cohesion	Interpersonal
Aesthetics	0	0	0	0	0	0	0	0	0	3
Asymmetric Environment	4.5	30.5	0	26	22	0	1	0	2	0
Complementary Obstsacle	1	38.5	0.5	80.5	121	0	5	0	9	2.5
Complementary Puzzle	19	53	7.5	192	461	0.5	36	2	20.5	8
Common Risk	4	3	0.5	0	9.5	0	0	0	0	0
Shared Puzzle	0	0	0	2.5	3	0	0	0	0	0

Fig. 6 *Keep Talking and Nobody Explodes* Cooperative Behaviors and Mechanics Correlation Heat Map

- 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.

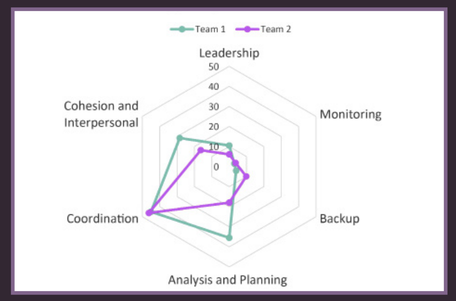

Fig. 7 *SnipperClips* Star Plot of Frequency of Cooperative Behaviors Between Teams

	Number of Puzzles	Avg Puzzle Time
Team 1	30	1:23
Team 2	25	2:38

Fig. 9 *SnipperClips* Team Performance

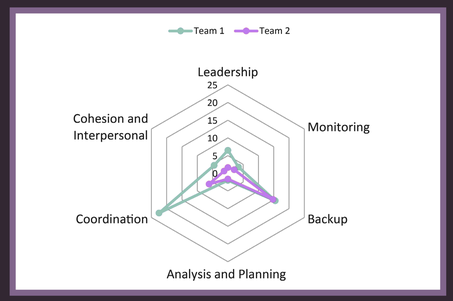

Fig. 8 *Don't Starve Together* Star Plot of Frequency of Cooperative Behaviors Between Teams

	Days Survived	Number of Deaths
Team 1	3	1
Team 2	2	0

Fig. 10 *Don't Starve Together* Team Performance

Discussion and Limitations

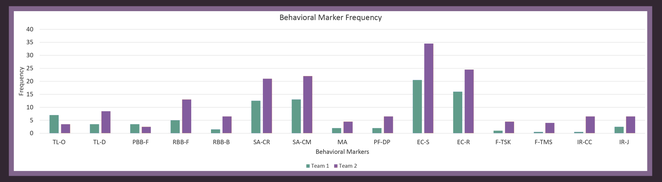

Fig. 11 *Portal 2* Frequency of Cooperative Behaviors Between Teams

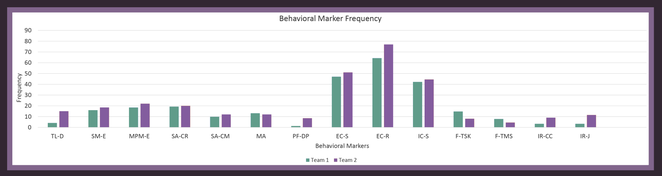

Fig. 12 *Overcooked 2* Frequency of Cooperative Behaviors Between Teams

- 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 explicit coordination than novice teams.
- A high frequency of cooperative behaviors **does not directly correlate** with a high team performance; **quality of behaviors** is also impactful.
- **Implicit coordination can be difficult to measure** in some games because annotators struggle to identify nonverbal behaviors.

Funded by NSF Grant 1757900

IOWA STATE UNIVERSITY

IOWA STATE UNIVERSITY
Virtual Reality Applications Center