The GGWP game: how matchmaking systems standardize the fun out of games.

In their work *Metagaming* (Boluk and LeMieux), Stephanie Boluk and Patrick LeMieux, argue that videogames can be mostly understood as a material which we use to play our own metagames, being more similar to balls and bats than to Baseball, for example. While doing so, they come up with the idea of the *standard metagame*, the “unquestioned ideology structuring twenty-first-century play”, trained and normalized by the videogame industry itself, with its marketing, hardware and software decisions, and, the focus of this paper, design choices. For competitive team based games, there is a noticeable change in the standard metagame before and after the 2010s. With the increasing popularity of esports, the game industry recognized the value in making players be able to compare themselves to each other (and with the professional players), and through the banner of balanced matches and data driven metrics, the search for the better rank became one of the mainstream ways of approaching these games.

Matchmaking has been a thing long before esports and computer games and even the invention of statistically precise systems such as Elo (which I’ll talk more about later). It exists prior to organized competitive sports and big budget boxing matches. We can extrapolate matchmaking as any process of matching two or more people together in the attempt to maximize some value regarding the formed group, we realize that it’s a process that is one fundamental to forming communities in general. But there’s also a deep connection between matchmaking and games and play, it’s one of the many processes in which people can form groups in the search of playing well together (DeKoven), like when choosing who to play with, or forming teams by picking captains and then drafting players to balance the teams (counting on the captains personal judgment of players’ skill).

In online multiplayer games there are additional factors that get into account for what is a good matchmaking experience other than searching for a balanced match, most importantly how long it takes for you to get into a match (Xu et al.) and the latency between the players (Agarwal et al.). These all also vary from game to game depending on the size of the player base and specific restrictions that can happen within a single game, such as game modes or players wanting to play a specific role in a match. The difficulty of balancing these factors, together with the search for a balanced match are what make matchmaking systems such a rich research field, especially with recent attempts to incorporate machine learning methodologies into it (Delalleau et al.).

Overwatch and Team Fortress 2 are two of the most successful games in the sub genres of first person shooters known as "hero shooters". In both games, there is a wide selection of characters with varying gameplay mechanics for players to ll in a selection of different in-game roles. This affords the possibility of a wide range of players with different play styles to be invested in playing such games, since they are not focused on the specific skill of aiming as much as most competitive multiplayer first person shooters are (Webster; Walker). Even though both games share a lot of similarities in how they are played, they are also great poster children of a big change that happened with multiplayer games over the decade that separates the launch of both games. Team Fortress 2 was originally played on dedicated servers, each with their own community rules and gameplay variants. Overwatch is a matchmade game, where players can choose to play between different predefined playlists, but mostly not choosing a community to play with. With centralized matchmaking, the game publisher has taken control over the process of finding people to play, taking away a social aspect of the online game experience, replacing it with pure competition.

In 2007, the year Team Fortress 2 came out, the most common method of playing multiplayer games on PC was through servers and a server browser. The way this worked was that any player had the ability

to have their own multiplayer game session, hosted on their own machine, and by making this information public, their game session showed up in a public server browser in the game's main menu (“Servers”). This quickly evolved to groups of players (usually within a larger community, or clan) with dedicated hardware to host their own shared game session with close to no downtime, these are known as dedicated servers (“GFL - Team Fortress 2”). Most of the public sessions available to play were created by some community of players with their own server rules, both in the form of code of conduct and sometimes in the form of in game rules and modifications (“Servers”). This meant that a public server usually had a constant group of people from the clan playing, but by it’s visible nature anyone could join as long as they followed the rules, with many casual servers not focusing on the competitive aspects of the game, being more of social spaces for the players to hang out and play. This is not to say competitiveness did not exist in these games, there were also servers focused on playing the game in the highest possible level, each implementing their own ranking structures and player selection mechanisms, not unlike one could find in professional and semi professional sports clubs (“How to get into Competitive TF2”).

A lot of what made servers work on PC was the existence of parallel structures available to support the game, especially in the form of forums (“GFL - Team Fortress 2”; “Team Fortress 2 General Discussions.”). Playing multiplayer games in the computer usually involves also being on the internet, consuming paratext on strategies on how to play, watching videos of interesting plays, reading through wikis, thus participating, passively or actively, in a broader gaming community (TL Taylor). Console games, thus, by nature of their separation from an internet browser, had less success with the forming of player clans. The developers took the opportunity to further improve the players experience by removing the server browser and instead created a system with party and skill based matchmaking for players to match with friends or random players of similar skill in the wanted game modes, initially implemented in the Xbox title Halo 2 (Nutt). The introduction of skill rating systems in console games later fed back into computer games, moving most games of the 2010s away from dedicated servers, with even Team Fortress 2 eventually implementing ranked matchmaking (Prescott), and Overwatch having it since close to launch (Skrebels).

A brief explanation on how rating systems works follows: The most well known of player skill rating systems is called Elo, created by Arpad Elo for the game of chess in the 1960s and later improved and adopted by the International Chess Federation in the 1970s (Elo). The basic premise of the system is move away from the previous systems of giving numerical rewards for specific achievements in matches or tournaments and towards a more statistical one, creating a system of prediction of the chance of a player winning in a match based on the ratings of each of the players, and then adapting each players ratings depending on the results. Essentially if a player unlikely to win wins, they will get more points than if it was likely that they would. This makes it so that the ELO rating is a global rating, and should be able to compare players that never played against each other before. The success of this system made it so it was the basis for most systems used in contemporary videogames and esports (Minka et al.), with varied degrees of success.

Skill rating systems serve two main purposes for the players of a game: be able to rank them and know who are the best and be able to properly know the balance of a match between players, thus affording the possibilities of matchmaking in order to search for fair matches for the players (“TrueSkill™ Ranking System”). One could argue that the search for fair matchmaking, especially in an online videogame, is a search for a well played game (DeKoven), a match of players of equal capability, giving their best to win within the bounds of play. It's one way of unifying the multitude of game communities of a same game into a single centralized structure and leaderboard. However, it’s also moving away from the more federated approach of dedicated servers, each with their own rules and

skill distribution. By standardizing how players get to play with each other, the game community also pushes forward their own idea of what is the standard metagame, which players unquestionably follow when just pressing “quick play” or when ranked play is the first alternative showing up in the game menu. It must also be understood that a consequence of the statistical nature of Elo is that in order to increase your rating you not only have to get better yourself, but you have to do so faster than the average player, which likely only happens for the players that are really looking to improve in the game. Therefore some games have different solutions on how to not make this completely transparent to players, and create convoluted points systems on top of the skill one to sort players ranking and incentivize players to play to increase their ranking (“Rank (League of Legends).”).

An additional problem with Elo as a ranking system which researchers try to solve (Minka et al.; Delalleau et al.) is that it is essentially a system for one versus one matches, and it does not translate as well for team based games, which is the scenario most common in modern multiplayer games. It makes sense to think that in a team based game, it is very difficult to separate individual performance from team performance, and thus it's hard to say that having a good individual performance rating actually compares a solo player to a professional athlete playing in a team. It creates a false sense of a non professional player, even if also non casual, being able to compare with professional players by having a common number both should be able to improve at. This also affects how players see their peers during a match, being able to blame bad performance on the low skill rating of a fellow player. It puts every player that plays a ranked game mode in the mindset that they are actually playing the game at a professional level, and raises the stakes of the match to increasing or decreasing their objective numerical skill rating.

The search for metrics to objectify values that are otherwise perceived by personal judgment, as in the earlier example of captains drafting their team in a casual team sport setting is “a cultural pattern of the recent decades” as Z. Muller explains in their book *The Tyranny of Metrics* (Z. Muller). They separate the key components of such fixation into three: *metrics* (the belief that it is desirable to replace judgement with numbers), *accountability* (the belief that publicizing the metrics is proof of responsibility) and *rankings* (the belief that the best way to motivate people is to reward based on measured performance). While the critique is being applied to education and work, it’s noticeable in the patterns explained above of the search of a user metric to quantify skill, even though skill should be treated as a multidimensional, and thus mostly opaque, value (Delalleau et al.). Another point that Z. Muller brings is that idealizing metrics “works even less well in organizations in which employees are oriented to a more idealistic mission”, which is easily understandable to the context of games, themselves a mostly idealistic endeavor, intended primarily as an object of entertainment (Tavinor).

Understanding that skill rating systems such as Elo are evaluators of your performance throughout multiple games, we can borrow Karhulahti discussion on defining a videogame (Karhulahti) as an “artifact that evaluates performance” to say that not only skill based matchmaking change the standard metagame, but they are also a videogame in themselves. The consequence of this is that by nature of them being an arbitrary evaluator, even if based on statistics and attempting to inform player skill, that players can always choose to play the internal game in order to maximize their performance on this other one dimensional metagame. This can result in cheating, toxicity, griefing, for example in certain systems where quitting is not taken into account, some players simply decide to leave the game before the match is over so that it doesn’t affect their rating and once players start doing, so, it reflects back to matchmaking itself, players have to take it back to their own hands and add their own judgment back into account (M. Moeller et al.). It’s a clear example where we can see Goodhart’s law in effect: “When a measure becomes a target, it ceases to be a good measure.” (Strathern).

Matchmaking is a process that happens regardless of there being a standardized matchmaker or not. It's a process that occurs before any game where a group of people decide to play together, and at the moment they decide how to split into multiple teams. It's one of the many meta layers that surrounds games, and thus, having control over it affords control over the form players play a game, the metagame. Making the evaluator transparent creates a whole new game on it's own, that players can also choose how to interact with, but the standard system is that of wanting to get on top of leaderboards, especially with no other means of interaction with the game. With the now archaic system of dedicated server, this part of the metagame was uncontrolled territory and so it was on the player to find their own play community. If players wanted to power game (TL Taylor) they could look for online communities of like minded players, finding specific servers where they could hone their skills and be better at the game. On the other hand, players looking for more casual fun could spend their time interacting with other game systems, even if this some times means joining an unbalanced match and trying to get that one kill out of the best player online. Balanced and fair games don't necessarily mean more fun games (Delalleau et al.), the fun is in playing well, in the form you decided to play with the community you joined.

With the current system of standardized matchmaking, game communities have to necessarily form outside of the game spaces, since players are usually looking for new matches with new players all the time, the system doesn't incentivize group formation, but individual progress. The incentive for individual performance also ends up not learning together and attempting to help each other get better during a match. Even though now it could be interpreted that all players form a single unified pool, different people have different objectives with a game, and there's no single incentive system that can solve that. Players are less of a big community looking to have fun or improve in a game and have become performers for the invisible matchmaking evaluator, imitating and attempting to be like the best players in the world, even if it means taking the fun away from the game itself.

References:

Agarwal, Sharad et al. "Matchmaking for Online Games and Other Latency-Sensitive P2P Systems". SIGCOMM Comput. Commun. Rev. 39. 4(2009): 315–326.

Alman, Josh, and McKay, Dylan, "Theoretical foundations of teammatchmaking",Proceedings of the 16th Conference on AutonomousAgents and MultiAgent Systems, International Foundation for Autonomous Agents and Multiagent Systems, 2017.

Boluk, Stephanie, and Patrick LeMieux. *Metagaming: Playing, Competing, Spectating, Cheating, Trading, Making, and Breaking Videogames*. University of Minnesota Press, 2017.

DeKoven, Bernie. *The Well-Played Game: a Player's Philosophy*. The MIT Press, 2013.

Delalleau, Olivier, et al. "Beyond Skill Rating: Advanced Matchmaking in Ghost Recon Online." *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 4, no. 3, 2012, pp. 167–177., doi:10.1109/tciaig.2012.2188833.

Elo, Arpad E. The Rating of Chessplayers: Past and Present. Ishi Press International, 2008.

"GFL- Team Fortress 2." *Games For Life*, gflclan.com/forums/forum/26-team-fortress-2/.

“How to Get into Competitive TF2.” *Steam Community*, steamcommunity.com/sharedfiles/filedetails/?id=117867854.

Karhulahti, Veli-Matti. “Defining the Videogame.” *Game Studies - Defining the Videogame*, Dec. 2015, gamestudies.org/1502/articles/karhulahti.

M. Moeller, Ryan, et al. “Cheesers, Pullers, and Glitchers: The Rhetoric of Sportsmanship and the Discourse of Online Sports Gamers.” *Game Studies*, Nov. 2009, gamestudies.org/0902/articles/moeller_esplin_conway.

Mingkuan Xu, et al. "Rule Designs for Optimal Online Game Matchmaking.", 2019

Muller, Jerry Z. *The Tyranny of Metrics*. Markus Haller, 2019.

Prescott, Shaun. “Team Fortress 2 Matchmaking and Competitive Mode Is Now Live.” *Pcgamer*, PC Gamer, 8 July 2016, www.pcgamer.com/team-fortress-2-matchmaking-and-competitive-mode-is-now-live/.

Ralf Herbrich, Tom Minka, and Thore Graepel. TrueSkill(TM): A Bayesian Skill Rating System. In Advances in Neural Information Processing Systems, pages 569–576. MIT Press, January 2007. https://www.microsoft.com/en-us/research/publication/trueskilltm-a-bayesian-skillrating-system/.

“Rank (League of Legends).” *League of Legends Wiki*, leagueoflegends.fandom.com/wiki/Rank_(League_of_Legends).

“Servers.” *Servers - Official TF2 Wiki | Official Team Fortress Wiki*, wiki.teamfortress.com/wiki/Servers.

Skrebels, Joe. “Overwatch Ranked Competitive Play Seemingly Delayed.” *IGN*, IGN, 2 May 2017, www.ign.com/articles/2016/06/01/overwatch-ranked-competitive-play-seemingly-delayed.

Strathern, Marilyn. “‘Improving Ratings’: Audit in the British University System.” *European Review*, vol. 5, no. 03, 1997, p. 305., doi:10.1017/s1062798700002660.

Tavinor, G. “Definition of Videogames”. Contemporary Aesthetics, 6 (1). 2008

“Team Fortress 2 General Discussions.” *Steam Community*, steamcommunity.com/app/440/discussions/.

“TrueSkill™ Ranking System.” *Microsoft Research*, 30 Oct. 2020, www.microsoft.com/en-us/research/project/trueskill-ranking-system/.

Walker, Robin. “A Heavy Problem.” *Team Fortress 2*, 1 July 2008, www.teamfortress.com/post.php?id=1670.

Webster, Andrew. “Overwatch and the New Wave of Friendly Online Shooters.” *The Verge*, 31 May 2016, www.theverge.com/2016/5/31/11821148/overwatch-blizzard-multiplayer-first-person-shooter.