

Effects of Social Behaviours in Online Video Games on Team Trust

van Rhenen, Jan-Willem ; Centeio Jorge, Carolina; Matej Hrkalic, Tiffany; Dudzik, Bernd

DOI

[10.1145/3505270.3558316](https://doi.org/10.1145/3505270.3558316)

Publication date

2022

Document Version

Final published version

Published in

CHI PLAY 2022 - Extended Abstracts of the 2022 Annual Symposium on Computer-Human Interaction in Play

Citation (APA)

van Rhenen, J.-W., Centeio Jorge, C., Matej Hrkalic, T., & Dudzik, B. (2022). Effects of Social Behaviours in Online Video Games on Team Trust. In K. Gerling, J. Iacovides, & R. Malaka (Eds.), *CHI PLAY 2022 - Extended Abstracts of the 2022 Annual Symposium on Computer-Human Interaction in Play* (pp. 159-165). ACM. <https://doi.org/10.1145/3505270.3558316>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Effects of Social Behaviours in Online Video Games on Team Trust

Jan-Willem van Rhenen*[†]
Delft University of Technology
Delft, South Holland, The Netherlands
j.w.vanrhenen@student.tudelft.nl

Tiffany Matej Hrkalic
Vrije Universiteit Amsterdam
Amsterdam, North Holland
Delft University of Technology
Delft, South Holland, The Netherlands
t.matejhrkalovic@vu.nl

Carolina Centeio Jorge*
Delft University of Technology
Delft, South Holland, The Netherlands
c.jorge@tudelft.nl

Bernd Dudzik
Delft University of Technology
Delft, South Holland, The Netherlands
B.J.W.Dudzik@tudelft.nl

ABSTRACT

In competitive multiplayer online video games, teamwork is of utmost importance, implying high levels of interdependence between the joint outcomes of players. When engaging in such interdependent interactions, humans rely on trust to facilitate coordination of their individual behaviours. However, online games often take place between teams of strangers, with individual members having little to no information about each other than what they observe throughout the interaction itself. A better understanding of the social behaviours that are used by players to form trust could not only facilitate richer gaming experiences, but could also lead to insights about team interactions. As such, this paper presents a first step towards understanding how and which types of in-game behaviour relate to trust formation. In particular, we investigate a) which in-game behaviour were relevant for trust formation (first part of the study) and b) how they relate to the reported player's trust in their teammates (the second part of the study). The first part consisted of interviews with League of Legends players in order to create a taxonomy of in-game behaviours relevant for trust formation. As for the second part, we ran a small-scale pilot study where participants played the game and then answered a questionnaire to measure their trust in their teammates. Our preliminary results present a taxonomy of in-game behaviours which can be used to annotate the games regarding trust behaviours. Based on the pilot study, the list of behaviours could be extended as to improve the results. These findings can be used to research the role of trust formation in teamwork.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in collaborative and social computing.**

*Both authors contributed equally to this research.

[†]This is the corresponding author

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI PLAY '22 EA, November 2–5, 2022, Bremen, Germany

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9211-2/22/11.

<https://doi.org/10.1145/3505270.3558316>

KEYWORDS

social behaviours; trust formation; perceived trustworthiness; teamwork; taxonomy creation; League of Legends

ACM Reference Format:

Jan-Willem van Rhenen, Carolina Centeio Jorge, Tiffany Matej Hrkalic, and Bernd Dudzik. 2022. Effects of Social Behaviours in Online Video Games on Team Trust. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '22 EA)*, November 2–5, 2022, Bremen, Germany. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3505270.3558316>

1 INTRODUCTION

Competitive multiplayer online video games have become more and more popular and serious over the years. This is reflected in the rapid growth of esports, in which professional gamers compete for monetary rewards, often watched by many spectators [23]. Games in many genres have a professional league or tournament setting. In particular, Multiplayer Online Battle Arenas (MOBAs - a genre of games in which players, each controlling one character, aim to destroy the opposing team's base.), such as *League of Legends (LoL)*, which boasts the largest esports scene and has the biggest active player base [23], gathering people from all around the globe to play in teams, both in amateur and professional ways. In these types of games, teamwork is of utmost importance, imposed by a high degree of interdependence [10] (i.e. the complementary relationships that two or more parties (in this case, the players) rely on to manage dependencies in joint activity) between the players' outcomes. That is, the game can only be won if players efficiently collaborate with each other. One key factor in effective teamwork is trust [25] and it is more important as more risk is involved [11]. In fact, trust is one of the decisive factors for one's decision to engage in an interdependent relationship with another entity [11]. Studying how and why strangers trust each other in MOBA games can contribute to the research and development of video games as well as social sciences and human-computer interaction.

Additionally, engagement is a big reason for people to play video games [26]. Players can become so immersed in playing these games (sometimes bordering near-obsessiveness [20]) that they perceive this to be a high-risk scenario (with the risk being that they lose a game and lose ranking). Furthermore, online video games are quick-paced, with reaction speed as the main physical distinctive characteristic between pro and amateur players [14], and thus,

high-pressure environments. This makes trust highly important in MOBA games. In cooperative video games, such as LoL, trust builds quickly based on player's interactions with the team[7].

In this work, we are interested in studying which social behaviours influence trust in team-oriented settings. In particular, we would like to study 1) which behaviours are related to trust's components (e.g. demonstration of skills shows ability of trustor) and 2) how these components relate to the formation of trust in teamwork. In order to answer these questions, we will use the video game LoL, as a model environment, as it requires high interdependence among teammates. In this game, a player is matched up with four strangers via the internet with whom they have to cooperate to win. As players play the game they may, implicitly or explicitly, make trust assessments through their teammate's in-game behaviours, such as teammates defeating opponents or scoring objectives. In order to study these social behaviours, we create a taxonomy of in-game behaviours, relating to trust as well as trust components. This taxonomy of behaviours can then be, in combination with a questionnaire on trust, used in order to explore trust formation in teamwork. In particular, we hypothesize that the found behaviours can be categorized according to the effect they have on specific components of trust.

This paper presents a taxonomy of social behaviours relating to trust and its components and a pilot study, to draw conclusions for further improvement of the taxonomy, as well as highlight its use for future research. Further sections in this paper will go over the relevant background and related work (Section 2), our method and experiments (Section 3), the results (Section 4) as well as our derived conclusion (Section 5).

2 BACKGROUND AND RELATED WORK

Teamwork has been broadly explored in literature. It can be defined as interrelated and interdependent reasoning, actions and behaviours of each team member that adaptive and dynamically combine to fulfill a shared and valued goal [21, 24]. There are "Big Five" elements of effective teamwork which include team leadership, mutual performance monitoring, team orientation, back-up behaviour and adaptability [25]. These elements are mainly driven by closed-loop communication, shared mental models, and *mutual trust*. Trust in human teams has been recently explored in contexts such as virtual teams [4], sports [8], and university group projects [22]. Trust has also been explored in video games, for example to examine the impacts of trust on teamwork[15], with different types of commitment as mediators. To the best of the authors knowledge, there is no literature regarding how trust forms in video games. In particular, we could not find literature on which behavioural cues or social behaviours lead (or not) to trust formation in video game teams.

Trust is a dyadic attitude or behaviour between a trustor (the one who trusts) and a trustee (the entity being trusted) and it can be defined as *"the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party"* [18]. Although trust is present in our everyday lives, it is not trivial to define, explain or, consequently, estimate it. Mayer et al. [18] propose that the two

main antecedents of trust are the trustor's propensity to trust in general, and the way the trustor perceives trustee's trustworthiness. In particular, the authors propose that the three main dimensions of this trustworthiness are ability, benevolence and integrity (ABI model).

Ability refers to the set of skills and competencies of the trustee, mostly associated with the domain. Benevolence, presents the overall belief that the trustor has about the trustee's intentions, usually associated with some attachment to the trustor, or (in)existing relationship (further explored in e.g., [9, 16]). Finally, integrity involves the set of principles and moral values that the trustor finds acceptable in the trustee. These principles can be such as honesty, truthfulness, sincerity, and ability to keep commitments (reliability/dependability) [19]. Ability is the component which is the most observable, and thus, far more addressed in the literature on trust formation. Furthermore, Mayer and Davis [17] suggested that factors in perceived trustworthiness may be more or less important given specific task characteristics. For some tasks, ability may be more important (e.g. hard and ambiguous tasks), whereas for others integrity may play a bigger role (e.g. moral/social dilemmas).

Measuring trust can be a challenge. In the literature, we can find instruments that follow the ABI model and measure, through questionnaires, propensity to trust [17] and perceived trustworthiness (of teammates) in military teams [1, 2]. Although we can find instruments to measure trust subjectively, it is of our interest to measure trust through observation. So far, studies have measured trust objectively through physiological signals, such as EEG (electroencephalography) and ECG (electrocardiography) and, sometimes, audio and EOG (electro-oculography) [3]. Here, we want to measure trust from social behaviours. The idea of relating social behaviours with trust is not new. For example, Breuer et al. [4] presented a taxonomy of behaviours that affect how teammates perceive each other's trustworthiness, in virtual human teams. Authors present behaviours for ability, benevolence, integrity, predictability and integrity. According to Breuer's study, ability can be observed through behaviours such as how successfully a task is performed (e.g. based on time or score of some kind), how much effort was put to do a task well, by continuously working thoroughly and accurately, and also in how appropriately the tools (such as technology) were used. For benevolence, authors mostly considered behaviours such as task support and granting autonomy to other teammates. Finally, integrity was considered observable through behaviours of confidentiality and ethical values, i.e., behaviours of honesty, commitment, credible communication and actions, privacy, etc [2, 18]. In this paper, we are mainly interested in behaviours associated with ability, benevolence and integrity, basing it on the ABI model.

All in all, literature guides us on how to explore behaviours that show trustworthiness as well as its components and, consequently affect trust in teams of strangers. However, little research has been done on trust formation in teams, as well as objectively research the relation between social behaviours and trust formation in teams.

3 METHOD

We undertook a two-part study. The first part consisted of a series of (semi-structured) interviews (see Sec. 3.2.1), that were undertaken to create a taxonomy (see Sec. 3.2.2) for in-game behaviour annotation

(see Sec. 3.2.3). This was done to infer which social behaviours are displayed during a LoL game and group them into categories relating to the ABI model of trust. The second part consisted of a pilot study in which we asked participants to play a LoL game and answer a questionnaire about trust, including their own propensity to trust and perceived trustworthiness of their teammates. The goal of the second part was to verify whether in-game behaviours, found in the taxonomy, related to ABI dimensions of the trust model. The study was approved by the Human Research Ethics Committee (HREC) of TU Delft.

3.1 Participants

To create a taxonomy for behaviour annotation, we interviewed 6 participants (5 male and 1 female), that had been actively playing the game for at least 1 year. The age groups of the participants were 18-25 (5) and 26-35 (1). All participants were European. The second part of the study consisted of 7 participants, recruited through the League of Legends subreddit.¹ We do not have demographic data regarding this sample.

3.2 Experimental Setup

3.2.1 Interview. In order to link behaviours to trust, a proper taxonomy for these behaviours was created. This was done by interviewing participants for this part of the study individually, in a semi-structured manner. First, participants were asked to fill in a questionnaire regarding their general propensity to trust (as used in Mayer and Davis [17]). Afterwards, they were explained the general goal of the study, as well as given definitions of trust, ability, benevolence and integrity (adapted from ABI [18] for the context of the experiment). This was done to ensure that the participants had knowledge of the types of behaviours relevant to this research. Then, participants were asked to list as many in-game behaviours they could think of as possible. Questions were asked in this unsupervised manner in order to keep interviewer bias as minimal as possible. However, when the participant could not come up with a sufficient number of behaviours, small suggestions were given in order to help them think of more behaviours. Suggestions were either made about specific parts of the game (e.g. *"Can you think of more early/mid/late game behaviours?"*) or about specific parts of the ABI model (e.g. *"Can you think of more behaviours relating to Ability/Benevolence/Integrity?"*). Finally, participants were asked to rate all behaviours on a 5-point Likert scale on their relevance in trust assessment. Only behaviours with a score of 3 or greater were considered for the taxonomy.

3.2.2 Taxonomy creation. A taxonomy was created using card sorting. Card sorting, mainly used to help guide website creation, is an effective tool to categorize an unstructured list into several, user defined, categories [27]. Three researchers were presented with a list of unsorted concepts (in this case, behaviours) and asked to sort these concepts into categories, in a manner that seems most logical to them, and name these categories. Using this method, all behaviours were grouped into a coherent taxonomy.

3.2.3 Game annotation guide. League of Legends (LoL) replay files have a specific format, that only allows for spectating in the game client. Therefore actual annotation proved not feasible. Instead, behaviours were counted and recorded in a separate file. These counts were done by two researchers, both having experience in playing League of Legends. Although we gathered the important behaviours for trust formation from interviews, there were several that were not sufficiently clear for annotation. Therefore, we have filtered and refined this list of behaviours for annotation.

3.2.4 Pilot. To research the relation between the created taxonomy and trust development, as well as improve the taxonomy, we ran a pilot study. In the pilot study, participants were asked to play a game of LoL and fill in a questionnaire regarding both their trust propensity (same questions as in the interview), their trust in the teammates (based on [1]), and send us the corresponding game replay file. The main objective of the pilot study was to illustrate the process of finding out which categories of behaviours are predictors trust dimensions (i.e. ability, benevolence, and integrity).

3.3 Measures

To measure trust, we use 1) a 15-item 5-point scale questionnaire, adapted from Adams et. al. [2] which measures the perceived trustworthiness of the teammates (ability, benevolence, integrity) and 2) an 8-item 5-point scale questionnaire [17] regarding the participant's propensity to trust (personal trait, transverse to others). We aimed at measuring trust mostly with the first questionnaire, but since individual differences in propensity to trust may play a role in how one perceives other's trustworthiness, we used the second questionnaire as a regulator.

4 RESULTS

In this section the results of the two parts of the study are presented. We present the developed taxonomy and game annotation guide resulting from the first study are presented, as well as an overview of the pilot results. Finally a discussion is started about the results.

In the first part of the pilot study, we interviewed 6 people. From these interviews we deduced a list of behaviours (in Table 1), which needed to be sorted in categories using the card sorting mechanism (in Table 2). The card sorting was done by three researchers, one having more experience in trust research, and the other two having more experience in the model environment League of Legends (LoL). The found categories reflected this, as one researcher defined 6 categories related to the ABI model, and the other two researchers defining 9 categories relating to LoL-related behaviours (e.g. Intentional malicious behaviour, Communication, etc.). The categories related to ABI model were: A+ (behaviour that shows positive ability), A- (behaviour that shows negative ability), B+ (behaviour that shows positive benevolence), B- (behaviour that shows negative benevolence), I+ (behaviour that shows positive integrity) and I- (behaviour that shows negative integrity). On the other hand, the categories related to the model environment were: Basic game knowledge, selfish/altruistic, strategy, communication, emotional, toxic, synergy, prejudice, and skills. Therefore, we decided to create a matrix of the two types of categories found, and fill them with the listed behaviours.

¹ https://www.reddit.com/r/leagueoflegends/comments/rg7llk/help_us_research_team_trust_in_lol/

Table 1: List of behaviours collected during the interviews.

E1 Jungler goes to red or blue	E30 How fed someone is	E58 People placing wards
E2 Bot and top help jungler	E31 Follow ups on my plays when I'm fed	E59 Playing to the strengths of your comp
E3 Mid (and bot) can wards	E32 Champ select	E60 Playing around the (fed) carry
E4 Jungler invade	E33 People on wrong role/champ	E61 Deliberately not helping/grieving
E5 Buying the right items	E34 Jungler in appropriate position	E62 Building a strange item
E6 Proper runes	E35 Countergank/gank other side of the map	E63 Repeatable dieing
E7 Counter invades	E36 A teammate dying 1v1	E64 Not listening to cha
E8 Everyone goes to their respective lanes	E37 Jungler contesting when their laners have no priority	E65 Not joining fight when the possibility is there
E9 Pings in general to communicate where to go	E38 People making the right tactical play	E66 Bad mechanics (Messing up an easy play)
E10 Dancing with character animation	E39 Know whether to contest obj in the midgame	E67 Not understanding your role as a champion
E11 The use of emotes	E40 Not playing around the fed carry	E68 ADC playing alone
E12 Chat is also important	E41 Understanding the wincon in the game	E69 Jungler not moving along with the rest
E13 Muting communication channels	E42 Toplaner tp'ing bot	E70 Engaging a fight without there being a reward
E14 Ganks	E43 Sacrificing your own resources to help teammates	E71 Mainly not understanding the game
E15 Communicating ganks through pings	E44 Getting ahead or behind w/ evenly matched laners	E72 Constant aggressive behaviour with pings/chat
E17 Asking for wards	E45 Choosing the wrong tactics because of grudges	E73 Threatening behaviour
E18 Communicating enemy wards	E46 Stigmas w.r.t. certain champions in certain roles	E74 Farming minions that are not 'your minions'
E19 Support is when in trades	E47 Flaming in chat/pinging	E75 Jungler going into enemy jungle when not supposed to
E20 Similar playstyle	E48 People inting	E76 Mid staying in lane too much
E21 Taxing after a gank	E49 Neutral objectives	E77 Support not warding/misplacing wards
E22 Follow up on a dragon fight (or stay in lane to farm)	E50 Jungler misssmiting	E78 Ganking
E23 Time to follow up engage in teamfight	E51 Mismanaging waves	E79 Pushing the wave as a jungler
E24 Late game: Strategic choices	E52 Expressing appreciation for a gank	E80 Ruining the lane state
E25 Where to ward	E53 Invading river, laners need to push for backup	E81 Leashing the jungler
E26 Adequate champion choice	E54 Support endangering themselves for the carries	E82 For teamfights, target the most valuable enemy
E27 Friendliness of enchanter players	E55 Falling behind early but have drakes as a contingency	E83 Jungler present in team objectives (dragons and barons)
E28 Towerdives	E56 Playing to the individual strengths of your teammates	
E29 Go outplay/commits, and not back out but go for it	E57 People buying control wards	

Table 2: All behaviours sorted using the card sorting method. On the left side are the categories relevant to LoL related behaviours and at the top are the categories related to the ABI model components. A+ would stand for a positive effect on ability and B- would stand for a negative effect on benevolence, etc.

	A+	A-	B+	B-	I+	I-
Basic game knowledge: Items/Runes	E1, E2, E5, E6, E8	E62	E81			
Selfish vs. altruistic	E54		E43			E21
Strategy: Wards/Ganking	E3, E4, E7, E14 E22, E24, E25, E60	E37, E67, E69, E71 E75, E76, E77		E68, E74		
Communication	E15, E17, E18		E18, E52, E9, E12	E13		
Emotional behaviour			E10, E11			
Toxicity						E45, E47, E48 E61, E64, E72, E73
Similar playstyle/Synergy	E19, E23				E20, E28	
Prejudice/Champion selection		E33, E46	E27			E26
Skill	E56	E36, E50, E51 E63, E66, E70				

Table 3: Final list of behaviours to be annotated.

G1 Leashing the jungler (<i>E2, E81</i>)	G11 Jungler not moving along with the rest (<i>E69</i>)
G2 Everyone goes to their respective lanes (<i>E1, E8</i>)	G12 Follow up on a dragon/baronfight (<i>E22, E65, E83</i>)
G3 Jungler invade (<i>E4, E75</i>)	G13 Warding/misplacing wards (<i>E25, E77</i>)
G4 Counter invade (<i>E7</i>)	G14 Jungler mismiting (<i>E50</i>)
G5 A teammate dying 1v1 (<i>E36, E44, E63</i>)	G15 ADC dying, playing alone (<i>E68</i>)
G6 Team vision score at 13:00 (<i>E3, E25, E57, E58</i>)	G16 Proper runes (<i>E6</i>)
G7 Ganks (<i>E14, E35, E78</i>)	G17 Building the wrong items (<i>E5, E62</i>)
G8 Taxing after a gank (<i>E21, E74, E79, E80</i>)	G18 Proper summoner spells (<i>E5, E6, E62</i>)
G9 Communicating ganks through pings (<i>E9, E15</i>)	G19 Number of people in a teamfight (<i>E23</i>)
G10 Jungler contesting & laners have no priority (<i>E37, E53</i>)	G20 People inting you (<i>E48</i>)

Table 4: This table contains the results of the annotation of the 7 games.

Category 1	Category 2	ID	Participants							
Basic	B+	G1	Yes	yes	Yes	Yes	Yes	No	Yes	
Basic	A+	G2	Yes	yes	Yes	Yes	Yes	Yes	Yes	
Strategy	A+	G3	0	0	4	3	0	0	1	
Strategy	A+	G4	1	5	1	1	1	2	1	
Skill	A-	G5	2	3	7	1	2	0	4	
	A+	G6	41	30	46	30	36	29	42	
Strategy	B+	G7	4	3	5	2	4	2	5	
Selfish	I-	G8	0	0	2	0	0	0	2	
Communication	A+	G9	0.50	0.33	0	1	0	0	0.4	
Strategy	A-	G10	2	0	2	1	1	0	0	
Strategy	A-	G11	1	0	0	1	0	1	0	
Synergy	A+	G12	4.5	3.33	2	2.8	2.25	3	2.17	
Strategy	I+	G13	0.83	0.63	0.33	0.8	0.25	0	0	
Skill	A-	G14	0	1.00	1	1	0	0	0.17	
Strategy	B-	G15	0	1.00	0	1	0	0	0	
Basic	A-	G16	0	0.00	0	0	1	0	0	
Basic	A-	G17	0	0.00	1	0	1	0	0	
Basic	A-	G18	1	0.00	1	0	0	1	0	
Synergy	A+	G19	0.67	0.25	0.04	0.4	0.06	0.5	0.3	
Toxicity	B-	G20	0	0.00	0	0	0	0	0	
Did they win the game			Yes	No	Yes	Yes	No	Yes	Yes	
Propensity			2.75	2.75	2.75	3.38	2.5	2.38	2.38	
Ability			2	1.2	3	3	4.4	2	2	
Benevolence			1.6	1.2	3.4	4	2.2	2	2	
Integrity			2.2	2.8	3.4	3.4	3.8	2	2.2	
Trust (ABI average)			1.93	1.73	3.27	3.47	3.47	2	2.07	

In order to create the video annotation guide, several behaviours were omitted or refined based on the discussion between annotators. The final list of behaviours to be annotated can be found in Table 3. Several behaviours (from Table 1) were either omitted, or changed, based on discussion that arose between annotators around a sample game. Behaviours were omitted for two major reasons: (1) behaviours were too ambiguous in order to be usable for annotation (e.g. E38 People making the right tactical play - as stated by an interviewee), or (2) behaviours were, with some modification, usable for annotation, but would be too costly to annotate, because of time constraints (e.g. E11 the use of emotes - all players would have to be closely tracked in order to find every emote used, which takes 5x the games' length, about two hours for this event only). Other behaviours were modified, or expanded upon based on edge cases found in the sample game. Also, many behaviours were

combined, because the annotators believed them to be different descriptions of the same event (see grouped events in Table 3).

We distinguish three types of behaviours: binary, counted and normalized. Each replay was annotated accordingly:

- Binary behaviours: some behaviours (G1 and G2) either occur or not throughout the game (count of 0 or 1).
- Counted behaviours: some behaviours happen multiple times during a game and were counted (e.g., G3, G4, G5).
- Normalized counted behaviours: some behaviours needed normalization in order to become relevant. For instance, some behaviours relate to the number of teammates displaying them across situations throughout the game. In that case the average number of teammates in all these situations was taken. (G8, G9, G12, G13, G14, G19)

Finally, we analyzed the results of the annotation (in Table 4). For this, we organized the counts of behaviours per participant and compared them with the participant's propensity, ability, benevolence, and integrity reported scores. Also the results of the questionnaire that can be linked to the replay files are listed below the matrix as propensity, ability, benevolence, integrity and trust (as an average of the three trust components). The current pilot does not in itself give a significant result, however it highlights the application of the created taxonomy as a basis for further research.

4.1 Discussion

The main takeaway from the results of the study were that there were several behaviour that people recognize and interpret as behaviours of trust and trustworthiness in League of Legends. Our results present a taxonomy that can be used for future work to further investigate how these behaviours can be predictors of trust. In the first part of the study we successfully collected qualitative data giving insights on how to further proceed to identify these behaviours automatically. The taxonomy presents these qualitative data in a structured way which may be useful to other domains and games as well. However, there were several behaviours that were difficult to annotate, leading us to discard a great amount of the qualitative data. We believe it may be important to further investigate behaviours that can be better annotated. One particular challenge for studying how the identified behaviours influence trust evaluations is that they comprise different levels of social complexity (e.g., isolated actions like E1, as well as complex events involving actions by multiple parties like E7). As such, it may be important to study behavioural dynamics at the micro level of interactions to understand (and perhaps model) actors' mutual influence each other's judgements and experiences (e.g., similar to decomposing the role of others' nonverbal behaviour [5] and situational context [6] in computational analysis of Face-to-Face settings). Furthermore, these behaviours can be further explored as to be automatically detected by artificial teammates in human-AI teams, leading to a better understanding of human teammates [13] and, consequently, appropriate mutual trust [12].

We also noticed that behaviours related to integrity and benevolence proved more difficult to strictly define for annotation. Unfortunately, the annotatable behaviours in these categories were also those with the fewest counts/behaviours, e.g. "G20 - People inting you", which had 0 counts. Finding other behaviours in these categories, that are more strictly annotatable are a point of great focus in future work. Literature on trust and the ABI model states that benevolence may take more time to develop and integrity depends on principles and values, which cannot always be easily linked to behaviours. This may justify our difficulties and deeper studies into these dimensions. In general, the results of the pilot study need more participants to be analysed. Regarding the trust measures gathered through questionnaire answers: the variation seen among all participants in all dimensions (propensity, ability, benevolence and integrity) seems promising.

5 CONCLUSION AND FUTURE WORK

In this work we have presented a first step towards studying the effect that behaviours have on trust among teammates. In particular,

we chose the online video game League of Legends (LoL) as a tool to investigate this. Our research was divided in two main parts which explored which, and how, in-game behaviours affected trust levels of players (who were strangers) after a game. For the first part, we conducted a small group qualitative study, which resulted in a taxonomy of in-game behaviour that affect trust. This taxonomy was then used to create a game annotation scheme. The second part of the study consisted of annotating participants' games and comparing these annotations with the players' self-reported trust in their teammates. For the second part, we only ran a pilot and analyzed the main takeaways, and reflected on what should be improved for the future work. Our results propose a taxonomy with 9x6 categories of in-game behaviours, which can be used in further studies. After annotating and running a pilot of the second part of the study, we believe, however, the amount of behaviours (not categories) in the taxonomy should be extended to improve the results. As such, in the future, we aim to extend the behaviours that can be annotated, particularly those related to benevolence and integrity, and run the second part of the study with more participants. We aim to analyze the relation between the annotated behaviours and the found levels of trust. These findings can be used for future research regarding teamwork and trust formation, both in human teams and in the emerging field of hybrid (also known as human-AI) teams.

ACKNOWLEDGMENTS

This research was (partially) funded by the TU Delft AI Lab AI*MAN Lab, part of TU Delft AI Initiative, <https://www.tudelft.nl/en/ai/aiman-lab>, and (partially) funded by the Hybrid Intelligence Center, a 10-year program funded by the Dutch Ministry of Education, Culture, and Science through the Netherlands Organisation for Scientific Research, <https://hybrid-intelligence-centre.nl>.

REFERENCES

- [1] Barbara D. Adams, Sonya Waldherr, and J. Sartori. 2008. Trust in Teams Scale, Trust in Leaders Scale: Manual for Administration and Analyses.
- [2] Barbara D. Adams and R. Webb. 2002. Trust in Small Military Teams. *Command and Control Research Program* (2002).
- [3] Ighoyota Ben. Ajenaghughrue, Sónia Sousa, and David Jose Ribeiro Lamas. 2020. Measuring Trust with Psychophysiological Signals: A Systematic Mapping Study of Approaches Used. *Multimodal Technol. Interact.* 4 (2020), 63.
- [4] Christina Breuer, Joachim Hüffmeier, Frederike Hibben, and G. Hertel. 2020. Trust in teams: A taxonomy of perceived trustworthiness factors and risk-taking behaviors in face-to-face and virtual teams. *Human Relations* 73 (2020), 3 – 34.
- [5] Bernd Dudzik, Simon Columbus, Tiffany Matej Hrkalic, Daniel Balliet, and Hayley Hung. 2021. Recognizing Perceived Interdependence in Face-to-Face Negotiations through Multimodal Analysis of Nonverbal Behavior. *Proceedings of the 2021 International Conference on Multimodal Interaction*, 121–130. Issue 1. <https://doi.org/10.1145/3462244.3479935>
- [6] Bernd Dudzik, Michel-Pierre Jansen, Franziska Burger, Frank Kaptein, Joost Broekens, Dirk K.J. Heylen, Hayley Hung, Mark A. Neerincx, and Khiet P. Truong. 2019. Context in Human Emotion Perception for Automatic Affect Detection: A Survey of Audiovisual Databases. *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 206–212. <https://doi.org/10.1109/ACII.2019.8925446>
- [7] Tobias Greitemeyer and Christopher Cox. 2013. There's no "I" in team: Effects of cooperative video games on cooperative behavior. *European Journal of Social Psychology* 43, 3 (2013), 224–228.
- [8] H. Huynh, C. E. Johnson, and Hillary S Wehe. 2019. Humble Coaches and Their Influence on Players and Teams: The Mediating Role of Affect-Based (but Not Cognition-Based) Trust. *Psychological Reports* 123 (2019), 1297 – 1315.
- [9] Theodore Jensen, Yusuf Albayram, Mohammad M.H. Khan, Ross Buck, Emil Coman, and Md A. Al Fahim. 2018. Initial trustworthiness perceptions of a drone system based on performance and process information. In *Proceedings of 6th International Conference on Human-Agent Interaction*.

- [10] M. Johnson. 2014. Coactive Design: Designing Support for Interdependence in Human-Robot Teamwork.
- [11] Matthew Johnson and Jeffrey M. Bradshaw. 2021. Chapter 16 - The role of interdependence in trust. In *Trust in Human-Robot Interaction*, Chang S. Nam and Joseph B. Lyons (Eds.). Academic Press, 379–403. <https://doi.org/10.1016/B978-0-12-819472-0.00016-2>
- [12] Carolina Centeio Jorge, Siddharth Mehrotra, Catholijn M. Jonker, and Myrthe L. Tielman. 2021. Trust should correspond to Trustworthiness: a Formalization of Appropriate Mutual Trust in Human-Agent Teams. In *Proceedings of the 22nd International Workshop on Trust in Agent Societies (TRUST 2021) Co-located with the 20th International Conferences on Autonomous Agents and Multiagent Systems (AAMAS 2021)*, London, UK, May 3–7, 2021 (CEUR Workshop Proceedings, Vol. 3022), Dongxia Wang, Rino Falcone, and Jie Zhang (Eds.). CEUR-WS.org. <http://ceur-ws.org/Vol-3022/paper4.pdf>
- [13] Carolina Centeio Jorge, Myrthe L. Tielman, and Catholijn M. Jonker. 2022. Artificial Trust as a Tool in Human-AI Teams. In *HRI 2022: ACM/IEEE International Conference on Human-Robot Interaction, Sapporo, Hokkaido, Japan, March 7 - 10, 2022*, Daisuke Sakamoto, Astrid Weiss, Laura M. Hiatt, and Masahiro Shiomi (Eds.). IEEE / ACM, 1155–1157. <https://doi.org/10.5555/3523760.3523956>
- [14] Denis Koposov, Maria Semenova, Andrey Somov, Andrey Lange, Anton Stepanov, and Evgeny Burnaev. 2020. Analysis of the Reaction Time of eSports Players through the Gaze Tracking and Personality Trait. In *2020 IEEE 29th International Symposium on Industrial Electronics (ISIE)*. 1560–1565. <https://doi.org/10.1109/ISIE45063.2020.9152422>
- [15] Chun-Chia Lee and Jen-Wei Chang. 2013. Does Trust Promote More Teamwork? Modeling Online Game Players' Teamwork Using Team Experience as a Moderator. *Cyberpsychology, behavior and social networking* 16 11 (2013), 813–9.
- [16] Dong-Jin Lee, Moonkyu Lee, and Jaebeom Suh. 2007. Benevolence in the importer-exporter relationship. *International Marketing Review* (2007).
- [17] Roger C Mayer and James H Davis. 1999. The effect of the performance appraisal system on trust for management: A field quasi-experiment. *Journal of applied psychology* 84, 1 (1999), 123.
- [18] Roger C Mayer, James H Davis, and F David Schoorman. 1995. An Integrative Model of Organizational Trust. *Source: The Academy of Management Review* 20 (1995), 709–734. Issue 3.
- [19] D. Harrison McKnight, Vivek Choudhury, and Charles J. Kacmar. 2002. Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Inf. Syst. Res.* 13, 3 (2002), 334–359. <https://doi.org/10.1287/isre.13.3.334.81>
- [20] Alison McMahan. 2003. Immersion, engagement, and presence: A method for analyzing 3-D video games. *The Video Game Theory Reader* (01 2003), 67–86.
- [21] B. Morgan, Albert S. Glickman, Elizabeth A Woodard, A. S. Blaiwes, and E. Salas. 1986. Measurement of Team Behaviors in a Navy Environment.
- [22] Andrew M. Naber, Stephanie C. Payne, and Sheila Simsarian Webber. 2018. The relative influence of trustor and trustee individual differences on peer assessments of trust. *Personality and Individual Differences* 128 (7 2018), 62–68. <https://doi.org/10.1016/j.paid.2018.02.022>
- [23] Newzoo. 2021. Global Esports & Live Streaming Market Report. (2021).
- [24] E. Salas, T. Dickinson, Sharolyn A. Converse, and S. Tannenbaum. 1992. Toward an understanding of team performance and training.
- [25] E. Salas, Dana E. Sims, and C. Burke. 2005. Is there a “Big Five” in Teamwork? *Small Group Research* 36 (2005), 555 – 599.
- [26] Henrik Schoenau-Fog et al. 2011. The Player Engagement Process-An Exploration of Continuation Desire in Digital Games.. In *DiGRA Conference*. Citeseer.
- [27] Donna Spencer and Todd Warfel. 2004. Card sorting: a definitive guide. *Boxes and arrows* 2, 2004 (2004), 1–23.